



**UiT** The Arctic University of Norway

Faculty of Science and Technology

**Multiple Signal Classification Algorithm:  
A computational microscopy tool for fluorescence  
microscopy**

**Sebastian Andres Acuña Maldonado**

A dissertation for the degree of Philosophiae Doctor

December 2023







UiT THE ARCTIC UNIVERSITY OF NORWAY  
FACULTY OF SCIENCE AND TECHNOLOGY

**Multiple Signal Classification  
Algorithm: a computational  
microscopy tool for fluorescence  
microscopy**

*Sebastian Andres Acuña Maldonado*

A DISSERTATION FOR THE DEGREE OF PHILOSOPHIAE  
DOCTOR

December 2023



# Abstract

Fluorescence Microscopy is still the most preferred tool for studying the cell's inner structures. One of the reasons for this preference is its great selectivity, which allows it to label specific types of structures and visualize them with high contrast. However, its resolution is conventionally limited due to the diffraction of light, which makes the study of cells at scales below 200 nm challenging. To go beyond this limitation, different techniques are required, such as Super-resolution microscopy.

The objective of this thesis is to explore and further develop one of these techniques: the Multiple Signal Classification Algorithm (MUSICAL), a computational method that enables the reconstruction of super-resolution images from a temporal stack of diffraction-limited fluorescent images. MUSICAL exploits the fluctuations in light intensity produced by the fluorescent emitters in the sample to improve resolution and contrast. It is based on the separation of a signal and noise component, followed by their subsequent recombination.

This thesis includes five publications that contribute to different aspects of MUSICAL. Paper I enhances the usability of the method, bridging the gap between the algorithm and the end-user through an easy-to-use ImageJ plugin. This plugin allows the customization of various parameters for quick testing and facilitates the processing of multi-channel data and long stacks of images. Papers II and III contribute to expanding the technical knowledge of the algorithm by extending its definition and providing further insights into its working mechanism. In these papers, soft MUSICAL and contrast enhancement (CE) are presented for the first time. These methods impose fewer constraints on the separation between signal and noise and rely less on the user's input.

Finally, the thesis concludes with Papers IV and V, which provide case studies of practical illumination mechanisms for microscopy that can potentially complement MUSICAL to achieve a resolution gain up to three times the conventional limit. These two systems correspond to chip-based microscopy and structured illumination, whose data was studied and then processed with MUSICAL.

Overall, this thesis contributes to the improvement of MUSICAL in three key areas and can be read as a story of the algorithm. Therefore, it also serves as a window briefly showcasing the tremendous potential behind MUSICAL, which will hopefully inspire further developments in super-resolution microscopy.



# Acknowledgements

Research is a collective endeavor aimed at expanding human understanding of the world. Therefore, we should express gratitude for all the time spent by great minds in the past that has brought us to this point. However, I want to focus on the individuals I have personally met, as they have played a significant role in helping me reach this point.

First and foremost, I would like to express my deepest appreciation for my mentor and supervisor, Professor Krishna Agarwal: without your patience and positive attitude, I would not have known how to navigate the vast sea of possibilities that a PhD can entail. Thanks for all the time spent with me through discussions and life advice. I will always treasure working under your guidance.

To my co-supervisors: Balpreet S. Ahluwalia, thank you for introducing me to the world of super-resolution and for establishing such an exciting research group in the far north. Fred Godtlielsen, thanks for your positive feedback every time I needed it.

To Ida S. Opstad, thanks for your proactivity in providing me with data and for testing MUSICAL so extensively. I would not have made as much progress if it weren't for your inquiries and motivation.

To Biswajoy Gosh, thanks for being a good friend. Your passion for research is contagious and has often helped keep me focused on my work.

To all the people I shared the office with: Luis Villegas, Nikhil Jayakumar, Farhad Niknam, Sunil Bhatt, and Daniel Hansen. Thank you for all the technical and not-so-technical discussions. You helped me maintain my spirits high during moments of doubt.

To the entire Nanoscopy group at UiT, you were always an inspiration.

To my parents, thanks for all the support over the years. I reached this point thanks to the values and education I received from you.

Finally, to the love of my life Millaray: thank you for your unconditional support, for the encouragement, and for always believing in me. Without you, I would have succumbed years ago. It is safe to say that your support has been an invaluable contribution to the completion of this thesis.



# List of Publications

## Publications included in the thesis

### I. MusiJ: an ImageJ plugin for video nanoscopy

**Sebastian Acuña\***, Florian Ströhl\*, Ida S. Opstad, Balpreet S. Ahluwalia, and Krishna Agarwal

Biomed. Opt. Express 11, 2548-2559 (2020), <https://doi.org/10.1364/BOE.382735>

### II. Soft thresholding schemes for multiple signal classification algorithm

**Sebastian Acuña**, Ida S. Opstad, Fred Godtlielsen, Balpreet S. Ahluwalia, and Krishna Agarwal.

Opt. Express 28, 34434-34449 (2020), <https://doi.org/10.1364/OE.409363>

### III. Deriving high contrast fluorescence microscopy images through low contrast noisy image stacks

**Sebastian Acuña\***, Mayank Roy\*, Luis E. Villegas-Hernandez, Vishesh K. Dubey, Balpreet S. Ahluwalia, and Krishna Agarwal.

Biomed. Opt. Express 12, 5529-5543 (2021), <https://doi.org/10.1364/BOE.422747>

### IV. Fluorescence fluctuation-based super-resolution microscopy using multi-modal waveguided illumination

Ida S. Opstad, Daniel H. Hansen, **Sebastian Acuña**, Florian Ströhl, Anish Priyadarshi, Jean-Claude Tinguely, Firehun T. Dullo, Roy A. Dalmo, Tore Seternes, Balpreet S. Ahluwalia, and Krishna Agarwal.

Opt. Express 29, 23368-23380 (2021), <https://doi.org/10.1364/oe.423809>

### V. Scalable-resolution structured illumination microscopy

Ankit Butola\*, **Sebastian Acuña\***, Daniel Henry Hansen, and Krishna Agarwal.

Opt. Express 30, 43752-43767 (2022), <https://doi.org/10.1364/OE.465303>

---

\* First authorship is shared.

## Publications not included in the thesis

1. **Chip-based multimodal super-resolution microscopy for histological investigations of cryopreserved tissue sections**

Luis E. Villegas-Hernández, Vishesh Dubey, Mona Nystad, Jean-Claude Tinguely, David A Coucheron, Firehun T. Dullo, Anish Priyadarshi, **Sebastian Acuña**, Azeem Ahmad, José M. Mateos, Gery Barmettler, Urs Ziegler, Ása B. Birgisdottir, Aud-Malin K. Hovd, Kristin A. Fenton, Ganesh Acharya, Krishna Agarwal, and Balpreet S. Ahluwalia.

Light Sci Appl. 2022 Feb 24;11(1):43.

2. **Highly Efficient and Scalable Framework for High-Speed Super-Resolution Microscopy**

Quan Do, **Sebastian Acuña**, Jon Ivar Kristiansen, Krishna Agarwal, and Phong Hoai Ha.

IEEE Access (2021).

3. **Blind super-resolution approach for exploiting illumination variety in optical-lattice illumination microscopy**

Krishnendu Samanta, Swagato Sarkar, **Sebastian Acuña**, Joby Joseph, Balpreet S. Ahluwalia, and Krishna Agarwal.

ACS Photonics 8 (9), 2626-2634 (2021)

4. **Artefact removal in ground truth deficient fluctuations-based nanoscopy images using deep learning**

Suyog Jadhav, **Sebastian Acuña**, Ida S. Opstad, Balpreet Singh Ahluwalia, Krishna Agarwal, and Dilip K. Prasad.

Biomed. Opt. Express 12, 191-210 (2021)

## Conferences

1. **MUSICAL – a versatile super-resolution microscopy solution for gentle imaging of physical characteristics of biological samples**

**Sebastian Acuña** and Krishna Agarwal.

4th International Conference On Optics, Photonics, And Lasers (OPL-2023). 4-7 Dec. 2023, Hiroshima, Japan. (Invited)

2. **MUSICAL – a versatile computational nanoscopy approach**

**Sebastian Acuña** and Krishna Agarwal,



European Summit On Laser Optics & Photonics Technology (ELOPS 2023), 25-27 Spe. 2023, Barcelona. (Invited).

3. **Multiscale resolution in structured illumination microscopy**

Ankit Butola, **Sebastian Acuña**, Daniel Henry Hansen, and Krishna Agarwal.

Advances in Microscopic Imaging IV, vol. 12630, pp. 32-34. SPIE, 2023.

4. **Quantitative phase imaging for tracing the motion of waveguide trapped bead particle**

Sunil Bhatt, Ankit Butola, **Sebastian Acuña**, Daniel Henry Hansen, Jean-Claude Tinguely, Dalip Singh Mehta, Balpreet Singh Ahluwalia, and Krishna Agarwal.

Imaging, Manipulation, and Analysis of Biomolecules, Cells, and Tissues XXI, vol. 12383, pp. 19-26. SPIE, 2023.

5. **Photonic chip: a high-throughput multimodal nanoscopy platform for histopathology**

Luis E. Villegas-Hernández, Vishesh K. Dubey, **Sebastian Acuña**, Hong Mao, Jean-Claude Tinguely, Krishna Agarwal, and Balpreet S. Ahluwalia.

Focus on Microscopy, Porto, 2-5 April 2023.

6. **MUSICAL: a family of computational techniques for fluorescence nanoscopy**

**Sebastian Acuña**, Ida S. Opstad, Luis E. Villegas-Hernández, Balpreet S. Ahluwalia, and Krishna Agarwal.

Focus on Microscopy, Porto, 2-5 April 2023.

7. **Optical waveguides for multimodal super-resolution microscopy**

Jean-Claude Tinguely, Ida S. Opstad, **Sebastian Acuña**, Luis E. Villegas-Hernández, Vishesh K. Dubey, Florian Ströhl, Krishna Agarwal, and Balpreet S. Ahluwalia.

Focus on Microscopy, Porto, 2-5 April 2023.

8. **MUSICAL: results on chip nanoscopy**

**Sebastian Acuña**, Luis E. Villegas-Hernández, Ida S. Opstad, Jean-Claude Tinguely, Balpreet S. Ahluwalia, and Krishna Agarwal.

Focus on Microscopy, Porto, 2-5 April 2023.

9. **Blind reconstruction of challenging raw data outperforms commercial 3DSIM**

Ida S. Opstad, **Sebastian Acuña**, Stefan Ströhl, Florian Ströhl, Balpreet S. Ahluwalia, and Krishna Agarwal.

Focus on Microscopy, Porto, 2-5 April 2023.

10. **From super-resolution imaging to cellular dynamics**

Ida S. Opstad, **Sebastian Acuña**, Åsa B. Birgisdottir, Balpreet S. Ahluwalia, and Krishna Agarwal.

Focus on Microscopy, Porto, 2-5 April 2023

11. **Multiscale resolution in Structured illumination microscopy**

Ankit Butola, **Sebastian Acuña**, Daniel H. Hansen, and Krishna Agarwal.

Focus on Microscopy, Porto, 2-5 April 2023.

12. **Integrated photonics for super-resolution microscopy**

Jean-Claude Tinguely, Ida S. Opstad, **Sebastián Acuña**, Luis E. Villegas-Hernández, Vishesh K. Dubey, Florian Strohl, Krishna Agarwal, and Balpreet S. Ahluwalia.

Integrated Photonics Research, Silicon and Nanophotonics, pp. ITu1B-3. Optica Publishing Group, 2022.

13. **Large area super-resolution TIRFM of arctic fish species**

Ida S. Opstad, Daniel H. Hansen, **Sebastian Acuña**, Florian Ströhl, Anish Priyadarshi, Jean-Claude Tinguely, Firehun T. Dullo, Tore Seternes, Roy A. Dalmo, Balpreet S. Ahluwalia, and Krishna Agarwal.

Focus on Microscopy (FoM), 10-13 April 2022.

14. **Photonic chip for multimodal super-resolution imaging in histopathology applications**

Vishesh K. Dubey, Luis E. Villegas Hernández, Mona Nystad, Jean-Claude Tinguely, David A. Coucheron, Firehun T. Dullo, Anish Priyadarshi, **Sebastian Acuña**, Azeem Ahmad, Jose M. Mateos, Gery Barnettler, Urs Zeigler, Åsa B. Birgisdottir, A-M K. Hovd, K. A. Fenton, G. Acharya, K. Agarwal, B. S. Ahluwalia.

2021 INTPART Winter Workshop on Microscopy, Bioimaging and Artificial Intelligence, 20-22 December 2021.

15. **Autothresholding in MUSICAL using non-orthogonal signal-noise separation** **Sebastian Acuña** and Krishna Agarwal.

Convened session “Women’s Contributions in Inverse Problems,” The International Union of Radio Science General Assembly and Scientific Symposium, URSI GASS 2021, Rome, Italy (Virtual), 28 August - 4 September 2021.

16. **Photonic-chip: a multimodal imaging tool for histopathology**

Luis E. Villegas-Hernández, Vishesh K. Dubey, Jean-Claude Tinguely, David A. Coucheron, Anish Priyadarshi, **Sebastian Acuña**, Krishna Agarwal, José M. Mateos, Mona Nystad, Aud-Malin K. Hovd, Kristin A. Fenton, and Balpreet S. Ahluwalia.

Proceedings, Biophotonics Congress 2021: Novel Techniques in Microscopy, Washington D. C., United States of America (Virtual), April 12–16, 2021.

17. **MusiJ – An imageJ tool for fluorescence nanoscopy of living cells**

**Sebastian Acuña**

Next generation live-cell microscopy, Tromsø, Norway, 17-19 February 2020.

18. **MUSICAL - concept and results**

**Sebastian Acuña**

Super-resolution Optical Fluctuation Imaging (SOFI) Meeting 2019, Bielefeld, Germany, 29-30 August 2019.

19. **Fluorescence nanoscopy using multiple signal classification algorithm (musical): from electromagnetics inverse scattering problem to optics inverse source problem**

**Sebastian Acuña**, Ida S. Opstad, Luis E. V. Hernandez, Mona Nystad, Milton O. Aguilera, Merete Storflor, Cristina I. Øie, Søren Abel, Åsa B. Birgisdottir, Trine Kalstad, Deanna L. Wolfson, Jean-Claude Tinguely, Truls Myrmel, Peter McCourt, Terje Johansen, Balpreet S. Ahluwalia, and Krishna Agarwal.

Progress in Electromagnetics Research Symposium, Rome, Italy, 17-20 June 2019.

20. **Multiple Signal Classification: challenges on the route from millimeter resolution to nanometer resolution**

**Sebastian Acuña**, G. M. A. Mehedi Hussain, Fred Godtlielsen, Balpreet S. Ahluwalia, Hoai Phuong Ha, Dilip K. Prasad, and Krishna Agarwal.

Progress in Electromagnetics Research Symposium, Rome, Italy, 17-20 June 2019.

## Patents

1. **Window-based parallelized method and system for producing super-resolution microscopy image from an image time-series**

**Sebastián Andrés Acuña Maldonado**, Krishna Agarwal, and Balpreet Singh Ahluwalia

Publication Number: WO/2023/209318

Available in: <https://patentscope.wipo.int/search/en/detail.jsf?docId=WO2023209318>

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Overview of included papers . . . . .	2
1.2	Structure of the thesis . . . . .	4
<b>2</b>	<b>Fluorescence microscopy</b>	<b>7</b>
2.1	Light . . . . .	9
2.2	The Optical Microscope . . . . .	10
2.3	Image formation . . . . .	11
2.4	Point-spread-function . . . . .	12
2.5	3D imaging . . . . .	14
2.6	Resolution, contrast and noise . . . . .	16
2.7	Fluorescence . . . . .	18
2.7.1	Ground and excited states . . . . .	19
2.7.2	Absorption and emission spectra . . . . .	20
2.8	Fluorescence Microscopy . . . . .	22
<b>3</b>	<b>Linear Algebra</b>	<b>25</b>
3.1	Vector spaces . . . . .	25
3.2	Subspaces and span . . . . .	27
3.3	Inner product . . . . .	27
3.4	Operating with vectors and matrices . . . . .	28
3.4.1	Addition and scalar multiplication . . . . .	28
3.4.2	Transpose . . . . .	28
3.4.3	Matrix multiplication . . . . .	29
3.4.4	Inverse of a matrix . . . . .	30
3.4.5	Euclidean inner product . . . . .	30
3.5	Orthogonal subspaces and matrices . . . . .	30
3.6	Matrix's column space and rank . . . . .	31
3.7	Linear transformations . . . . .	31
3.8	Image formation as matrix multiplication . . . . .	32
3.9	Singular value decomposition . . . . .	33

<b>4</b>	<b>Computational microscopy</b>	<b>35</b>
4.1	The inverse problem . . . . .	36
4.2	Sample simulation . . . . .	37
4.3	Deconvolution . . . . .	38
4.4	Single molecule localization . . . . .	40
4.5	Intensity fluctuation based methods . . . . .	42
4.5.1	SOFI . . . . .	44
4.5.2	MUSICAL . . . . .	45
4.5.3	SRRF . . . . .	47
4.5.4	SACD . . . . .	49
4.6	Qualitative comparison . . . . .	50
<b>5</b>	<b>MUSICAL</b>	<b>55</b>
5.1	Imaging model and decomposition . . . . .	56
5.2	Indicator function . . . . .	60
5.3	Eigenimages . . . . .	63
5.4	3D MUSICAL . . . . .	65
<b>6</b>	<b>Summary of Research</b>	<b>69</b>
6.1	Paper I . . . . .	69
6.2	Paper II . . . . .	70
6.3	Paper III . . . . .	73
6.4	Paper IV . . . . .	74
6.5	Paper V . . . . .	78
<b>7</b>	<b>Conclusion and future work</b>	<b>81</b>
7.1	Future work . . . . .	81
7.2	Final remarks . . . . .	82
	<b>References</b>	<b>83</b>
	<b>Appendix: Included articles</b>	<b>91</b>
	Paper I . . . . .	93
	Paper II . . . . .	107
	Paper III . . . . .	129
	Paper IV . . . . .	155
	Paper V . . . . .	193

# List of Figures

2.1	Scale of things <sup>†</sup> . . . . .	8
2.2	Magnifying lens diagram <sup>†</sup> . . . . .	11
2.3	Example of image formation as a convolution between the ideal image and the PSF . . . . .	13
2.4	Airy patterns for two objective lenses with different NA <sup>†</sup> . . . . .	14
2.5	Gibson and Lanni model for focus scanning . . . . .	15
2.6	Comparison of resolution criteria between Abbe's and Rayleigh's . . . . .	17
2.7	Illustrative example of the effect of noise in contrast . . . . .	18
2.8	Jablonski diagram . . . . .	20
2.9	Absorption and emission spectra theoretical fluorophore <sup>†</sup> . . . . .	21
2.10	Fluorescence microscopy image of mitochondrial structures . . . . .	23
3.1	Relation between the 4 subspaces induced by a linear transformation $\mathcal{T}$ . . . . .	32
4.1	Forward and inverse imaging problem in widefield fluorescence microscopy <sup>†</sup> . . . . .	36
4.2	Single emitter's ON-state and OFF-state simulation . . . . .	38
4.3	Deconvolution applied to a microscopy image . . . . .	39
4.4	Single molecule localization microscopy principle (SML) <sup>†</sup> . . . . .	41
4.5	Sparse and dense fluctuations . . . . .	43
4.6	Fluorescence Fluctuation-based method: processing steps . . . . .	51
4.7	Fluorescence Fluctuation-based method: results . . . . .	52
4.8	Fluorescence Fluctuation-based method: profile . . . . .	53
5.1	Sliding window . . . . .	56
5.2	MUSICAL processing workflow . . . . .	57
5.3	Singular values . . . . .	59
5.4	Simplified scheme for point versus signal and noise space for a 2-pixel detector . . . . .	61
5.5	Indicator function applied on simulated data for two different noise levels . . . . .	64
5.6	Example of eigenimages in decreasing order and their respective projections . . . . .	66
5.7	4D array into 2D array for MUSICAL processing . . . . .	67
5.8	3D reconstruction using MUSICAL . . . . .	67
6.1	MusiJ complete workflow . . . . .	71
6.2	Simulated dataset of 3 cylinders with emitters . . . . .	72

6.3	Contrast enhancement applied on tissue sample . . . . .	74
6.4	Chip-based microscopy system's illumination patterns . . . . .	75
6.5	Chip-based microscopy system's scanning of laser . . . . .	76
6.6	MUSICAL applied on chip-based microscopy system for fish scale's cells . .	77
6.7	Correlation as a function of the distance between frames . . . . .	78
6.8	Scalable-resolution structured illumination microscopy system . . . . .	79
6.9	Eigenimages of different orders for the entire field of view . . . . .	80



# Chapter 1

## Introduction

Life happens at all scales. Some of it we can observe with our eyes while some other phenomena occur in complete secrecy for most of us, including fundamental processes that keep us alive. At the core of these processes is the cell, the smallest functional unit of our body. Therefore, if we want to understand life, studying the structure of the cell is essential as it can explain how we get sick, how we heal, how we grow, and how we die. This is the purpose of life sciences, in which the microscope stands out as one of its most important tools. In its purest form, this instrument can be simply defined as a tool that allows human beings to study and observe small objects. Small enough that our eyes cannot see them. However, while the concept is fairly simple to explain, its evolution is nothing but a series of incredible milestones [1]. Currently, the field of microscopy is wide and not only relegated to the study of forms and shapes. It can be used to measure acidity, concentration of molecules, temperature, motion, and the list goes on.

Among microscopy techniques, fluorescence microscopy stands out from the rest by providing high specificity and contrast.

However, in its simplest form, the diffraction of light imposes a limit on its achievable resolution. This is because conventional fluorescent microscopes are unable to resolve between objects separated by less than 200 nm. To obtain smaller details, more complicated optical systems are required such as structured illumination microscopy (SIM) [2], stimulated emission depletion (STED) [3], single molecule localization (SML) [4], or most recently minimal photon fluxes (MINFLUX) [5]. However, there exists another route through which the optical system does not need to change. This is done by exploiting the stochastic behavior of photon emission of fluorescent molecules which can be captured through temporal acquisition of multiple images of the same sample. This can be compared to SML except that the number of frames is reduced enormously. The earliest algorithm that exploited this concept was super-resolution optical fluctuation imaging (SOFI) [6]. From then, several others followed [7, 8, 9, 10] as this approach has several advantages over other systems, especially regarding total exposure time. This thesis is centered on one of such techniques: multiple signal classification algorithm, also called MUSICAL [11]. When compared to other techniques, its main advantage is its broader applicability as it

has been proven to work under a wide range of structured illumination systems and sample conditions. Regarding its working principles, it is based on singular value decomposition (SVD) [12], a matrix decomposition algorithm that MUSICAL uses to separate the signal and noise components on the observed data. These two components are then combined using a function called *indicator function* into a single resolution-increased microscopy image. Further details of the algorithm are left for Chapter 5.

The work presented in this thesis starts where the first work published on MUSICAL left in 2016 [11] when Agarwal showed how an algorithm used by the electromagnetic community (MUSIC) [13] could be adapted to fluorescence microscopy. Since then, a dozen related research papers have been published. This thesis presents 5 of them.

Paper I presents the first ImageJ [14] implementation of MUSICAL, which contributed towards improving its usability and therefore to disseminate the algorithm as an easy-to-use tool for researchers. Paper II and Paper III contributed through deeper knowledge of the indicator function and provided variations of it that extend the capabilities of the algorithm. Finally, Papers IV and V contribute to the study of the versatility of MUSICAL by presenting the results of two types of illumination-engineered systems. An extended discussion of these articles will be presented in Chapter 6.

The selected papers tell the story of MUSICAL from its envisioning and can be considered in terms of these three contributions:

- Extend MUSICAL’s usability: Paper I.
- To develop the technique further; Papers II and III.
- Analyze its usage on different optical systems: Paper IV and V.

## 1.1 Overview of included papers

### Paper I

**Title:** MusiJ: an ImageJ plugin for video nanoscopy

**Authors:** Sebastian Acuña, Florian Ströhl, Ida S. Opstad, Balpreet S. Ahluwalia, and Krishna Agarwal.

**Overview:** This work presents an open-access implementation of MUSICAL for ImageJ. ImageJ is a wide-spread platform for image analysis and manipulation that among other features allow biologist to work on microscopy images. The implementation provided here allows quick installation and usage of the MUSICAL algorithm for testing purposes. Before this implementation, MUSICAL was available only in MATLAB which means an extra layer of complexity to reach the final user. This work also presents the first attempt at computational optimization as otherwise was not possible to deliver similar times to MATLAB.

**Contribution:** My contribution was the implementation of MUSICAL in Java and the optimization of its code. In addition, I provided the results, comparisons, and benchmarks while contributing to writing the manuscript.

## Paper II

**Title:** Soft thresholding schemes for multiple signal classification algorithm

**Authors:** Sebastian Acuña, Ida S. Opstad, Fred Godtliebsen, Balpreet Singh Ahluwalia, and Krishna Agarwal.

**Overview:** MUSICAL is based on the selection of a threshold that separates signal and noise into two non-overlapping sets. The purpose of this paper was to study different mechanisms for dealing with the fact that signal and noise may overlap in the microscopy domain due to Poissonian noise statistics. This was accomplished by extending the definition of the *indicator function* used in MUSICAL. The study comprises the analysis of conventional MUSICAL and different variations of indicator function using already existing experimental datasets combined with simulated-based studies. As a result, different functions are tested that classify signal and noise in overlapping sets. The simulations correspond to simple geometrical 3D structures in combination with a point-spread function generated using the Gibson and Lanni model [15]. Among the different methods developed, **Soft-MUSICAL** has proven to be one of the most consistent ones for future work.

**Contribution:** My contribution was in the envisioning of the indicator function, implementation of the new indicators functions in Python, simulation of the fluorescent samples, analysis of the results, and leading the manuscript writing.

## Paper III

**Title:** Deriving high contrast fluorescence microscopy images through low contrast noisy image stacks

**Authors:** Sebastian Acuña, Mayank Roy, Luis E. Villegas-Hernandez, Vishesh K. Dubey, Balpreet Sing Ahluwalia, Krishna Agarwal

**Overview:** This work continues the track of Paper II. It proposes and studies a new set of indicator functions that use only the signal part of the decomposition. The new indicator functions use the singular values to weight the contribution of the signal components and since it does not use the noise in the denominator, reconstruction is less prone to artifacts. This results in a more robust algorithm that can deal with more challenging situations than conventional MUSICAL can do due to its dependence on noise. The cost is that resolution gain is less pronounced and therefore it is a recommended tool for cases where resolution is not the main problem but out-of-focus light and noise.

**Contribution:** I guided a Master's student to implement the modifications needed for this and analyzed all the results provided on the paper.

## Paper IV

**Title:** Fluorescence fluctuation-based super-resolution microscopy using multimodal waveguided illumination

**Authors:** Ida S. Opstad, Daniel H. Hansen, **Sebastian Acuña**, Florian Ströhl, Anish Priyadarshi, Jean-Claude Tinguely, Firehun T. Dullo, Roy A. Dalmo, Tore Seternes, Balpreet S. Ahluwalia, and Krishna Agarwal

**Overview:** This work explores the utilization of MUSICAL on datasets of fluorescence datasets produced on a chip-based system. It uses an experimental dataset of fish scale cells across two different magnifications to compare results. MUSICAL on its soft version (Paper II) shows that reconstructions of images taken at 10x 0.3 NA can be comparable to conventional 60x 1.0 NA objective. This means a large field of view without compromising resolution.

**Contribution:** I performed the reconstructions and the analysis of the number of frames versus. I designed the plots to show a correlation between close frames to conclude a minimal number of frames for correct reconstructions.

## Paper V

**Title:** Scalable-resolution structured illumination microscopy

**Authors:** Ankit Butola, **Sebastian Acuña**, Daniel Henry Hansen, and Krishna Agarwal.

**Overview:** Structured illumination requires in general full knowledge or at least an estimation of parameters prior to performing a reconstruction. Here we explore MUSICAL as a blind reconstruction algorithm for structured illumination microscopy. Our system is based on a line-based pattern that can be projected at different scales in the sample. MUSICAL enables high resolution for this type of system without knowledge of the pattern itself.

**Contribution:** I designed the illumination patterns and performed an analysis of the number of frames or patterns required to obtain a proper reconstruction and the analysis in the Fourier domain. I share the first co-authorship of this article.

## 1.2 Structure of the thesis

This thesis is intended to cover all the required material for understanding MUSICAL and my doctoral work on MUSICAL. Chapter 2 describes the basic concepts surrounding microscopy and resolution. It covers image formation as it is a critical component for understanding the algorithm. It covers resolution as it is the main problem that the algorithm tackles. And finally, it covers fluorescence microscopy, which is the microscopy technique that may benefit from MUSICAL. Chapter 3 covers the essential algebraic operations required for understanding MUSICAL. It introduces the notations used for describing the algorithm and covers basic operations such as matrix multiplication and matrix decompo-

sition which form the essential underlying concepts of MUSICAL. Chapter 4 introduces computational methods applied to fluorescence microscopy. Here, MUSICAL is introduced for the first time in the thesis. Chapter 5 finally uncovers specific details of MUSIC and the indicator function. It provides deeper insight into the algorithm and the progress made by papers II and III. The last chapter is a summary of every selected paper, focusing on its main findings, applications, and how I contributed to it.



## Chapter 2

# Fluorescence microscopy

Microscopy plays a pivotal role in modern scientific research, enabling the investigation of intricate structures and phenomena at the micro and nanoscale, across various fields, including biology, materials science, physics, and chemistry. The main tool of Microscopy is a microscope, an instrument that utilizes a variety of physical phenomena to magnify and make visible the otherwise *invisible world*. These are, objects that are beyond the capabilities of the naked eye's *resolution*. Some common structures with their respective size are shown in Figure 2.1.

It is generally accepted that the first microscope was developed around 1590 by Hans Janssen or his son Zacharias Janssen [16]. This simple but functional instrument is described as being a tube with two lenses. Ever since, scientists and engineers coming from Physics, Chemistry, and most recently Computer Science, have worked together toward improving microscopes' design and capabilities, to the point at which the instrument is present in around 97% of articles in cell biology journals [17].

At present, there is a diverse range of microscopes whose imaging capabilities are enabled by a wide range of different physical phenomena. Hawkes and Spence roughly classify them in their book [18] according to the source phenomenon of the imaging process. For instance, the imaging process can involve electrons, photons, or even acoustic waves, to mention a few. As a result, each type of microscope has its own limitations due to the nature of the phenomena they are based on. Not only that, in addition to the design of the device, the techniques may need special preparation protocols such as fixation of the sample to avoid its free motion, or most recently enlargement.

Regardless of the method, it is possible to mention specifications that are universal across the different techniques. One of them is the ability to resolve two close objects as different. This is expressed under the term *resolution* and corresponds to a central topic in this work. Different microscopes have different resolutions depending on their working mechanisms.

In terms of pure resolution, recent advancements place the Electron Microscope at a resolution of 1 Å [19]. While this may be enough to observe even molecular structures in materials, it cannot be used for studying life due to the harsh conditions required during

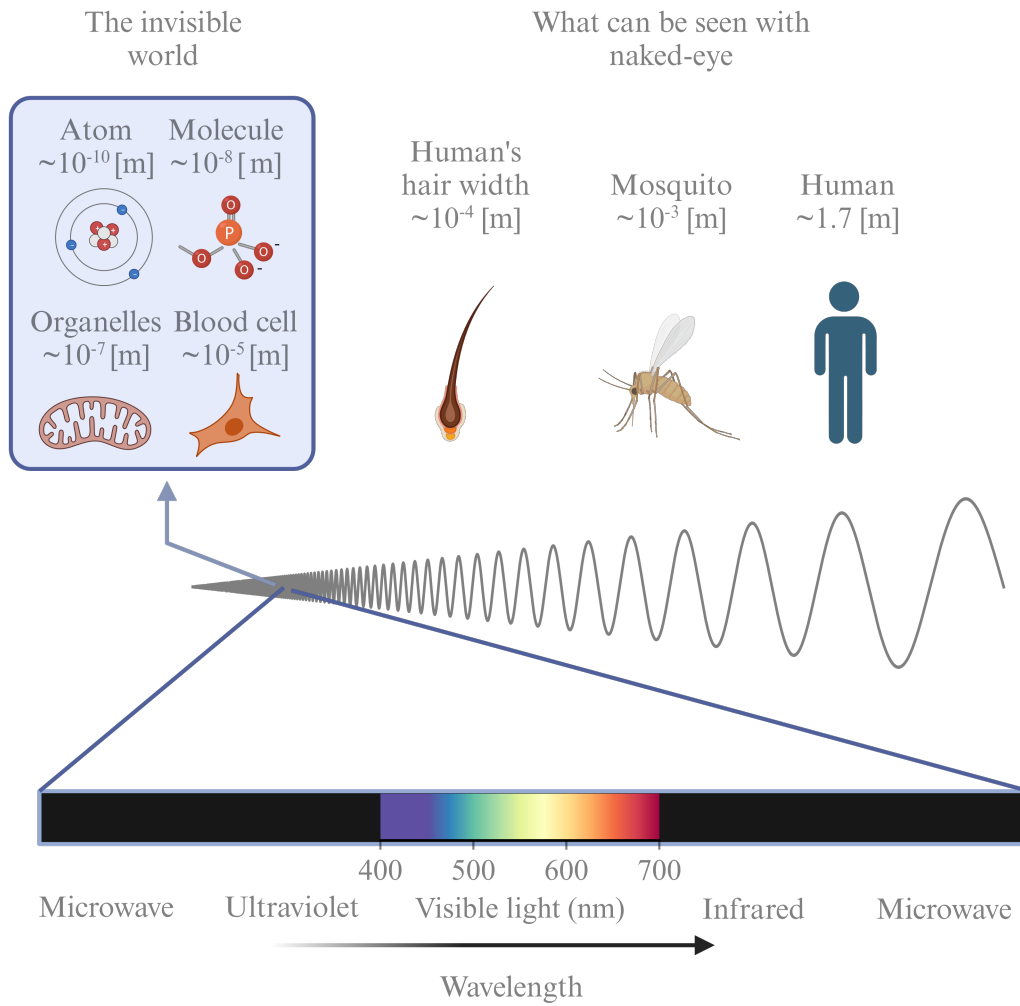


Figure 2.1: Scale of things. The human naked-eyed is not capable of observing details smaller than  $10^{-4}$  m. Therefore its resolution is limited. However, the human eye can perceive differences in the wavelength of light which produces the sense of colors.



sample preparation and posterior imaging. For the study of life and cells, in particular, the most common type of microscope falls under the umbrella of optical microscopes. They require gentler sample preparation and can often support gentler, even living state-friendly, imaging.

This chapter is centered around optical microscopy, with a specific emphasis on fluorescence microscopy. The primary objective of this chapter is to provide an overview of an optical microscope system and delve into the principles of fluorescence.

## 2.1 Light

The most fundamental concept in microscopy is Light. However, due to the complexity and breadth of the topic, the discussion is contained and simplified to the terms needed for discussing the techniques described here.

Light is a type of electromagnetic radiation and therefore, it is associated with the flux of energy in free space. Currently, there are two ways of describing its behavior. In the context of Classical Physics, it can be described using Maxwell's Equations and modeled as an Electromagnetic Wave. This model can be used to explain observed phenomena associated with light such as:

- Reflection: Light bounces at a polished surface.
- Refraction: Light bends its trajectory when going from one material to another.
- Diffraction: Light bends when passing around the edges of objects and is a critical concept to understand why Microscopes are limited in resolution.

This model also describes the light in terms of its frequency or wavelength, similar to any oscillating wave. Colors are the manifestation of this property with blue being light oscillating at 400 nm, and red at 700 nm. This range is referenced to as the *visible range* which can be observed by our eyes directly. This is only a partial range of the total Electromagnetic spectrum. Visible light stands between ultraviolet and infrared as can be seen in Figure 2.1.

The study of light under the principles of Electromagnetic Waves is called *Electromagnetic Optics*. On the other hand, light can also be modeled as the flux of a discrete number of energy packets called *photons*. The study of light as photons corresponds to the field of Quantum Optics.

Photons according to Plank have specific energy given by their wavelength  $\lambda$ . This relation can be described by the following formula:

$$E = \frac{hc}{\lambda} \quad (2.1)$$

Here  $h$  corresponds to the Plank's constant and  $c$  to the speed of light. The wavelength  $\lambda$  corresponds to the length between peaks observed in the oscillating field transported by

the photon. Light is formed by photons with wavelengths between roughly between 400 nm and 700 nm, which corresponds to a small region of the total *electromagnetic spectrum* called the *visible spectrum*. The spectrum is illustrated in Figure 2.1 as a reference.

In the context of microscopy though, photons can be used to explain photoluminescence and fluorescence in particular. fluorescence corresponds to the phenomena under which some materials may absorb light at a particular wavelength and re-emit light on a different one. Section 2.7 will delve into more details on Fluorescence as it plays a key component in the work presented in this thesis.

## 2.2 The Optical Microscope

Optical microscopy is a family of techniques that utilizes a microscope to observe and examine samples at a microscopic level using visible light or other forms of electromagnetic radiation in the optical range. This already presents an advantage when compared to other techniques as this spectrum is compatible with life and our eyes.

In simple terms, an optical microscope is a device that samples the spatial distributions in the real world and converts them to a signal that can be captured either with eyes or another type of detector. Roughly speaking, the imaging process can be separated into two parts that occur sequentially, namely illumination and acquisition.

The first part is the illumination of the sample through which the object to be observed is exposed to a source of light. The source may be sunlight, a lamp, an LED, a laser, or a combination of them. The path from the source to the sample is referred to as the *illumination path* and may or may not be part of the microscope design. The space in which the specimen is located is called the *sample space*.

The second part encompasses the acquisition of the information coming from the sample space and is normally referred to as the *imaging path*. This corresponds to the optical pathway through which the light transmitted or emitted by the specimen is collected, magnified, and focused onto a detector, such as an observer's eye, or a digital sensor such as a camera. The space in which the light is collected is referred to as the *image space*.

There exists a wide variety of optical systems and microscopes. This list includes bright-field microscopy, phase microscopy, and fluorescence microscopy, to mention a few. Each one is tailored to specific applications and provides unique insights into the microscopic world.

The simplest way of magnifying the sample space is by using a magnifying lens as shown in Figure 2.2. The system allows to capture the light produced at distance  $d_f$  and bend the light to produce an image at  $d_i$ . These distances are measured along an imaginary line called the *optical axis* and they define what is called the *focal* and *image* planes respectively. These distances are not fixed but can be related by the thin lens equation:

$$\frac{1}{d_i} + \frac{1}{d_o} = \frac{1}{f} \quad (2.2)$$

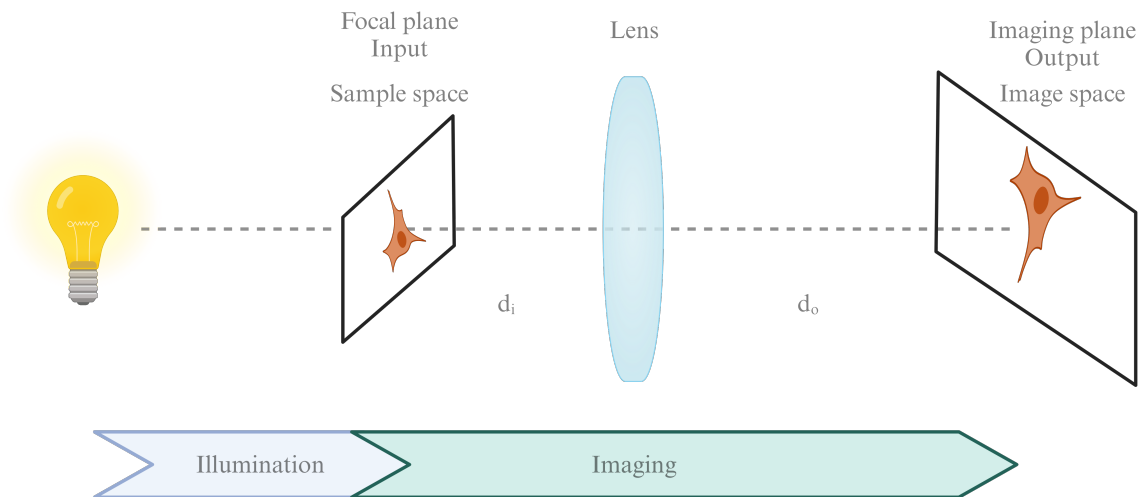


Figure 2.2: Magnifying lens diagram. The illumination path provides the sample with light. The imaging path transmits the light coming from the sample towards the sensor that captures the image.

Here,  $f$  refers to the *focal length* which is a value specific to the lens. This equation and some algebra can be used to define the magnification  $M$  for a given pair of focal and image planes as:

$$M = -\frac{d_o}{d_i} \quad (2.3)$$

This indicates that a point observed at height  $y_i$  in the focal plane will be observed at  $y_o$  in the image. Therefore, it indicates a measurement of enlargement between the object being observed and the final image.

However, magnification alone is not enough to characterize the output of the system. A most useful characterization is the point-spread function which can be used with any optical system and allows to modeling of complex optical systems by only focusing on the focal and image planes of any optical system.

In a microscope, the lens that gathers the information from the sample is called the *objective lens*.

## 2.3 Image formation

A microscope can be modeled as a linear system [20] by considering its input-output. For the purpose of this thesis, the input and output correspond to the intensity of the electric field at the focal plane and image plane. This can be done for incoherence systems where interference can be omitted [21].

On an ideal system with no loss of information, the intensities at input  $I_i$  and output  $I_o$  planes would be characterized only by the magnification:

$$I_o(x, y) = \frac{1}{|M|} I_i\left(\frac{x}{M}, \frac{y}{M}\right) \quad (2.4)$$

Here,  $x$  and  $y$  correspond to the spatial coordinates at the output plane. It is possible to improve this model by considering diffraction. This can be included in the system by using the impulse response of the system which in the context of optical systems is referred to as point-spread-function (PSF) of the system. The PSF is a function that maps the contribution of a punctual source in the focal plane to the image plane and depends on the optical system configuration. For a shift-invariant system where only the relative distance between the source and the evaluation is relevant, this can be expressed as a convolution:

$$I_o(x, y) = \int \int G(x - \xi, y - \eta) U_i(\xi, \eta) d\xi d\eta = G(x, y) * I_i(x, y) \quad (2.5)$$

To simplify notation, the magnification was omitted from  $I_i(x, y)$ . Therefore, the expression above allows to model the resulting image as the convolution between the perfect magnified image and the PSF  $G$ . This expression also indicates that a perfect PSF is given by an impulse function.

Since Equation 2.5 is a convolution, it is possible to use the convolution property of the Fourier transform on it to analyze the behavior of the optical system in terms of the spatial frequencies of the input image:

$$I_o(u, v) = G(u, v) I_i(u, v) \quad (2.6)$$

The variables  $u$  and  $v$  correspond to spatial frequencies. In this representation is easy to see how  $G$  operates as a frequency filter over the ideal image. The Fourier Transform of the PSF is called the *optical transfer function* (OTF). An example of the effect of the PSF both in spatial and frequency domains is presented in Figure 2.3. This example shows how the convolution filters high frequency components which results in the loss of details. For this example, the PSF corresponds to the Airy function which will be discussed in the next section.

## 2.4 Point-spread-function

The point spread function (PSF) of an optical system refers to the spatial response of an imaging system to an ideal point source of light. It essentially describes how the light coming from a point source spreads in the image produced by the optical system. Hence the name. Therefore, it represents a mapping between the sample space and the image space that depends on the specific system being used. The spread occurs due to diffraction, lens imperfections, and other aberrations and conforms to a characteristic pattern of a particular optical system.

Using scalar theory, it is possible to describe the PSF  $G(x, y)$  of the optical system through the Fourier transform of the aperture [21]:

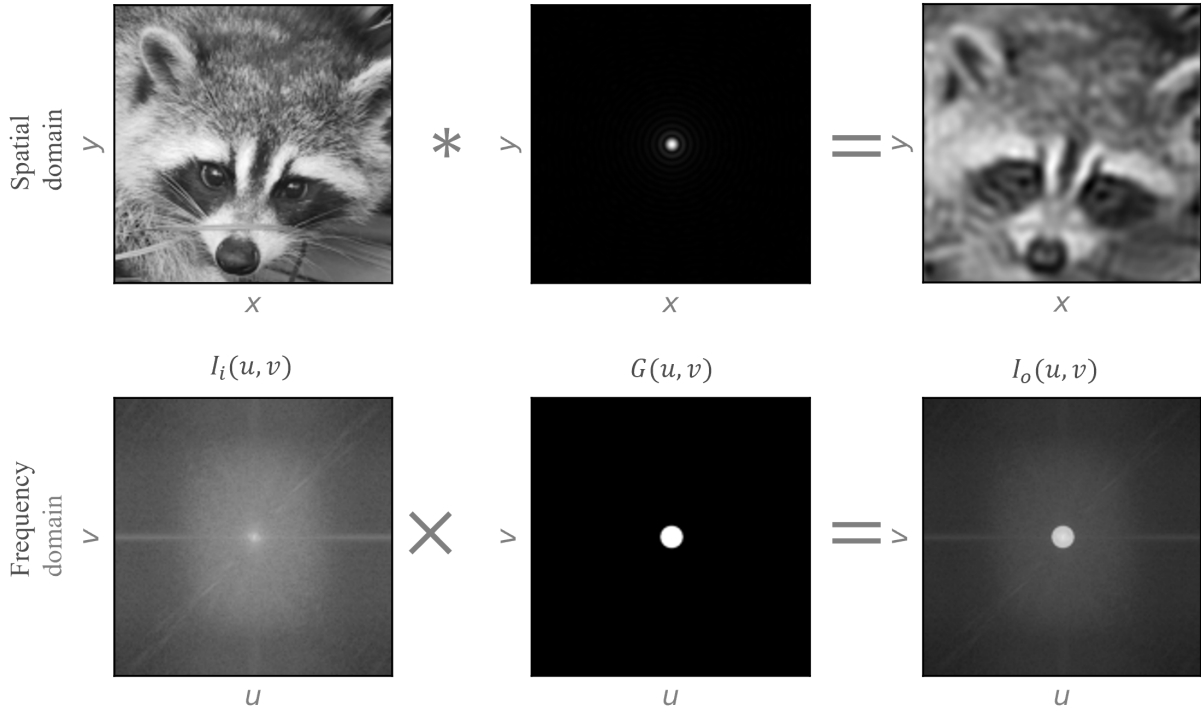


Figure 2.3: Example of image formation as a convolution between the ideal image and the PSF. The upper row presents the process through convolution in the spatial while the lower row presents the process as a multiplication. The resulting image shows less details than the original one which is also apparent from the attenuation of high-frequency components due to  $G(u, v)$ .

$$G(x, y) = \left| \frac{A}{\lambda z_i} \iint P(u, v) e^{-i \frac{2\pi}{\lambda z_i} (ux + yv)} dx dy \right|^2 \quad (2.7)$$

For a lens with a circular aperture and omitting the scalar constant, this can be written as [22]:

$$G(x, y) \propto \left| \frac{J_1 \left( \pi \frac{\sqrt{x^2 + y^2}}{r_0} \right)}{\pi \frac{\sqrt{x^2 + y^2}}{r_0}} \right|^2 \quad (2.8)$$

Here  $J_1$  corresponds to the first order of the Bessel function of the first kind. The constant  $r_0$  is related to the magnification of the configuration but also to the wavelength of light.

Equation 2.7 provides a straightforward relation between the size of the aperture and the spread of the PSF. One of the consequences of this relation is that a bigger aperture will provide a sharper PSF. For an objective lens, this can be characterized through its numerical aperture. This parameter is given by  $NA = n \sin(\theta)$  and is a measure of quantifying the angle of collection of the lens.  $\theta$  corresponds to the maximum angle that can enter the lens measured from the optical axis while  $n$  corresponds to the *index of refraction* of the media in which the lens is working. This last parameter, corresponds to the ratio

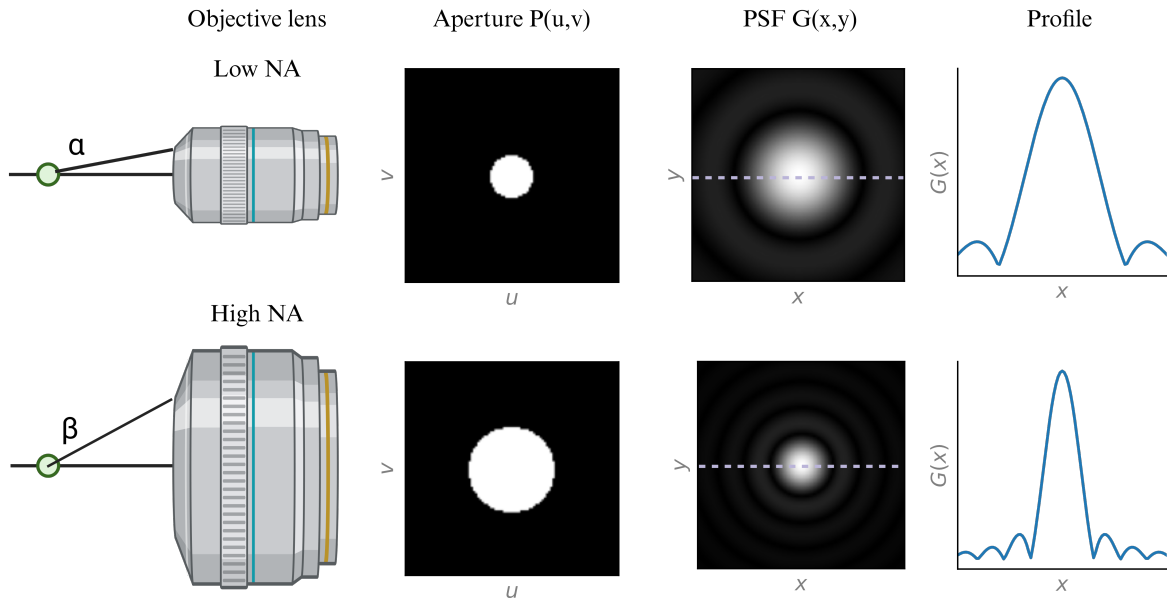


Figure 2.4: Airy patterns for two objective lenses with different NA. A high NA objective allows more information to reach the image and therefore produces a sharper PSF. The higher NA is a consequence of a higher angle  $\beta$  with  $\alpha$  corresponding to the angle for the objective with lower NA in this example. The dashed line in  $G(x, y)$  is used to draw their corresponding profiles along the horizontal. A sharper PSF means better resolution as object can be resolved at shorter distances.

between the speed of light in vacuum and the speed in the media. An illustrative example of the PSF under different NA objectives is given in Figure 2.4 where both  $P(u, v)$  and  $G(x, y)$  have been included. The figure shows two different objective lenses and provides an example of why a high NA objective allow to obtain sharper images.

The pattern given by Equation 2.8 and shown in Figure 2.4 as  $G(x, y)$  receives the name of the *Airy pattern*. The main lobe is commonly referred to as the Airy disk and is often used as the primary component to describe why microscopes are limited in resolution.

The Airy pattern shown before is valid for a planar sample located at the focal plane which means that the sample is *in focus*.

## 2.5 3D imaging

The model described previously considers point sources at a single plane, the focal plane, and in general it can be used with thin samples. However, most samples are three-dimensional, and therefore light may be produced and reach the image plane from different off-focus regions and converge at different distances than  $d_f$ . In optical microscopes with widefield illumination, 3D images can be taken by sequentially acquiring 2D images of different planes. Each plane corresponds to an optical section of the sample and the final image then corresponds to a stack of planes.

In this new scenario, is still possible to use the linear model by extending the dimen-

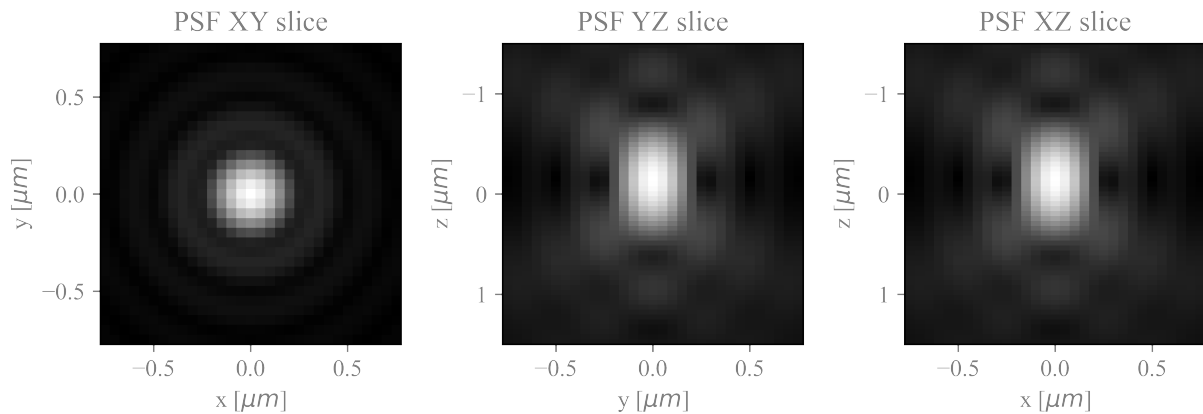


Figure 2.5: Gibson and Lanni model for focus scanning. This example corresponds to a single particle located at  $0.1 \mu\text{m}$  from the coverslip. The coverslip is moved with respect to the objective lens (a negative value in  $z$  indicates that the sample is closer to the objective). This figure was adapted from [24].

sionality of the convolution:

$$I_o(x, y, z) = G(x, y, z) * I_i(x, y, z) \quad (2.9)$$

The simulations done in Paper III use the scalar model described by Gibson and Lanni [15]. This model is scalar and therefore does not take into account light properties such as polarization. However, it is a model created for microscopy systems that considers several parameters related to the imaging process such as the refraction index of the media and glass. The Gibson-Lanni PSF is given by [15]:

$$G(x, y, z|z_p, \Psi) = \left| A \int_0^1 e^{iW(x, y, z|z_p, \Psi)} J_0(k\sqrt{x^2 + y^2}NA\rho) \rho d\rho \right|^2 \quad (2.10)$$

In this equation,  $A$  is a complex constant,  $W$  represents the optical path distance,  $z_p$  is the axial location of the source relative to the coverslip used to keep the sample pressed in the slide, and  $z$  is the location of the focal plane. The optical parameters such as numerical aperture and indices of diffraction are contained in  $\Psi$ . The simulations performed in Paper III relied on the Gibson and Lanni implementation published with [23]. The model was used to simulate the effect of out-of-focus particles with respect to a fixed focal plane.

Besides Gibson and Lanni, there are a variety of models that are used in microscopy such as the scalar model provided by Born and Wolf in [25]. It is also possible to derive the PSF following a vectorial model [26]. However, this will not be discussed here as polarization is not discussed on the papers included in this thesis.

## 2.6 Resolution, contrast and noise

The *resolution* of an optical system corresponds to the smallest distance at which two distinct objects can be distinguished. On a perfect system, this would be almost zero. However, as described previously, an optical system acts as a band pass filter due to diffraction which results in a loss of high-frequency components. This affects the ability to resolve small structures and therefore the system is said to be diffraction-limited.

There is not a single way of characterizing an exact estimate in the resolution of an optical system as different criteria have been proposed. The first reported criterion was presented by Abbe [27]:

$$r_{Abbe} \approx \frac{1}{2} \frac{\lambda}{NA} \quad (2.11)$$

Another common definition was given later by Rayleigh [28]:

$$r_{Rayleigh} = 0.61 \frac{\lambda}{NA} \quad (2.12)$$

This comes directly from studying Equation 2.8 as it is equivalent to the radius of the main lobe of the Airy disk. It can be understood as the distance such that the second source is sitting just at the first zero of the PSF projected by the first source. Both criteria are shown in Figure 2.6 where it is possible to appreciate how Rayleigh's resolution is slightly more conservative.

However, resolution is not the only thing that limits the information in the image. Another important aspect is the amount of photons that can be captured by the microscopy with respect to the background. This can be referred to as the contrast of the image.

One common definition of contrast in terms of the intensities found on a digital image  $I$  is the following [29]:

$$C = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}} \quad (2.13)$$

The subscript here indicates the maximum (max) and minimum (min) intensities of the image. In the context of microscopy images where structures vary across different regions, this definition may be used locally, and therefore in general it is not possible to define a global measurement of contrast. It is also possible to find alternatives [30] but in this work we will normally refer to contrast as shown in Equation 2.13.

Another way of referring to contrast is through the concept of *signal-to-noise ratio* or SNR. This can be defined as the ratio of what is considered a signal on the image, to the noise. Formally it may be defined as [31]:

$$SNR = \frac{\text{variance of image}}{\text{variance of noise}} \quad (2.14)$$

A similar concept is the signal-to-background ratio (SBR) which is conceptually different from SNR. The SBR considers specifically the signal from the background which may



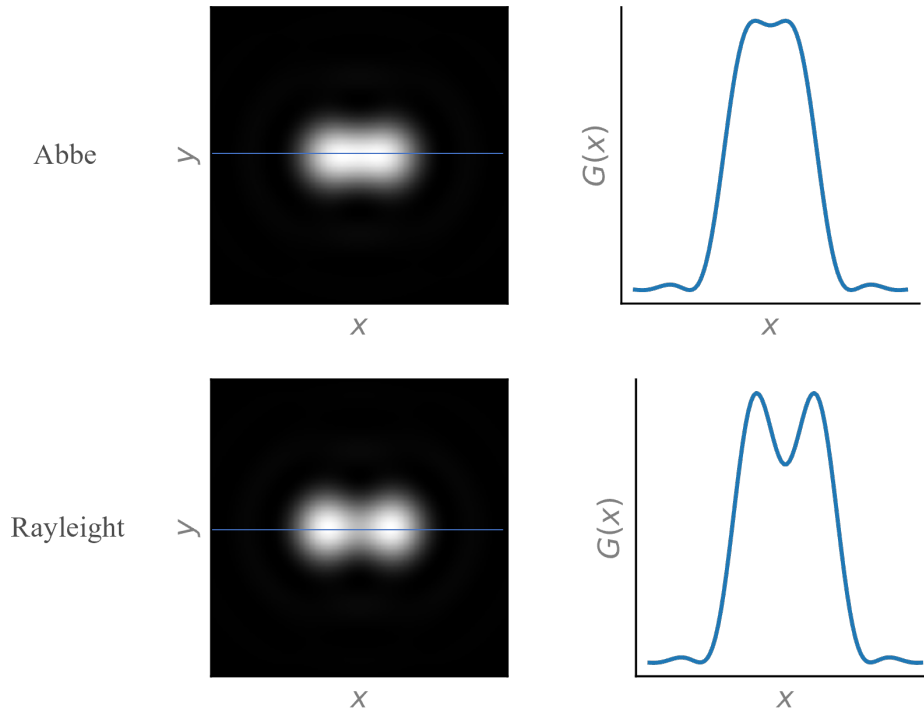


Figure 2.6: Comparison of resolution criteria between Abbe's and Rayleigh's. This illustrates how Rayleigh is a more conservative criterion. The example corresponds to a sample of two sources imaged through the same optical system. For each case, the distance was set exactly to the resolution estimated by both criteria.

not be empty and actually contain information that may contribute to the interpretability of the image. Therefore, the concept of SBR will be preferred over the SNR.

The concept *noise* refers to the corruption that affects the signal that is intended to be measured by the optical system. Therefore, it is considered an undesired modification to the *true* signal which usually involves random processes. In the case of microscopy, there are normally two sources of noise. These are called *read-noise* and *shot-noise*. The first one describes the effect of the sensor such as variations in thermal noise and does not depend on the input signal. It is normally modeled as an additive component that follows a Gaussian probability distribution. The second is produced by the quantum nature of photons [32]. This is modeled using a Poisson distribution and is a function of the signal intended to be measured:

$$p_x(x) = \text{Poisson}(x|x_0) \quad (2.15)$$

This indicates the probability density function of the measured image  $x$  given the true value  $x_0$ . For a Poisson distribution,  $x_0$  corresponds also to the average of the Poisson distribution. Therefore, when obtaining multiple measurements of  $x$ , the mean value would be an unbiased estimator of  $x_0$  due to the properties of the Poisson distribution [33].

This model, which uses the Poisson distribution, was used in Paper III to generate samples for testing the contrast and resolution of MUSICAL.

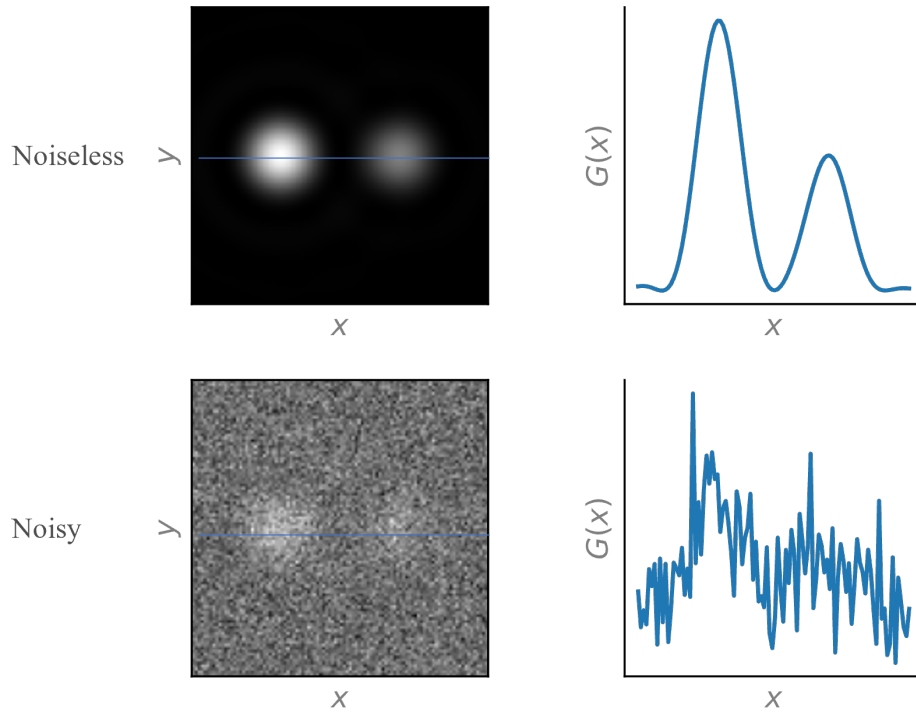


Figure 2.7: Illustrative example of the effect of noise in contrast. The example displays a sample with two different objects that are distinguishable from each other in the noiseless case (upper row). However, once additive noise is considered differentiating between structure and background becomes challenging.

The interested reader is referred to [34] for a larger discussion on noise models.

As an example, Figure 2.7 illustrates a comparison between two measurements where resolution is not an issue. The structure is formed by two distinct objects with different intensities. In this case, the addition of noise makes it difficult to differentiate the dimmer structure from the background.

## 2.7 Fluorescence

Fluorescence has been a subject of fascination and scientific exploration for over a century. This process represents a distinctive form of *photoluminescence*, wherein specific materials and molecules possess the ability to absorb photons of one energy and subsequently emit photons at a different, longer wavelength, often in the visible or ultraviolet spectrum.

Fluorescence microscopy makes use of this phenomenon to create vibrant and highly specific contrast within biological and material samples.

What differentiates fluorescence from other photoluminescence phenomena is that the absorption of a photon involves a change in the *energy level* of the molecule which involves a transition from a *ground state* to an *excited state* referred to as a single state. The emission of a photon occurs when the molecule decays from the single to the ground state. This will be discussed in the next section.

### 2.7.1 Ground and excited states

Molecules are quantum mechanical systems whose energy can take only discrete values of energy [35]. A molecule may exist in different electronic configurations which in turn may have different levels of vibrational and rotational energy. The total energy of the system is a combination of translational, rotational, vibrational, and electronic energy. The state of lowest possible energy is called the ground state and is normally referred to as the  $S_0$  state. Under normal conditions, this is where an unexcited molecule resides [22]. Some molecules have the ability to change their electronic state through the absorption of a photon and then decay again to its lowest energy level producing a photon in return. The molecules that have this property are called *fluorophores*.

There exist many levels of energy levels for a molecule, but the name and properties of the state may differ depending on the spin of the excited electron. For a pair of electrons occupying an orbital level, it is known that they must possess different spins, as described by Pauli's exclusion principle. When energy in the form of a single photon is provided to the system in such a way that it matches the difference between the ground state and an excited state, the molecule may respond by increasing the energy and therefore orbitals of one of the electrons. If there is no change of spin, the new state is called a *Singlet* state and denoted as  $S_N$  with  $N$  indicating the energy level. If the spin changes, the state is called a *Triplet* state and is denoted as  $T_N$ . Due to the low chances of occurrence, the triplet state is also referred to as a *forbidden* state.

Once the electron is excited either to singlet or triplet state, it tends to immediately lower its energy and drop to the ground state which in turn releases energy. This may happen as radiation if the energy is released in the form of a photon, or a non-radiative phenomenon as in the vibrational relaxation. The wavelength of such a photon is in general higher than the absorbed one due to a slight loss of energy.

The phenomenon of emitting a photon due to relaxation from the singlet state is called Fluorescence. The emission is immediate as the molecule tends to be in the singlet state for a very short time. If the radiation occurs from a triplet state, it is called Phosphorescence. The time-scale though is much larger than Fluorescence as the electron remains for a longer time in the triplet state before returning to a lower energy configuration. As a result, luminescence produced by Fluorescence is stronger and easier to see than the one produced by Phosphorescence.

Regardless of the state, for Fluorescence and Phosphorescence to happen the electron must first go down to the lowest vibrational level of an excited state. This reduces the energy of the electron and therefore increases the wavelength of the potential photon soon to be released.

The whole process can be summarized using a Jablonski diagram as shown in Fig. 2.8. This diagram shows the energy level at the vertical axis which allows a graphical explanation of the different levels of energy that the molecule may exist in.

During the triplet state, the molecule can be considered as a non-emitting source due

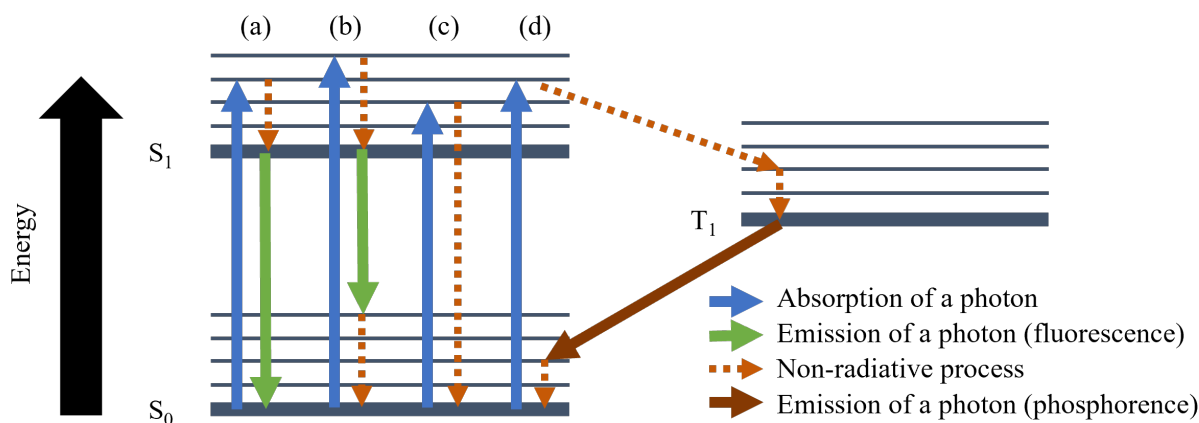


Figure 2.8: Jablonski diagram that allows to explain some of the transitions experienced by excited molecules. The thick lines correspond to different electronic levels and the thin lines correspond to different vibrational levels. The horizontal length of the arrows allows to indicate the energy involved in the process. (a) A molecule in its ground state is excited to some higher energy level. From there it goes through non-radiative relaxation until it reaches a lower electronic level. From there, a photon can be emitted. (b) The absorbed photon has higher energy than the one in (a). From there the molecules go through non-radiative relaxation and then emit a photon with even less energy than in (a). The emission in (a) and (b) is called fluorescence. (c) The absorption does not produce a photon. (d) After absorption, the molecules transit to a triple state where they can stay longer. From there it may emit a photon as well but the process is called phosphorescence.

to the lower probability of emission of a photon and its longer stay.

Using this property of the fluorophores, it is possible to abstract two distinct states. If the fluorescent molecule is able to emit fluorescent, the molecule can be considered active or in an *ON* state. Whenever the molecule enters the forbidden state, in effect it can be considered as inactive or as being in an *OFF* state. However, these states have been considered as transitory. It is also possible for the molecule to completely change its structure as a product of these transitions and permanently lose its fluorescent properties. This degradation is called *photobleaching* [36]

## 2.7.2 Absorption and emission spectra

Since there are many possible excited states, each one with its own vibrational levels, there is also a range of wavelengths that can be absorbed. In Figure 2.8 examples (a) and (b) illustrate this phenomenon and how there exist a vast number of options for energy absorption and emission. However, some transitions are preferred over others and therefore the probability is not uniform. Such preference can be explained using the Frank-Condon principle [37] but in practical terms, it can also be described by looking at the *absorption* and *emission* spectra. Each of these spectra indicates the probability of absorbing a photon as a function of its wavelength, or the probability of emitting a photon respectively.

One immediate observation is that the emission spectra are shifted to the right (higher wavelengths) compared to the absorption spectra due to the loss of energy described in

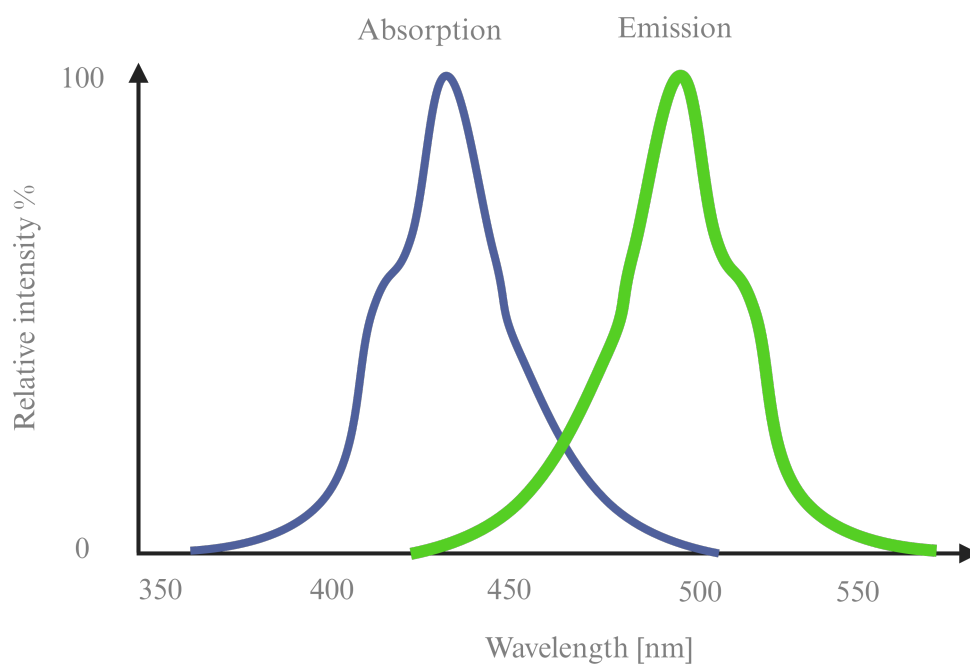


Figure 2.9: Absorption and emission spectra for a theoretical fluorophore. This illustrates the overall shapes of the spectra. The horizontal axis represents the wavelength and the vertical axis the relative intensity. This example shows how emission tends to happen to the right of the absorption due to energy losses. Generally speaking, absorption and emission can be characterized by the wavelengths where the respective peaks occur.

the previous section. The second observation is that in general there is some degree of symmetry between both which can also be explained by the Frank-Condon principle.

## 2.8 Fluorescence Microscopy

Fluorescence can be used to improve the contrast of microscopy images and to measure different properties of a sample. This is the purpose of Fluorescence Microscopy, an imaging technique that allows imaging of specific parts of the sample by attaching fluorophores to those specific parts. An example of such specificity is shown in Figure 2.10 where two structures are labeled and clearly imaged.

The imaging process in this case does not measure the sample directly but indirectly by capturing the light emitted by the fluorescent molecules. An exception though is in the case of auto-fluorescence in which the sample is able to produce light without external agents. In addition to providing structural detail of microscopic data, fluorophores can be used to image different properties of the sample such as acidity and temperature [39]. However, this will not be discussed further as it is out of the scope of this thesis.

In terms of implementations, there exists a vast range of microscope system designs. However, they all share certain characteristics when describing their illumination and acquisition paths. The illumination path is used to excite the fluorophores in the sample and needs to take into account the absorption spectrum of the fluorophore. Similarly, the acquisition path needs to consider the emission spectrum of the fluorophore. Since both spectra may overlap, the selection of fluorophores and color filters is crucial. Additionally, it is possible to use multiple markers to image multiple structures of interest which may provide even better insight into the biological structures in the sample.

There exists a vast range of ways to produce illumination and acquisition. The most common type though in fluorescence microscopy is through epi-illumination, wide-field illumination, and collection. The first means the illumination and collection occur on the same side of the sample. The second corresponds to illuminating the entire sample at once and capturing the emitted light all at once. Another way uses a rastering mechanism to sequentially illuminate and capture the information around a small volume of the sample. This is referred to normally as a confocal system. A third alternative is by using so-called structured illumination. In this case, the sample is illuminated with a non-uniform light pattern. The acquisition part is similar to how a conventional wide-field image is taken but it normally involves digital processing of the image. Examples are HiLO [40], structure illumination microscopy (SIM) [2], and random illumination microscopy (RIM) [41]. A fourth type is called a light sheet. In this type of microscope, the illumination is concentrated along a single plane in the sample which allows exceptional optical sectioning [42].

One key aspect of fluorescent microscopy is that captured light is incoherent as emitters photons independent of each other. Therefore, it is possible to use Equation 2.5 where the



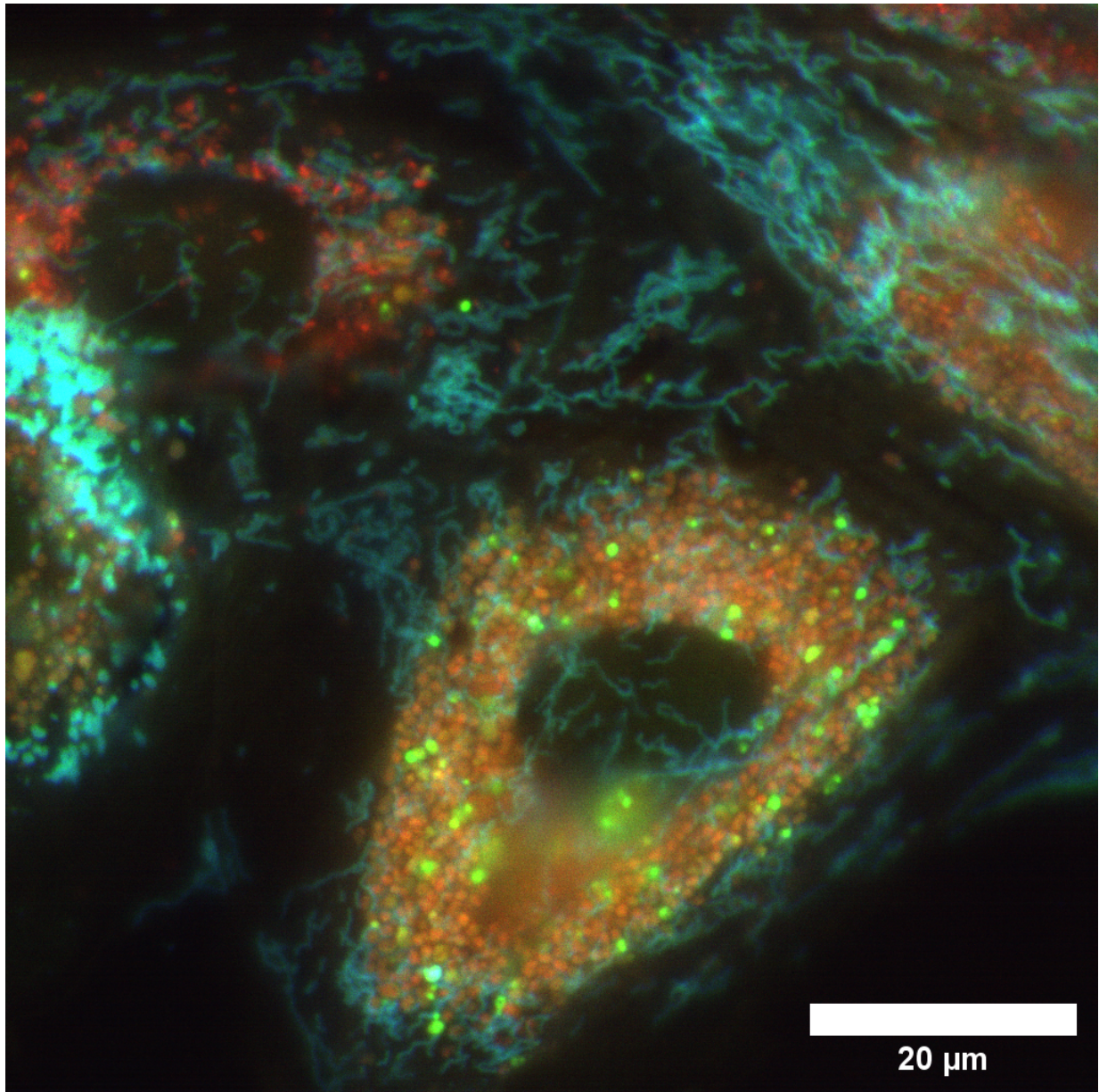


Figure 2.10: Fluorescence microscopy image of mitochondrial structures. The structure has been labeled for specific imaging of lysosomes (red) and mitochondria's outer membrane (green and blue). The dual marker enables the study of the acidity of the media through the bleaching of green light while keeping the structure visible in blue. Data available in [38]

ideal image is given just by  $I_i$ . However,  $I_i$  is in general proportional to the intensity of the illumination used to excite the fluorophores in the sample. Therefore Equation 2.5 can be used as long the illumination is uniform. However, if the illumination is non-uniform, as in confocal, structured, or light-sheet illuminations, the imaging model can be modified as follows:

$$I_o(x, y) = G(x, y) * (F(x, y)I_i(x, y)) \quad (2.16)$$

Here, the term  $F(x, y)$  denotes a non-uniform illumination pattern over the sample  $I_i$  which allows for modulation of the sample's frequency components. In the case of SIM,  $F(x, y)$  corresponds to a sinusoidal pattern. Since this pattern has two strong peaks in the frequency domain it allows twice the resolution. However, if this pattern is strong enough it can induce even more harmonics and therefore increase the theoretical resolution [43]. Regardless of the method, non-uniform illumination requires in general more than a single image and an accompanying computational algorithm that allows the reconstruction of the final microscopy image. Later on Papers IV and V, it will be shown that MUSICAL can be used in this context as well. However, the critical property of fluorophores that is used by MUSICAL is the random fluctuation in the number of photons emitted by each fluorescent molecule as a function of time. More details on this will be given in Chapters 4 and 5.

As described previously, fluorescence microscopy relies on fluorophores, which are molecules capable of producing fluorescence. In general, these molecules need to be added to the sample through a process known as *staining*. This must be done following the protocol provided by the manufacturer of such a fluorophore. The mechanism through which these molecules are attached to the sample is beyond the scope of this thesis. However, it's important to note that fluorophores are commonly used to label specific cellular structures such as membranes and actin.

So far,  $I_i(x, y)$  has been considered as the ideal image of the fluorescent distribution. It is possible to be more specific for a sample labeled with  $N$  emitters distributed along the space. A single source  $e$  located at point  $(x_e, y_e)$  can be considered to be ideally represented by  $b_e\delta(x - x_e, y - y_e)$ . Here,  $b_e$  corresponds to the emission intensity of the source  $e$  multiplied by a shifted Dirac delta function  $\delta(x, y)$ <sup>1</sup>. Therefore, a fluorescent sample can be modeled by:

$$I_i(x, y) = \sum_{e=1}^N b_e\delta(x - x_e, y - y_e) \quad (2.17)$$

---

<sup>1</sup>The Dirac delta function is not a function but a *distribution*



# Chapter 3

## Linear Algebra

MUSICAL is based extensively on Linear Algebra to convert a temporal stack of diffracted-limited images into a single super-resolved image. A key aspect of this process is to consider 2D images (or even 3D images as it will be discussed in Chapter 5) as one-dimensional vectors. This enables the treatment of an entire temporal stack as a rectangular matrix and, therefore, makes possible the utilization of linear algebra operations on it. Additionally, the use of matrices in the treatment of the image formation model based on convolution with the PSF by the ideal image into a single matrix multiplication. This chapter covers briefly concepts such as vectors, matrices, and the decomposition of matrices. In the process, it will introduce some of the notation and concepts used for understanding MUSICAL. Most of the notations and concepts here introduced are based on Chapter 2 of Eldar’s book [44].

### 3.1 Vector spaces

A vector space consists of two essential components. First, it includes a set or collection of elements known as *vectors*. Second, it encompasses a set of operations that can be performed on this set while adhering to certain defining properties. Although there are various types of vector spaces, this thesis primarily focuses on the vector space formed by real-valued vectors and matrices.

A vector space, denoted as  $\mathcal{V}$  over  $\mathbb{R}$ , is defined as a set of real elements equipped with operations for addition and scalar multiplication. To distinguish vector elements from scalar numbers, the elements within this space are represented in bold type. An illustrative example of a vector space is  $\mathbb{R}^N$ , consisting of an infinite collection of vectors, where each element  $\mathbf{u}$  is a tuple comprising  $N$  real values. The  $i^{\text{th}}$  element of  $\mathbf{u}$  is referred as  $u[i]$ :

$$\mathbf{u} = \begin{bmatrix} u[1] \\ \vdots \\ u[i] \\ \vdots \\ u[N] \end{bmatrix} \quad (3.1)$$

Another space is formed by elements from  $\mathbb{R}^{M \times N}$ . Each element in this space is termed a *matrix* and will be denoted using bold uppercase letters. The matrix  $\mathbf{U}$  is defined as follows:

$$\mathbf{U} = \begin{bmatrix} U[1,1] & \cdots & U[1,N] \\ \vdots & \ddots & \vdots \\ U[M,1] & \cdots & U[M,N] \end{bmatrix} \quad (3.2)$$

The elements of the matrix can be associated with their locations using a pair of indices, denoted as  $(i, j)$ , where  $i$  indicates the row and  $j$  represents the column. If  $M = N$  then the matrix is termed a *square* matrix. If the non-zero elements are only along the diagonal of the matrix ( $i = j$ ), the matrix is said to be *diagonal*.

Regardless of the actual element and as stated previously, a vector space requires addition and scalar multiplication. These operations act on the vectors to produce another element on the set. The first one is the *sum*, noted using the symbol  $+$ , that operates on two vectors: for  $u$  and  $v$  in the set  $\mathcal{V}$ , their sum corresponds to  $u + v$ . The second is scalar multiplication. This operates over a scalar value that belongs to  $\mathbb{R}$ , and a vector  $u$  in  $\mathcal{V}$ . The main condition is that the operations are *closed*. This is: for any two vectors  $u$  and  $v$  in the vector space  $\mathcal{V}$ , their sum  $u + v$  is also an element of  $\mathcal{V}$ . For any scalar  $\alpha$  in  $\mathbb{R}$ ,  $\alpha v$  is also in  $\mathcal{V}$ .

In order to be considered a vector space, the set must hold the following properties:

1. The sum is commutative: for any vectors  $u, v$ , the order of the sum does not affect the result. This is  $u + v = v + u$ .
2. The sum is associative: for any vectors  $u, v$ , and  $w$  in  $\mathcal{V}$ ,  $(u + v) + w = u + (v + w)$ .
3. The operations are distributive: For any vectors  $u, v$  in  $\mathcal{V}$  and scalar  $\alpha$ ,  $\alpha(u + v) = \alpha u + \alpha v$ .
4. Existence of additive identity in the set: There exists a special vector called the zero vector in  $\mathcal{V}$ , denoted as  $0$ , which satisfies the property that for any vector  $v$  in  $\mathcal{V}$ ,  $v + 0 = v$ .
5. Existence of additive inverse in the set: For every vector  $v$  in  $\mathcal{V}$ , there exists another vector  $-v$  in the space such that  $v + (-v) = 0$ .
6. Multiplicative identity: For any vector  $v$  in  $\mathcal{V}$ ,  $1v = v$ .

While the term vector is used in the context of Algebra as the elements of a vector space, it is more commonly used to indicate elements of  $\mathbb{R}^N$  (or  $\mathbb{C}^N$ ). This convention will also be followed from now on.

## 3.2 Subspaces and span

A vector space  $\mathcal{W}$  is termed a *subspace* of the vector space  $\mathcal{V}$  if all the elements of  $\mathcal{W}$  belong to  $\mathcal{V}$  and if  $\mathcal{W}$  is closed under sum and scalar multiplication. An example of subspace, that is also relevant for MUSICAL, is the *span* of a set of vectors. Let's consider a finite set  $\mathcal{X} = \{x_1, x_2, \dots, x_N\}$  of elements in  $\mathcal{V}$  that may not necessarily be a vector space. The span of  $\mathcal{X}$  corresponds to the set formed by all the elements that can be obtained through the *linear combination* of the vectors of  $\mathcal{X}$ . Using a list of scalars  $\alpha_i$  this can be expressed as:

$$\text{span}(\mathcal{X}) = \left\{ \sum_{i=1}^N \alpha_i x_i \right\} \quad (3.3)$$

The span of a set of vectors  $\mathcal{X}$  is a vector space and will be written as  $\text{span}(\mathcal{X})$ .

Of particular interest is the element 0 in  $\text{span}(\mathcal{X})$ . We say that the vectors of  $\mathcal{X}$  are linearly independent if:

$$0 = \sum_{i=1}^N \alpha_i x_i \rightarrow \alpha_i = 0 \quad \forall i \quad (3.4)$$

In other words, the only way the zero vector can be obtained is if all the scalars are zero.

If the elements of  $\mathcal{X}$  are linearly independent and  $\text{span}(\mathcal{X}) = \mathcal{V}$  (referred too as  $\mathcal{X}$  being *complete*), then  $\mathcal{X}$  is said to be a basis of  $\mathcal{V}$ .

## 3.3 Inner product

The inner product is an operation that allows a geometrical perspective of vector spaces. The operation acts over two elements of the vector space  $\mathcal{V}$  to produce a scalar value. Only the case for real scalar will be considered. For  $u$  and  $v$  in  $\mathcal{V}$ , the inner product is written as  $\langle u, v \rangle$  and must hold the following properties:

1. Symmetry: For  $u$  and  $v$  in  $\mathcal{V}$ ,  $\langle u, v \rangle = \langle v, u \rangle$ .
2. Linearity: For  $u, v$  and  $w$  in  $\mathcal{V}$ ,  $\langle u, \alpha v + \beta w \rangle = \alpha \langle u, v \rangle + \beta \langle u, w \rangle$ .
3. Positive-definiteness: For  $v$  in  $\mathcal{V}$ ,  $\langle v, v \rangle \geq 0$ . If  $\langle v, v \rangle = 0$ , then  $v$  must be the zero vector.

If the inner product between two non-zero vectors is zero, we say that those vectors are perpendicular or *orthogonal* between each other.

With the inner product, it is possible to *induce* the *norm* of a vector  $v$ . A norm can be interpreted as the *length* of a vector and its definition is not unique. In the case of this thesis, when dealing with inner products the norm ( $\|\cdot\|$ ) of a vector  $\mathbf{u}$  in  $\mathcal{U}$  will be defined as:

$$\|\mathbf{u}\| = \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle} \quad (3.5)$$

## 3.4 Operating with vectors and matrices

In what follows, the focus will be on vector spaces formed by vectors and matrices. This is, spaces whose elements belong to  $\mathbb{R}^M$  and  $\mathbb{R}^{M \times N}$  respectively. A single matrix in  $\mathbb{R}^{M \times N}$  can be seen as a collection of vectors in  $\mathbb{R}^M$  or  $\mathbb{R}^N$ . These can be referred to as *row vectors* or *column vectors* respectively. As the name indicates, row vectors are associated to the rows of the matrix and column vectors to its columns. These types of vectors can be considered as the time as matrices where either  $M$  or  $N$  are equal to one.

### 3.4.1 Addition and scalar multiplication

Let's consider matrices  $\mathbf{U}, \mathbf{V}$  and  $\mathbf{W}$  in  $\mathbb{R}^{M \times N}$ . The addition  $\mathbf{W} = \mathbf{U} + \mathbf{V}$  is obtained as an element-wise sum between real numbers:

$$W[i, j] = U[i, j] + V[i, j] \quad (3.6)$$

Similarly, the scalar multiplication  $\mathbf{V} = \alpha \mathbf{U}$  is defined as:

$$V[i, j] = \alpha U[i, j] \quad (3.7)$$

The multiplication can also be used to define the division  $\mathbf{V} = \frac{\mathbf{U}}{\alpha}$  which is defined as long as  $\alpha \neq 0$ .

### 3.4.2 Transpose

The transpose  $\mathbf{V}$  of a matrix  $\mathbf{U}$  is written as  $\mathbf{V} = \mathbf{U}^T$  and is defined as:

$$V[i, j] = U[j, i] \quad (3.8)$$

From this definition, the transpose of a column vector is a row vector, and vice versa.

### 3.4.3 Matrix multiplication

The matrix multiplication is a non-commutative operation defined between two matrices  $\mathbf{U}$  and  $\mathbf{V}$  that produces a new matrix  $\mathbf{W}$ . It is simply written as  $\mathbf{W} = \mathbf{UV}$  and is only defined for  $\mathbf{U} \in \mathbb{R}^{M \times N}$  and  $\mathbf{V} \in \mathbb{R}^{N \times K}$  with  $\mathbf{W} \in \mathbb{R}^{M \times K}$ . It is defined as:

$$W[i, k] = \sum_{j=1}^N U[i, j]V[j, k] \quad (3.9)$$

This definition can also be used to multiply matrices with vectors as long the dimensions match. For  $\mathbf{U} \in \mathbb{R}^{N \times M}$  and  $\mathbf{v} \in \mathbb{R}^{M \times 1}$ , the matrix-vector multiplication gives as result  $\mathbf{w} \in \mathbb{R}^{N \times 1}$  is given by:

$$\mathbf{w} = \mathbf{Uv} \quad (3.10)$$

This expression corresponds to the linear combination of the columns of  $\mathbf{U}$  with the weights being the elements of  $\mathbf{v}$ . Therefore, the resulting vector  $\mathbf{w}$  belongs to the span of the columns of  $\mathbf{U}$ . The same expression can be used to obtain the  $i$ th column of  $\mathbf{W}$  by taking only the  $i$ th column of  $\mathbf{V}$ .

The matrix multiplication operation has an identity element called the *identity matrix*, denoted as  $\mathbf{I}$ . The key property of  $\mathbf{I}$  is that for any matrix  $\mathbf{U}$ :

$$\mathbf{UI} = \mathbf{U} \quad (3.11)$$

The identity matrix  $\mathbf{I}$  must have a valid number of dimensions depending on whether it is on the left or right of  $\mathbf{U}$  and is defined as a square matrix whose elements are  $I[i, j] = 1$  for  $i = j$  and 0 otherwise.

Another interesting property of the multiplication is given by the transpose of the multiplication of several matrices. For a series of  $N$  matrices  $\mathbf{A}_i$ :

$$(\mathbf{A}_1 \mathbf{A}_2 \cdots \mathbf{A}_N)^T = \mathbf{A}_N^T \cdots \mathbf{A}_2^T \mathbf{A}_1^T \quad (3.12)$$

### 3.4.4 Inverse of a matrix

If  $\mathbf{A}$  is multiplied by  $\mathbf{X}$  and its result is the identity matrix, then  $\mathbf{X}$  is said to be the inverse of  $\mathbf{A}$ . This is written as  $\mathbf{X} = \mathbf{A}^{-1}$ . Depending on the side by which the inverse is used it is possible to describe a left inverse and a right inverse. For a square matrix, both are the same. However, not every matrix has an inverse in which case it is called *singular* matrix. A matrix whose columns are linear independent is assured to have an inverse.

### 3.4.5 Euclidean inner product

For column vectors, the most common option for the inner product is the Euclidean inner product. This is normally written using a single dot ( $\cdot$ ) and is therefore called the *dot product*. The Euclidean inner product between two vectors  $u$  and  $v$  in  $\mathbb{R}^{N \times 1}$  is defined as:

$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T \mathbf{v} \quad (3.13)$$

This inner product induces the Euclidean norm:

$$\|\mathbf{u}\| = \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle} = \sqrt{\mathbf{u}^T \mathbf{u}} = \sqrt{\sum_{i=1}^N u[i]^2} \quad (3.14)$$

The interpretation of the Euclidean norm for vectors in  $\mathbb{R}^N$  is of the length. The easiest way of understanding this is when considering the position of an object in a three-dimensional space. In this case, the Euclidean norm is the total distance of the object to the center of the frame of reference.

When a vector space is paired with its inner product they are also referred to as *Hilbert spaces*.

## 3.5 Orthogonal subspaces and matrices

Two vectors are said to be orthogonal when their inner product is zero. Similarly, there exists orthogonality between subspaces. Formally, let  $\mathcal{H}$  be a Hilbert space, and let  $\mathcal{U}$  and  $\mathcal{V}$  be subspaces of  $\mathcal{H}$ .  $\mathcal{U}$  and  $\mathcal{V}$  are said to be orthogonal if for every element  $u \in \mathcal{U}$  and  $v \in \mathcal{V}$ , the inner product is zero:

$$\langle u, v \rangle = 0 \quad \text{for all } u \text{ in } \mathcal{U} \text{ and } v \text{ in } \mathcal{V}. \quad (3.15)$$

Since  $\mathcal{U}$  and  $\mathcal{V}$  are Hilbert spaces as well, they must each contain the zero-vector. Therefore, a condition of orthogonality is that  $\mathcal{U} \cap \mathcal{V} = \{0\}$ . If the union of these two orthogonal spaces gives  $\mathcal{U} \cup \mathcal{V} = \mathcal{H}$ , then  $\mathcal{V}$  is said to be the *orthogonal complement* of  $\mathcal{U}$ . This can be noted as  $\mathcal{U}^\perp = \mathcal{V}$ .

Let's consider a matrix  $\mathbf{U}$  and its columns  $\mathbf{u}_i$ . If all of its columns are orthogonal between each other, the set of columns is said to be orthogonal. This is equivalent to:

$$\langle \mathbf{u}_i, \mathbf{u}_j \rangle = 0 \quad \text{for } i \neq j \quad (3.16)$$

If, in addition, for every column  $\|\mathbf{u}_i\| = 1$ , then the matrix  $\mathbf{U}$  is said to be *orthogonal*. This means that the following is true :

$$\mathbf{U}\mathbf{U}^T = \mathbf{U}^T\mathbf{U} = \mathbf{I} \quad (3.17)$$

Therefore,  $\mathbf{U}^{-1} = \mathbf{U}^T$ .

### 3.6 Matrix's column space and rank

The column space of a matrix corresponds to the span of the set formed by all the columns of such matrix. With this in mind, it is possible to define the *rank* of the matrix. The rank corresponds to the minimum number of elements that are allowed to span the column space. For a matrix  $\mathbf{U} \in \mathbb{R}^{M \times N}$  that contains  $N$  column vectors, then  $\text{rank}(\mathbf{U}) \leq N$ . If the vectors in  $\mathbf{U}$  form a set that is linear independent, then  $\mathbf{U}$  is said to be fully ranked with  $\text{rank}(\mathbf{U}) = N$ . If the matrix is square and full-ranked, then its columns span the entire  $\mathbb{R}^M$  space.

### 3.7 Linear transformations

A linear transformation  $\mathcal{T}$  is a mapping between two Hilbert spaces  $\mathcal{X}$  and  $\mathcal{Y}$  that is linear. This means that for  $x \in \mathcal{X}$  and  $y \in \mathcal{Y}$ ,  $\mathcal{T}(x) = y$ , and for scalars  $\alpha$  and  $\beta$  the mapping complies with the following:

$$\mathcal{T}(\alpha x_1 + \beta x_2) = \alpha \mathcal{T}(x_1) + \beta \mathcal{T}(x_2) = \alpha y_1 + \beta y_2 \quad (3.18)$$

In the special case of  $\mathcal{X} = \mathbb{R}^N$  and  $\mathcal{Y} = \mathbb{R}^M$ , it is possible to define the transformation as a matrix. This is  $\mathcal{T} = \mathbf{T}$  where  $\mathbf{T} \in \mathbb{R}^{M \times N}$ . This is normally written as:

$$\mathcal{T}(\mathbf{x}) = \mathbf{T}\mathbf{x} = \mathbf{y} \quad (3.19)$$

This matrix maps then vectors in  $\mathcal{R}^N$  space into the  $\mathcal{R}^M$  space. This is discussed here as it can be used to model several processes such as sampling and convolution. In particular, the convolution stated in Equation 2.5 can be treated as a matrix multiplication as it will shown later in Section 3.8. Using multiplication, the convolution becomes a linear mapping between the discrete real space and the discrete image space. The dimensions of the former will be defined by the sampling rate considered to model the real space, while the later will depend on the number of sensors.

There are 4 important subspaces associated with a linear transformation  $\mathcal{T} : \mathcal{U} \rightarrow \mathcal{V}$  [44]:

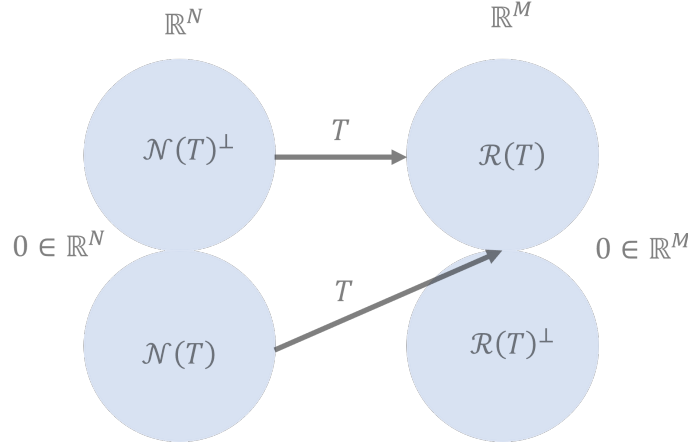


Figure 3.1: Relation between the 4 subspaces induced by a linear transformation  $\mathcal{T}$ . Figure adapted from [44].

- Null space (kernel)  $\mathcal{N}(\mathcal{T})$
- Orthogonal complement of the null space  $\mathcal{N}(\mathcal{T})^\perp$
- Range space  $\mathcal{R}(\mathcal{T})$
- Orthogonal complement of the range space  $\mathcal{R}(\mathcal{T})^\perp$

The focus now is on the case when  $\mathcal{T} = \mathbf{T}$ . In this case,  $\mathcal{N}(\mathbf{T})$  corresponds to the subspace given by all the vectors  $\mathbf{u}$  such that:

$$\mathbf{T}\mathbf{u} = \mathbf{0} \quad (3.20)$$

Here  $\mathbf{0} \in \mathbb{R}^M$  represents the zero-vector.  $\mathcal{N}(\mathcal{T})^\perp$  corresponds then to all the elements that produce a non-zero vector. A similar analysis can be done with the range. However, it is possible to be even more explicit. The range is given by all the values  $\mathbf{y}$  that can be obtained through the linear combination of the columns of  $\mathbf{T}$ . Its complement is then all the values that cannot be obtained from the linear transformation. Figure 3.1 presents a visualization of the relation between these spaces.

### 3.8 Image formation as matrix multiplication

In Equation 2.5 the image formation was presented as a convolution between the image  $I_i(x, y)$  and the PSF  $G(x, y)$ . In the discrete case where the number of sensor elements is  $M$ , and the number of sample points is  $N$ , the same model can be written as:

$$\mathbf{i}_o = \mathbf{G}\mathbf{i}_i \quad (3.21)$$

The matrix  $\mathbf{G}$  contains in its columns the image produced by points located at different locations of the sample. The vector  $\mathbf{i}_i$  corresponds to the discretized sample space and  $\mathbf{i}_o$



corresponds to the image obtained in the sensor elements. The column  $i$  of  $\mathbf{G}$  is then the image that a point located in the position associated with  $i_o[i]$ . In this scenario, it is clear that  $\mathbf{G}$  acts as a linear operator. This allows the description of its subspaces in an intuitive way. In this case, the range space  $\mathcal{R}(\mathbf{G})$  corresponds to all the possible images that can be obtained with the optical system modeled by  $\mathbf{G}$ . Its orthogonal complement is then the images that are not possible to observe. On the real space given by the discrete domain, the complement of the kernel indicates all the samples that can be sampled with the system. The kernel on the other hand occurs with samples that are invisible. These are non-zero samples that produce the zero vector. This topic of invisible objects will not be covered further as it is not discussed in the work presented in this thesis.

An important aspect of  $\mathbf{G}$  is that in general it does not have an inverse and therefore the transformation is not invertible. This means that to perform the inverse operation for obtaining  $\mathbf{i}_i$  back is not trivial.

### 3.9 Singular value decomposition

Matrix decomposition, also known as matrix factorization, corresponds to expressing a given matrix as a product or sum of simpler and more structured matrices. It involves breaking down a matrix into its constituent parts, which can provide insights into its properties, structure, and relationships with other matrices. One of them is Singular Value Decomposition, referred to simply as SVD [45].

SVD decomposes a matrix  $\mathbf{A} \in \mathbb{R}^{M \times N}$  into the product of three matrices:  $\mathbf{U}$ ,  $\mathbf{S}$ , and  $\mathbf{V}$  such that  $\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^T$ .  $\mathbf{U}$  and  $\mathbf{V}$  are square orthogonal matrices belonging to  $\mathbb{R}^{M \times M}$  and  $\mathbb{R}^{N \times N}$  respectively.  $\mathbf{S}$  is a rectangular diagonal matrix in  $\mathbb{R}^{M \times N}$  whose diagonal elements are called the *singular values* of  $\mathbf{A}$ . The singular values have the property that are always non-negative. Any rectangular matrix can be decomposed using SVD which is why this decomposition is quite useful.

Using the fact that  $\mathbf{S}$  it is possible to refer only to its diagonal as the vector  $\mathbf{s}$ . In that case, the decomposition can be written explicitly as the following weighted sum :

$$\mathbf{A} = \sum_{i=1}^{\min(M,N)} s[i] \mathbf{u}_i \mathbf{v}_i^T \quad (3.22)$$

This expression shows explicitly that every column of  $\mathbf{U}$  and  $\mathbf{V}$  are linked through a single singular value forming a tuple  $(s, \mathbf{u}, \mathbf{v})$ . It also shows that their contribution can be measured by such value since larger values will contribute more information to form  $\mathbf{A}$ .

One of the applications of SVD is to find the rank of a matrix. This can be done by only analyzing the singular values in  $\mathbf{s}$  as the rank is given by the number of non-zero singular values. The algorithm for obtaining SVD normally sorts the singular values in descending order. Therefore, the vectors  $\mathbf{u}_i$  and  $\mathbf{v}_i$  are also sorted in terms of importance. This is helpful when analyzing their meaning in terms of the column and row space of the matrix

**A.** In practical terms, this means that not all the vectors in  $\mathbf{U}$  and  $\mathbf{V}$  may be needed but only the first few. This means that:

$$\mathbf{A} = \sum_{i=1}^{\text{rank}(\mathbf{A})} s[i] \mathbf{u}_i \mathbf{v}_i^T \quad (3.23)$$

From the decomposition itself, it is possible to conclude that the column space is given by the span of the set formed by the vectors  $u$  with non-zero singular values. The same follows for the row space which can be written as the span of the vectors  $v$  with non-zero singular values.

SVD is tightly linked to a decomposition referred to normally as *eigendecomposition* which is designed to work only on square matrices. The eigendecomposition of a square matrix  $\mathbf{B}$  is given by:

$$\mathbf{B} = \mathbf{Q} \mathbf{\Lambda} \mathbf{Q}^{-1} \quad (3.24)$$

$\mathbf{Q}$  is an orthonormal matrix whose columns are normally referred to as *eigenvectors*.  $\mathbf{\Lambda}$  is a diagonal matrix whose values at the diagonal are referred to as the *eigenvalues*. Using that  $\mathbf{A} = \mathbf{U} \mathbf{S} \mathbf{V}^T$  then SVD and eigendecomposition can be shown to be related as:

$$\mathbf{A} \mathbf{A}^T = \mathbf{U} \mathbf{S} \mathbf{V}^T \mathbf{V} \mathbf{S} \mathbf{U}^T = \mathbf{U} \mathbf{S} \mathbf{U}^T = \mathbf{Q} \mathbf{\Lambda} \mathbf{Q}^{-1} \quad (3.25)$$

Therefore, if  $\mathbf{B} = \mathbf{A} \mathbf{A}^T$ , then  $\mathbf{U} = \mathbf{Q}$  and  $\mathbf{S}^2 = \mathbf{\Lambda}$ .

# Chapter 4

## Computational microscopy

Microscopes are based on the utilization of optical setups to capture the information provided by the light coming from a sample and redirect it toward the imaging plane, which can be the retina or a digital sensor. The advantage of a digital sensor, however, is that information can be permanently stored in digital format. Visualization can be done as a subsequent process in which the digital information is then displayed on a screen, potentially accompanied by data processing. For example, applying a clipping function can increase the contrast between faint structures or reduce intensities in the background. It is also possible to use a mapping function that allows for the normalization of intensities across the image.

Currently, computers are more than mere visualization tools. They are active elements in the digital image formation process, used to *reconstruct* microscope images from *raw* measurements obtained through an optical system, most commonly via a camera sensor.

Some examples of computational methods include quantitative phase imaging (QPI) [46], in which the raw data are interferograms of the sample. The final image is a phase map of the sample which is obtained through a phase retrieval computational algorithm.

A similar process occurs in digital ptychography [47], where the raw data corresponds to different images of the sample where each has been obtained with illumination at a different angle. These intensity measurements are then combined computationally to estimate the phase of the light passing through the sample.

Another example is light field microscopy [48], a technique based on adding an array of microlenses at the end of the optical path. The raw data corresponds to a grid of images, each showing a different angle of the sample. This single image is then processed to obtain different perspectives of the sample and even refocusing as a post-processing step.

This chapter focuses on computational techniques that can be applied to conventional fluorescence microscope systems with widefield illumination. However, it's important to note that these techniques are not limited to such systems.

The chapter will begin by introducing the framework of an inverse problem and then proceed to describe various methods within this context. It briefly covers deconvolution and single localization microscopy. Subsequently, it delves into methods that utilize tem-

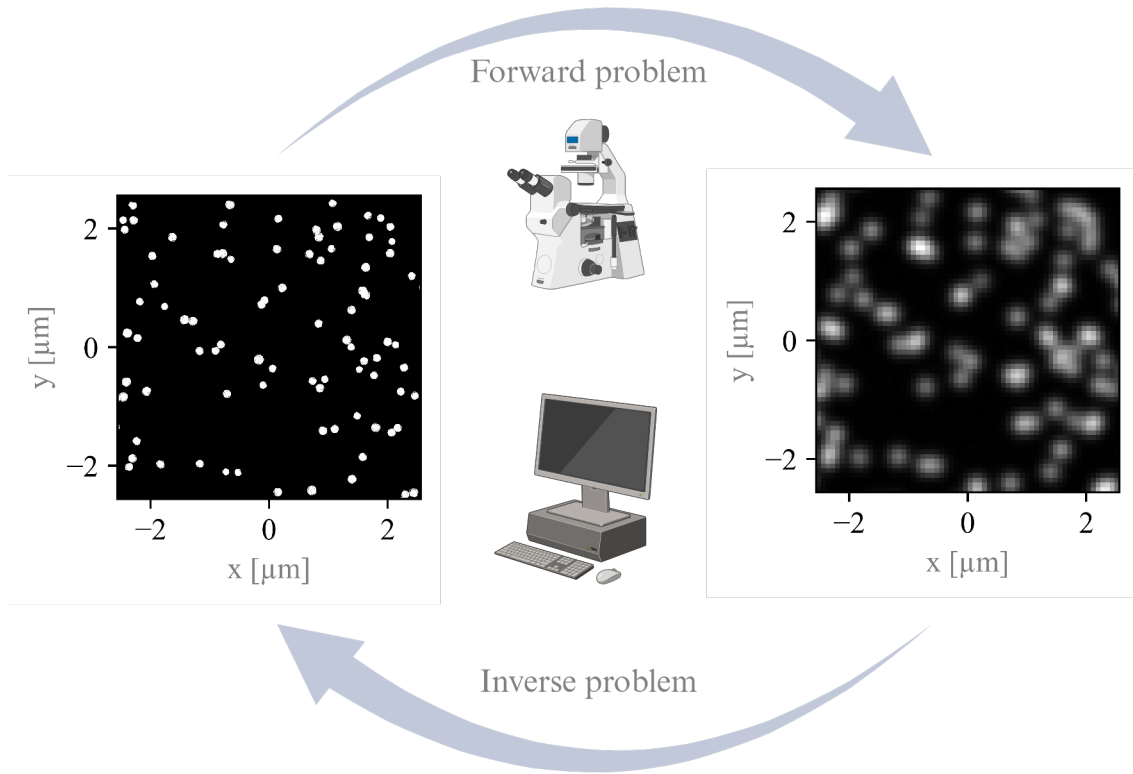


Figure 4.1: Forward and inverse imaging problem in widefield fluorescence microscopy, Forward problem: from a known distribution of fluorophores the problem is to predict the image acquired with the microscope. Inverse problem: from a diffracted limited measurement, predict the original configuration of fluorophores.

poral information encoded in the data, including SOFI, SRRF, SACD, and MUSICAL. A concise explanation of MUSICAL will be provided, with detailed information reserved for Chapter 5.

## 4.1 The inverse problem

Section 2.3 presents the imaging process as a convolution. As long as the Point Spread Function (PSF) is known, this allows us to predict the process of imaging a known sample for a given optical system. This simple model also explains why systems are diffraction-limited, as the PSF imposes a band-pass filter that attenuates high-spatial frequency components.

The situation just described corresponds to the *forward problem*. In contrast, the *inverse problem* considers the challenge of determining what the input looks like based on the measurements taken.

In the case of the imaging system discussed in this thesis, the unknown corresponds to the spatial distribution of fluorescence emission, with the raw data representing the diffraction-limited images.

On a discrete system as seen before in Section 3.8, the imaging process can be modeled

as a matrix multiplication (Equation 3.21). This model suggests that it may be possible to find  $\mathbf{i}_i$ . However, the loss of high frequencies due to diffraction makes  $\mathbf{G}$  non-invertible. Computational microscopy can be defined as the methods designed to estimate the sample  $i_i$  from the diffracted-limited image  $i_o$ .

The methods discussed here are designed to enhance the resolution of the raw data beyond the limits imposed by the Point Spread Function (PSF). Therefore, they can be categorized as *super-resolution* methods. Figure 4.1 illustrates the difference between the forward and inverse problems in microscopy.

## 4.2 Sample simulation

In order to assess the quality of reconstructions of different methods, it is desirable to have an image of the true distribution of emitters. This is normally referred to as *ground truth*. One way of obtaining this is by means of simulation. This procedure was used for Paper II. In this section we describe briefly this procedure in addition to two tools that can be used with this purpose. These tools share the property that can simulate intensity fluctuation datasets and therefore are suitable for the intensity fluctuation-based discussed in this chapter.

There exist several tools that allow the simulation of fluorescence microscopy data but not all of them implement fluctuations in the number of photons emitted through time which is key for the intensity fluctuation-based method. One tool that implements this type of simulation are SOFI simulation tool [49]. This tool is open-source and available for MATLAB platform. It allows the simulation of different structures from a selection of different 2D distributions. The user then needs to select a structure and some optical parameters such as pixel size and density of pixels. A different alternative is TestSTORM [50]. Similar to the previous tool, this software is open-source and available for MATLAB. However, it presents several advantages such as 3D structures, different models of PSF, and far more parameters regarding the fluorophore property. The example used in this section was created using TestStorm.

Both simulators present a similar model for simulating the transition between states. They use two variables given by the user. These correspond to the average time a fluorophore may stay in the ON state and OFF state. The duration in each state is modeled as random variable following a log-normal distribution. Two parameters are used to control the mean duration of each state, namely  $\tau_{ON}$  and  $\tau_{OFF}$  in Paper III. The rough illustration of the sequence is shown in Figure 4.2 where final sequence contains the number of units that each state lasts (this example omits photobleaching which corresponds to a permanent off-state). The final number of photons depends on the rate but this can be assumed constant. Once the number of photons is known for each frame, it is possible to positionate the fluorescent molecule in the image using the PSF of the system.

For paper II, this procedure was coded in Python to create noise free images. To simplify

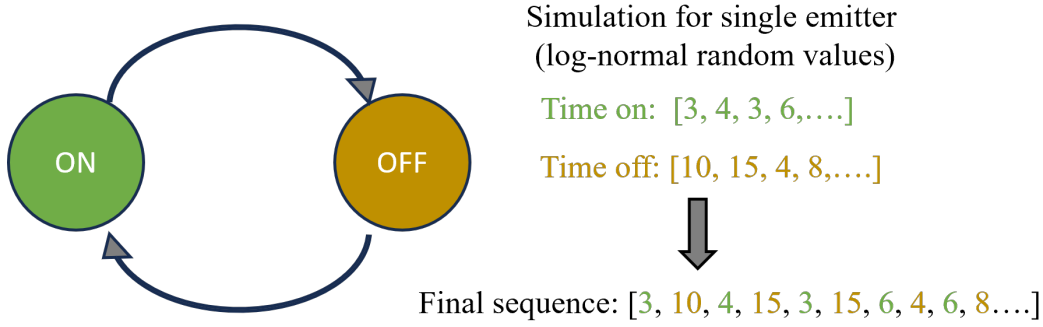


Figure 4.2: Single emitter's ON-state and OFF-state simulation. A single emitter can be either ON or OFF. To simulate the transitions the time on each state can be modeled using a log-normal distribution. These values are always positive and allow to use the average time that each state lasts. In green is shown a sequence of ON-state duration. Brown shows a sequence for OFF-states. The final sequence is the combination of both.

the analysis, photobleaching was not considered in the simulation. The final image, requires the addition of noise. This was done through the Poisson random distribution as follows. However, before the noise is added, the signal-to-background (SBR) is used to scale the image:

$$I^N(x, y, t) = (H - L) \frac{I(x, y, t) - \min I(x, y, t)}{\max I(x, y, t) - \min I(x, y, t)} + L \quad (4.1)$$

The scaled image  $I^N(x, y, t)$  has now minimum value  $L$  and maximum value  $H$ . These parameters are given by the user and allow the explicit definition of the SBR as  $\frac{H}{L}$ . The final step consists in incorporating noise through the Gaussian distribution where each pixel is sampled from a Poisson distribution with mean given by the noise-free image.

### 4.3 Deconvolution

As the name implies, deconvolution is the process of reversing the effects of convolution. The most straightforward method involves working in the frequency domain, which allows us to transform the convolution operation in space into a multiplication operation in the frequency domain.

In this scenario, it may appear straightforward to find an estimate  $\tilde{I}_i(u, v)$  since:

$$\tilde{I}_i(u, v) = \frac{I_o(u, v)}{G(u, v)} \quad (4.2)$$

This expression is well-defined as long as  $G(u, v) \neq 0$ , corresponding to the region where the filter permits information to pass. However,  $G(u, v)$  is band-limited, and therefore, deconvolution is still constrained by diffraction.

Another approach to deconvolution that does not rely on the Fourier domain involves optimization. In this setup, the problem can be formulated as the following minimization problem:

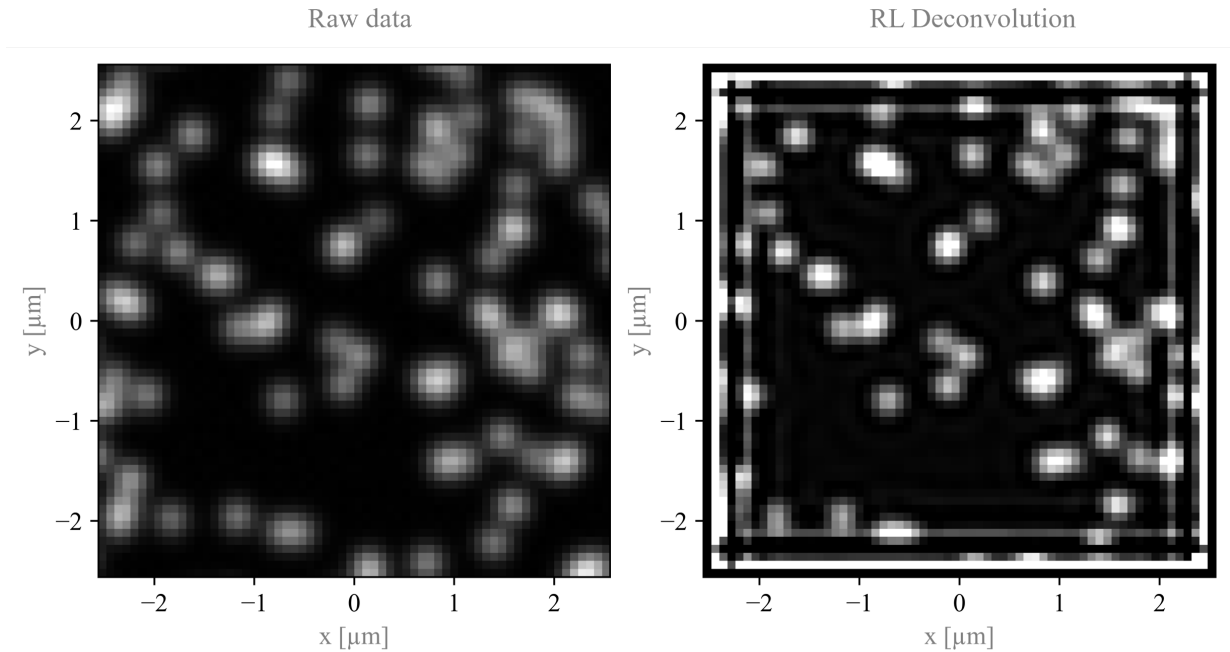


Figure 4.3: Deconvolution applied to a microscopy image. The left image shows the simulated raw data. The right image is result of applying Richardson-Lucy deconvolution. The result shows a slight improvement in the structural information but resolution gain is not clear.

$$\tilde{I}_i(x, y) = \arg \min_{\tilde{I}_i} (||G(x, y) * I_i(x, y) - I_o(x, y)||^2) \quad (4.3)$$

This is a simple expression that does not take into account either noise or possible regularization factors. More complex functions can be used depending on the knowledge of the sample, such as sparsity and continuity.

Some methods do consider noise in their models. Examples include Weiner deconvolution [51] and Richardson-Lucy deconvolution [52, 53]. The first method works well under Gaussian noise and can be performed in a single step in the Fourier domain. In contrast, Richardson-Lucy is an iterative algorithm that operates directly in the spatial domain and works effectively for Poisson noise.

Up to this point, these expressions have assumed knowledge of the PSF. Methods that enable deconvolution without prior knowledge of the PSF are termed *blind*. In this context, blind means that they do not require exact knowledge of the PSF as the process itself estimates both the PSF and the true image.

Figure 4.3 presents an example of deconvolution. This corresponds to a simulated sample of vesicles where the PSF is known. The deconvolution algorithm corresponds to Richardson-Lucy with 30 iterations. The contrast is improved making some clusters of sources easier to distinguish.

## 4.4 Single molecule localization

Single molecule localization (SML) corresponds to a family of methods that rely on converting the information acquired through multiple diffracted-limited images into a super-resolved image. These methods typically require widefield illumination of the sample to induce fluorescence, followed by the estimation of the positions of individual point sources [4]. The achievement of super-resolution in SML relies on the high-precision estimation of the position of a single emitting source, as long as its Point Spread Function (PSF) does not overlap with others. In other words, as long as the PSFs of different emitters do not overlap at a specific instant, they can be resolved as independent sources, even if their separation is below the limit imposed by a system constrained by diffraction. This can be achieved if, during the exposure time, only a small fraction of the total population of sources is active. This is referred to as the *ON* state, indicating that the fluorophore is actively emitting light. This contrasts with the *OFF* state, in which a fluorophore cannot emit fluorescence. Given the large number of fluorophores, with ideally only a small subset imaged at a time, the number of required images typically ranges in the order of thousands. The various SML methods differ in the type of fluorophores they utilize, which in turn defines how their states can be controlled. For example, photoactivatable localization microscopy (PALM) [54] uses UV illumination to activate the fluorescent emitters (fluorescent proteins) attached to the sample. Similarly, stochastic optical reconstruction microscopy (STORM) [55] uses synthetic dyes on a special buffer solution that allows the dyes to photoswitch. A third example with a slightly different working mechanism is point accumulation in nanoscale topography (PAINT) [56]. This technique uses free-diffusing fluorescent emitters that move around the sample until they collide with the structure of interest. While in motion the emitters cannot form an image due to their speed compared to the exposure time (*OFF* state). However, upon collision, they stop moving which allows the formation of the emitter's image (*ON* state).

In terms of reconstruction, the images are typically processed individually in a workflow that may involve denoising and background suppression, followed by localization. The localization step can be achieved through methods like pattern matching of the PSF, identifying local minima, or employing more complex algorithms. The final result is presented as a table of localizations, which can then be converted into an image by applying binning.

The concept can be also applied to the reconstruction of 3D images. This requires the introduction of a PSF that encodes the axial location of the fluorescent source in a non-ambiguous manner making it easier to extract its location in all the spatial dimensions. This can be done through astigmatism, dual helix or by increasing the amount of focal planes.



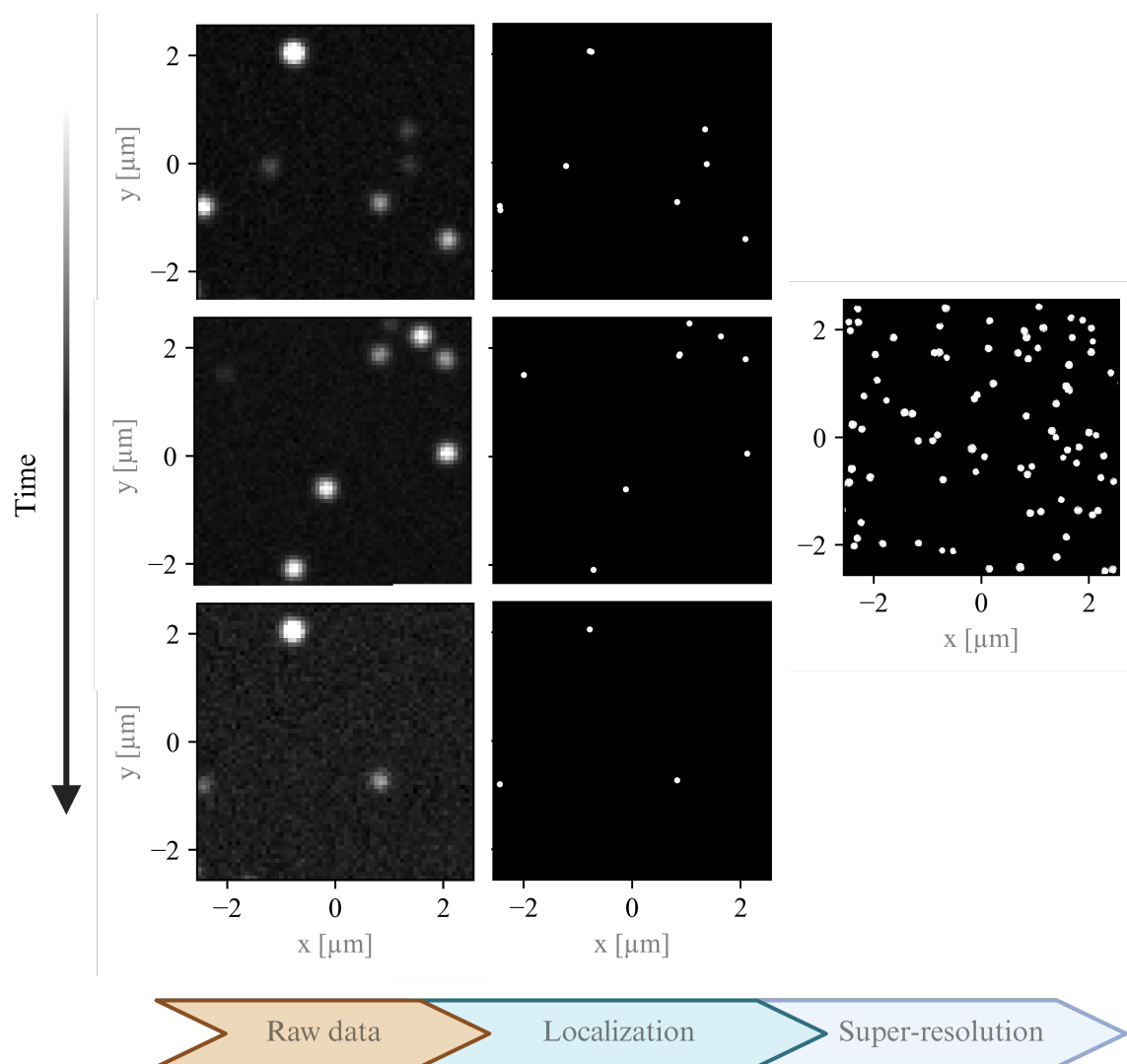


Figure 4.4: Single molecule localization microscopy principle (SML). If the number of sources is small then localization of every single frame is not difficult. The localization step finds the position of every emitter that is active on that particular frame and requires the analysis of thousands of images. The final step is the combination of all the localizations.

## 4.5 Intensity fluctuation based methods

SML methods rely on high spatio-temporal sparsity between emitters which enables localization of individual emitters. This requires specific dyes and imaging conditions. However, in general, this does not make explicit utilization of the natural temporal fluctuation in the number of photons produced by fluorescent emitters. An example of the variation in the intensity emission of a collection of fluorophores is given in Figure 4.5. Methods that exploit the temporal fluctuation in the intensity of emission of the fluorophores in the sample can achieve super-resolution with fewer images and are less dependent on spatiotemporal sparsity.

The inclusion of a temporal variable means that images are a function of time as well. Therefore, the image of a sample at any particular moment can be modeled:

$$I_o(x, y, t) = G(x, y) * I_i(x, y, t) \quad (4.4)$$

An important observation is the  $G$  does not depend on time and therefore the convolution occurs only in the spatial domain. Combining this with Equation 2.17, a fluorescent sample with  $N$  emitters can be expressed as:

$$I_o(x, y, t) = G(x, y) * \sum_{e=1}^N b_e(t) \delta(x - x_e, y - y_e) \quad (4.5)$$

Using the sifting property of the impulse, the image can be expressed also as:

$$I_o(x, y, t) = \sum_{e=1}^N G(x - x_e, y - y_e) b_e(t) \quad (4.6)$$

Similarly, using matrix multiplication, this can be written as:

$$\mathbf{I}_o = \mathbf{GB} \quad (4.7)$$

In this expression, each column of  $\mathbf{I}_o$  corresponds to the image taken at a particular time. Therefore, is it termed the *image stack matrix*. Regarding  $\mathbf{B}$ , this contains the temporal intensities of all the  $N$  emitters. For  $T$  images this means that these matrices are defined as:

$$\mathbf{I}_o = [\mathbf{i}_o^1 \ \mathbf{i}_o^2 \ \cdots \ \mathbf{i}_o^T] \quad \mathbf{B} = \begin{bmatrix} b_1(1) & b_1(2) & \cdots & b_1(T) \\ b_2(1) & b_2(2) & \cdots & b_2(T) \\ \vdots & \vdots & & \vdots \\ b_N(1) & b_N(2) & \cdots & b_N(T) \end{bmatrix} \quad (4.8)$$

Each column of  $\mathbf{I}_o$  is termed a *frame*.

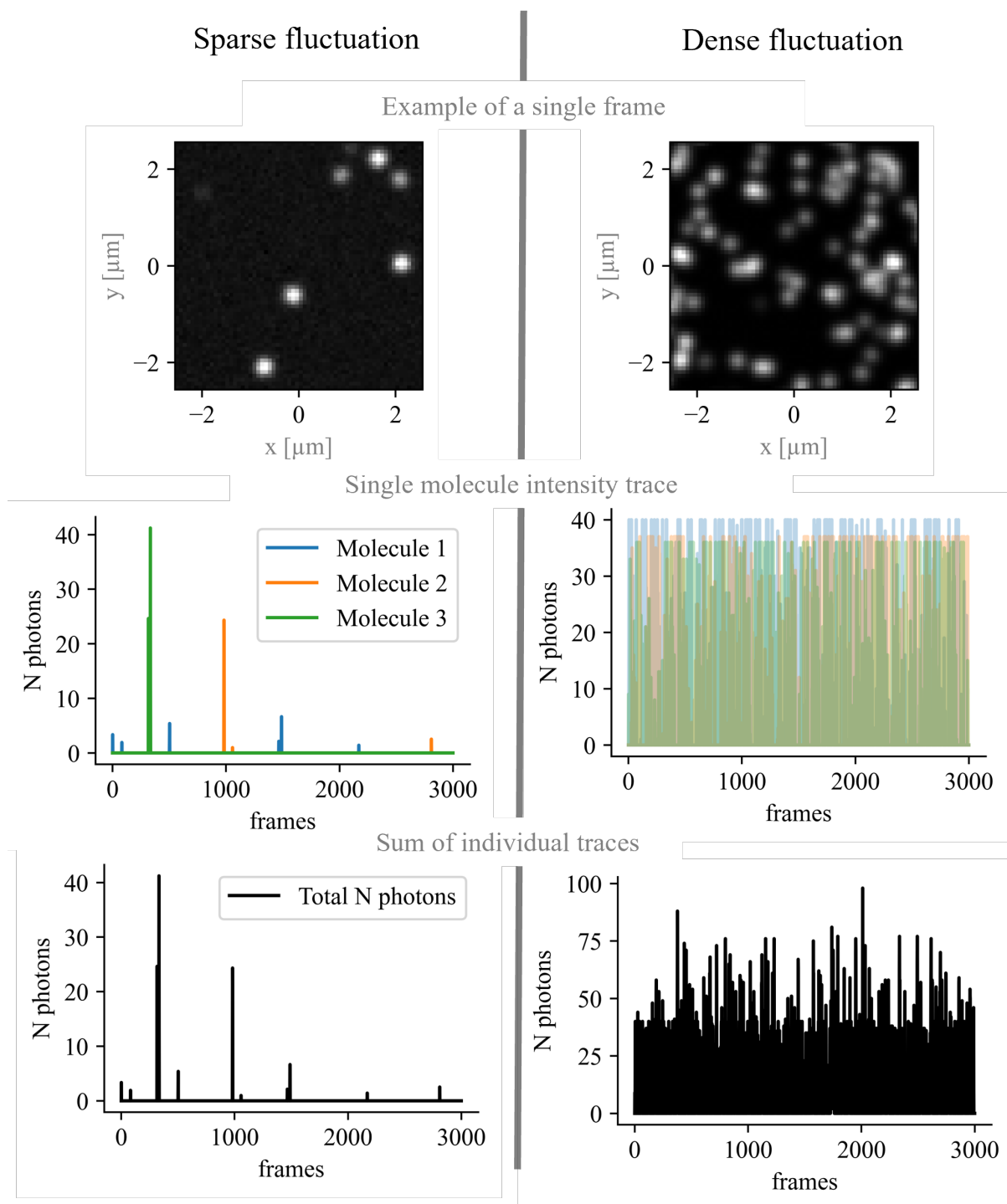


Figure 4.5: Sparse and dense fluctuations. This simulated example illustrates the temporal photon emission profile for 3 different emitters. The sum can be understood as the measurement obtained at a single pixel close to the three of them. If the sample is sparse enough then SML methods suffice. However, for dense samples intensity fluctuation-based methods behave better.

### 4.5.1 SOFI

Super-resolution optical fluctuation imaging (SOFI) was published in 2009 [6], cementing the floor for algorithms that exploit photo-switchable fluorophores. Similar to single-molecule localization, SOFI requires a stack of images where the following properties need to be met:

- The fluorophores exhibit two optically distinguishable states.
- The transition between states occurs independently between emitters.
- The pixel size must be smaller than the PSF of the system.

The algorithm is based on a descriptive quantity for random distributions called *cumulants* [57]. Similar to moments, cumulants allow the description of random variables and the relation between several of them. SOFI uses cumulants and cross cumulants as the key processing step to effectively reduce the size of the PSF. The method is blind to the PSF and allows the increase of the resolution by several times depending on the order of the used cumulant.

In the simplest version, SOFI uses the second-order cumulant over the content of each individual pixel trace. This allows the reduction of the explanation of the method as the second-order cumulant coincides with the variance. The raw data corresponds to the images given by  $I_o(x, y, t)$ , this is, the image obtained at the sensor. From there, the processing part starts with a normalization of the temporal information. Assuming a sample formed by  $N$  temporal fluctuating fluorescent sources as in Equation 4.6, then the normalized image  $\tilde{I}_o(x, y, t)$  is given by:

$$\tilde{I}_o(x, y, t) = I_o(x, y, t) - \langle I_o(x, y, t) \rangle_t \quad (4.9)$$

$$= \sum G(x - x_e, y - y_e)(b_e(t) - \langle b_e(t) \rangle_t) \quad (4.10)$$

$$= \sum G(x - x_e, y - y_e)\tilde{b}_e(t) \quad (4.11)$$

The tilde on top is used to denote that the variable is normalized so its expected value is zero. The next step is to obtain the temporal variance of each pixel in  $\tilde{I}_o(x, y, t)$ . Since the variable is normalized it becomes just the expectation:

$$\text{Var}\{\tilde{I}_o(x, y, t)\} = E\{\tilde{I}_o(x, y, t)^2\} \quad (4.12)$$

$$= \sum_e \sum_{e'} G(x - x_e, y - y_e)G(x - x_{e'}, y - y_{e'})E\{b_e(t)b_{e'}(t)\} \quad (4.13)$$

The assumption of independence means that  $E\{b_e(t)b_{e'}(t)\} = 0$  for  $e \neq e'$ . Therefore, the expression can be simplified. What remains is the expression for the SOFI image of order 2:

$$I^{SOFI2}(x, y) = Var\{\tilde{I}_o(x, y)\} = \sum_e G(x - x_e, y - y_e)^2 E\{b_e(t)^2\} \quad (4.14)$$

The SOFI image of order 2 is equivalent to the convolution of the ideal  $I_i$  by an equivalent PSF that is the square of the original which results in an increased resolution. The expression above describes a resolution of  $\sqrt{2}$ . This can be improved by deconvolution and then convolution. For second-order SOFI this can be expressed as follows [58, 59]:

$$I_L^{SOFI}(x, y) = \mathcal{F} \left\{ I_L^{SOFI}(u, v) \frac{G(2u, 2v)}{G(u, v) * G(u, v) + \alpha} \right\} \quad (4.15)$$

The expression indicates a deconvolution by the PSF square, followed by a convolution with a PSF that is half the original. This process then makes the relation between the increase in resolution and the order of the SOFI image linear. The process is similar for higher-order cumulants as well.

The process just described does not explain how subpixelation is achieved. This is done through the concept of cross-cumulants that estimate the cumulant in a virtual location. It is worth mentioning that a more complex version of SOFI also exists called balance SOFI (bSOFI) [60] which combines the information of multiple SOFI images from different orders. For a more detailed description of the algorithm, the interested reader is referred to [61]. It is worth mentioning that a more complex version of SOFI also exists called balance SOFI

## 4.5.2 MUSICAL

Multiple Signal Classification algorithm (MUSICAL) was published in 2016 [11]. Similar to SOFI, it enables the enhancement of image resolution by consolidating information from different frames through matrix decomposition.

The procedure begins with Equation 4.7. The first step involves decomposing the image stack matrix using Singular Value Decomposition (SVD):

$$\mathbf{I}_o = \mathbf{GB} = \mathbf{USV}^T \quad (4.16)$$

As discussed in Section 3.9,  $\mathbf{U}$  encodes the column space of  $\mathbf{I}_o$ , and  $\mathbf{V}$  encodes the row space. Therefore, it is stated that  $\mathbf{U}$  encodes the spatial information, while  $\mathbf{V}$  encodes the temporal information.

MUSICAL is currently designed to exploit the information from  $\mathbf{U}$  and  $\mathbf{S}$ , leaving  $\mathbf{V}$  aside. This means that extracting meaningful information from  $\mathbf{V}$  is still an open challenge. More on this will be discussed in Section 7.

In the context of MUSICAL, the columns of  $\mathbf{U}$  are referred to as the *eigenimages*, and they contain the structural information available in the stack. The values in the diagonal of  $\mathbf{S}$  are referred to as the *singular values* or *eigenvalues*, depending on the algorithm used for the decomposition (See Section 3.9 for a discussion about the similarities between SVD

and eigendecomposition). The set of eigenimages will be denoted as  $\mathcal{U}$ .

From Equation 4.7, it is possible to observe that in the ideal case, the columns of  $\mathbf{I}_o$  belong to the span of the columns of  $\mathbf{G}$ . Simultaneously, the same columns belong to the span of  $\mathbf{U}$ .

The assumption in MUSICAL is that it is possible to separate the eigenimages into two disjoint sets, namely the *signal* and *noise* sets. These sets, in turn, span the signal and the noise spaces, respectively. The signal set is denoted as  $\mathcal{U}_S$  and the noise set as  $\mathcal{U}_N$ . Since  $\mathbf{U}$  is orthonormal then all the elements of  $\mathcal{U}$  are mutually orthogonal which means that  $\mathcal{U}_S$  is orthogonal to  $\mathcal{U}_N$ .

The signal space corresponds to all the possible images that can be generated from the underlying structural information contained in the sample, regardless of the temporal fluctuations. Everything that is not part of the signal space belongs to the noise space and represents the unwanted content in the measurements.

Assuming that the separation between the signal and the noise is indeed possible, then the signal space is equivalent to the span of the columns of  $\mathbf{G}$ , which implies that the space spanned by  $\mathbf{G}$  is also orthogonal to every element of the noise space.

This fact is used to define MUSICAL's *indicator function*. To do this, it is necessary to first define the real-valued vector  $\mathbf{g}_i$  related to the image induced by a theoretical source located at some coordinate indexed by  $i$ .

For a system defined by the Point Spread Function (PSF)  $G(x, y)$ , a sensor array with  $M$  elements (pixels) whose central coordinates are  $(x_j, y_j)$ , and an emitter located in the sample plane at  $(x_i, y_i)$ , a test point's image is given by:

$$\mathbf{g}_i = \begin{bmatrix} G(x_1 - x_i, y_1 - y_i) \\ G(x_2 - x_i, y_2 - y_i) \\ \vdots \\ G(x_M - x_i, y_M - y_i) \end{bmatrix} \quad (4.17)$$

With this, it is possible to define an *indicator function* that can be evaluated at any location  $(x_i, y_i)$  in the image space:

$$f(x_i, y_i) = \sqrt{\frac{\sum_{\mathbf{u} \in \mathcal{U}_S} |\mathbf{u} \cdot \mathbf{g}_i|^2}{\sum_{\mathbf{u} \in \mathcal{U}_N} |\mathbf{u} \cdot \mathbf{g}_i|^2}}^\alpha \quad (4.18)$$

In this expression,  $\alpha$  is a parameter that controls the contrast. However, it will not be discussed much in detail as for most practical cases this can be set to 4 as recommended in [11].

For a test point located in a position that matches the position of an actual emitter, the denominator of this function will tend to be zero. This occurs because, for such a test point,  $\mathbf{g}_i$  will reside in the signal space. As a consequence,  $\mathbf{g}_i$  will be orthogonal to the noise space. Therefore,  $|\mathbf{u} \cdot \mathbf{g}_i| = 0$ . This means that if the test point is correct, the indicator function will tend to infinity.

The MUSICAL reconstruction is obtained by evaluating  $f$  at multiple locations in the sample space, which also allows subpixelation in the reconstruction. MUSICAL's processing workflow is based on processing smaller patches of the image called the "sliding window" and then stitching the reconstructed windows. The numerator of the indicator function contributes to the seamless stitching of these.

So far, the signal and noise sets have been defined. In practice, they need to be classified by the algorithm. In [11], this is done through the manual selection of a threshold. This threshold is directly applied to the singular values. With this procedure, the signal space is the one spanned by all the eigenimages with their corresponding singular values larger than the given threshold. The noise space is composed of the remaining eigenimages. As will be shown in Paper II and Paper III, this can be improved by using *soft-thresholding*.

MUSICAL is not a quantitative method, as the numerical results cannot be directly translated into physical information about the sample. However, it can be used to detect structures that are not visible with conventional means, providing increased resolution and improved contrast in fluorescent images. The reported resolution gain has ranged between 2 and 3 times the diffraction limit, with the initial report achieving a resolution of 50 nm.

### 4.5.3 SRRF

Super-resolution Radial Fluctuation (SRRF) [8] is an algorithm based on two parts. The first part involves the computation of the so-called *radiality maps* for every image in the stack. The second part is the combination of these maps.

Computing the radiality maps first requires the computation of the gradient  $\nabla I_o(x, y)$ . This can be expressed as estimating the following functions:

$$\nabla_x I_o(x, y) = \frac{\partial I_o(x, y)}{\partial x} = \frac{I_o(x + \Delta x, y) - I_o(x - \Delta x, y)}{2\Delta x} \quad (4.19)$$

$$\nabla_y I_o(x, y) = \frac{\partial I_o(x, y)}{\partial y} = \frac{I_o(x, y + \Delta y) - I_o(x, y - \Delta y)}{2\Delta y} \quad (4.20)$$

In this expression,  $\Delta x$  and  $\Delta y$  correspond to a fixed quantity that could represent the pixel size in the discrete domain. However, they can be considered equal to 1, provided that  $x$  and  $y$  correspond to pixel units. The gradient maps, when combined, represent the direction (as a vector) of the maximum intensity increase when looking at a particular spatial location.

$$\nabla I_o(x, y) = \begin{bmatrix} \frac{\partial I_o(x, y)}{\partial x} \\ \frac{\partial I_o(x, y)}{\partial y} \end{bmatrix} \quad (4.21)$$

The gradient maps are then interpolated to estimate values at arbitrary points and combined to construct the *radiality map*.

A radiality map is an estimation of the degree of convergence towards a point of interest, which includes first estimating the direction and then the magnitude of the convergence.

This point can be evaluated at any location, allowing for the creation of a finer grid than the one provided by the original image.

For a point of interest located at  $\mathbf{r}_i = (x_i, y_i)$ , the convergence is estimated by measuring whether the gradients at points surrounding  $\mathbf{r}_i$  point towards or away from that point. For  $N$  points with coordinates  $\mathbf{r}_n$ , they are given by:

$$\mathbf{r}_n = \mathbf{r}_i + r_s \begin{bmatrix} \cos(\frac{2\pi}{N}n) \\ \sin(\frac{2\pi}{N}n) \end{bmatrix} \quad \text{For } n = 0..(N - 1) \quad (4.22)$$

Here, the parameter  $r_s$  is specified by the user and corresponds to the radius of a ring.

From each point in the ring, the direction of convergence is measured in terms of the angle formed by the gradient and the vector  $\tilde{\mathbf{r}}_n = \mathbf{r}_n - \mathbf{r}_i$ :

$$\cos(\theta_n) = \frac{\nabla I_o(x_n, y_n) \cdot \tilde{\mathbf{r}}_n}{|\nabla I_o(x_n, y_n)| |\tilde{\mathbf{r}}_n|} \quad (4.23)$$

If  $\tilde{\mathbf{r}}_n$  lies parallel to  $\nabla I$  and both have the same direction, then  $\cos(\theta_n) = 1$ , and the gradient is said to be divergent. If they are opposite, then the gradient is convergent, and  $\cos(\theta_n) = -1$ . In [8], this angle is used only to check whether the gradient is converging or diverging. Additionally, it computes the distance between the point and the line projected by the gradient. The line can be expressed as:

$$(y - y_n) = \frac{\nabla_x I_o(x_n, y_n)}{\nabla_y I_o(x_n, y_n)} (x - x_n) \quad (4.24)$$

From this expression, the distance  $d_i$  is given by:

$$d_i = \frac{|(x_i - x_n)\nabla_y I_o(x_n, y_n) - (y_i - y_n)\nabla_x I_o(x_n, y_n)|}{\sqrt{\nabla_x I_o(x_n, y_n)^2 + \nabla_y I_o(x_n, y_n)^2}} \quad (4.25)$$

A value  $d_i = 0$  means that  $\mathbf{r}_i$  lies on the line of the gradient, either diverging or converging. Once the distance is obtained, the convergence from point  $n$  to center  $i$  is calculated as:

$$c_n = \text{sgn}(\cos(\theta_n)) \left(1 - \frac{d_i}{r_s}\right)^2 \quad (4.26)$$

The direction is given by the signum function  $\text{sgn}$  defined as 1 for non-negative values and -1 for negative. Therefore,  $c_i$  can assume values between  $-1$  and  $1$ , being 0 the case where the segment between  $\mathbf{r}_s$  and  $\mathbf{r}_c$  are perpendicular.

Finally, the radially at  $\mathbf{r}_i$  is obtained by averaging the convergence from all the points in the ring.

$$C_i = \frac{1}{N} \sum_{n=1}^N c_n \quad (4.27)$$

It is possible to further refine this value by weighting the magnitude of the gradient on a secondary ring, but that will not be covered in this thesis. The radially maps correspond to



a temporal stack of radial maps, which can be denoted as  $R(x, y, t)$  where  $(x, y)$  corresponds to the coordinate  $\mathbf{r}_i$  in the previous derivation. The final SRRF image is obtained through the combination of images in this stack. The available methods are:

- Temporal radiality averaging (TRA): Uses the average of the pixel maps:

$$I^{SRRF}(x, y) = \frac{1}{T} \sum_{t=1}^T R(x, y, t) \quad (4.28)$$

- Temporal radiality maximum (TRM): Uses the maximum of the pixel maps:

$$I^{SRRF}(x, y) = \max_t R(x, y, t) \quad (4.29)$$

- Temporal radiality pairwise product mean (TRPPM): Uses the following formula:

$$I^{SRRF}(x, y) = \frac{2}{T(1+T)} \sum_{t=1}^T \sum_{s=t}^T R(x, y, t)R(x, y, s) \quad (4.30)$$

- Auto-cumulant (TRAC): uses cumulants in a similar way to SOFI [6].

The radiality map aids in reducing the size of the PSF and serves as a deconvolution step and therefore may increase the the resolution slightly.

#### 4.5.4 SACD

The Super-resolution method using auto-correlation with two-step Lucy-Richardson deconvolution (SACD) [62] is based on a two-step deconvolution process. First, deconvolution is applied to each individual frame, and then auto-correlation is performed across the deconvolved frames, followed by a second deconvolution step using the theoretical Point Spread Function (PSF) after auto-correlation.

To achieve smaller pixel sizes, the images are preprocessed using Fourier interpolation. This involves applying a Fourier transform, padding the image with zeros at the edges, and then performing the inverse Fourier transform. Since the image is band-limited due to the PSF, adding zeros in the Fourier domain does not modify the content of the image. After padding, the image is transformed back to the spatial domain, resulting in an image that retains frequency information but has smaller pixels. A similar approach has been reported in [61] for SOFI images.

The primary advantage of this method is a notable reduction in the number of frames required for reconstruction. The authors have reported successful reconstructions using as few as 16 images.

## 4.6 Qualitative comparison

To conclude this section, a comparison of the methods in terms of processing steps and resulting images is included. This is not meant to be an exhaustive comparison but an illustration of their similarities and differences. A more extensive analysis can be found in [63].

Figure 4.6 displays the most critical steps for each of the methods discussed previously. The resulting images are shown in Figure 4.7 where the two points highlighted are separated by a distance close to 150 nm. The profile is shown in Figure 4.8. From the methods, only MUSICAL was able of obtaining such distance. However, it needs to be highlighted that SRRF has several parameters that can be optimized and SACD has a limit to the maximum number of frames.

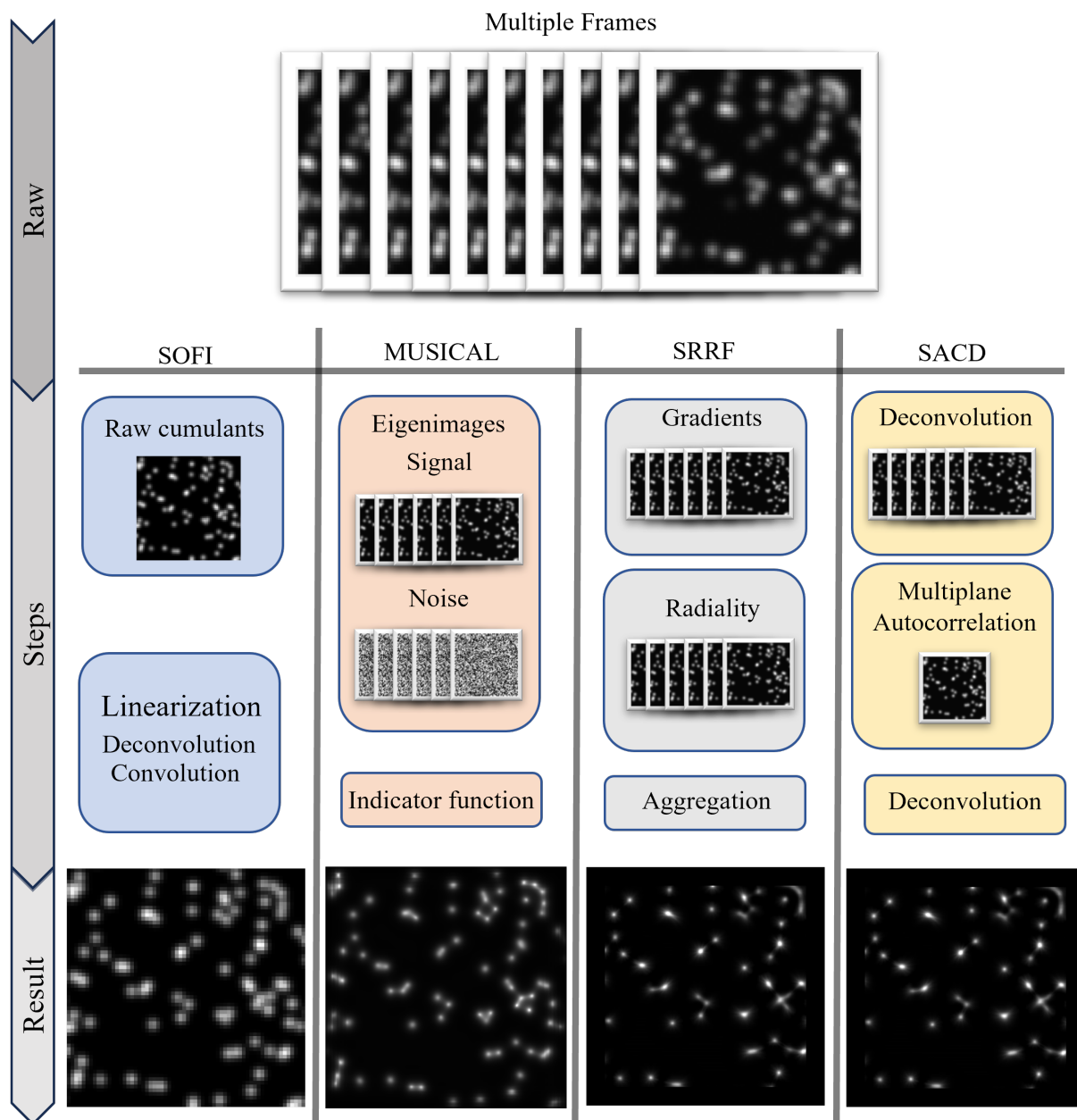


Figure 4.6: Fluorescence Fluctuation-based method: processing steps. All the methods start by considering all the frames in the raw data. The processing steps are shown sequentially downward.

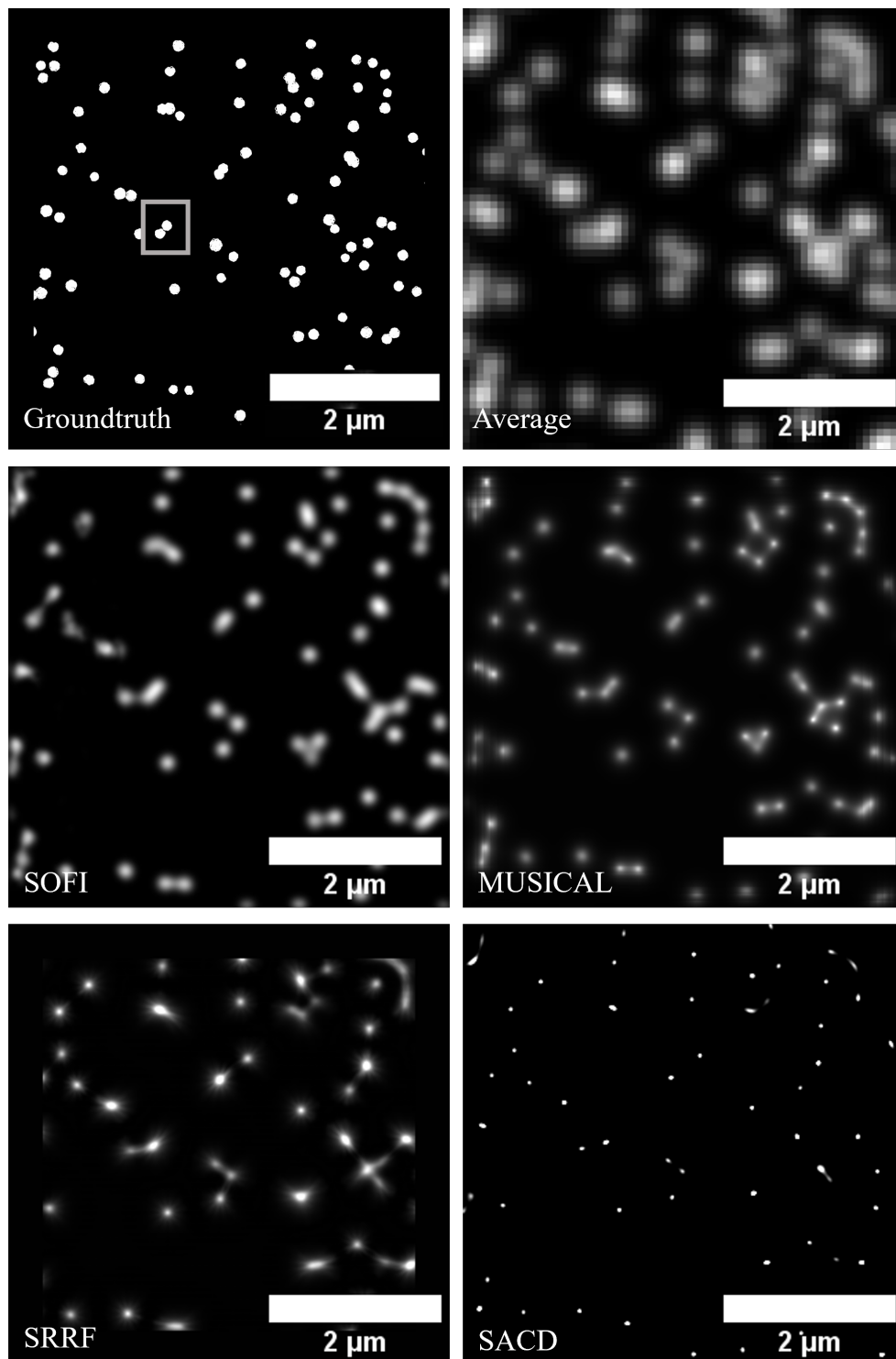


Figure 4.7: Fluorescence Fluctuation-based method: results. The different methods produce images that are qualitatively different. SOFI result was obtained with its MATLAB implementation and corresponds to bSOFI. MUSICAL reconstruction corresponds to soft MUSICAL, computed in a Python implementation. SRRF was obtained using all the default parameters. SACD was computed using 16 frames in its MATLAB implementation.

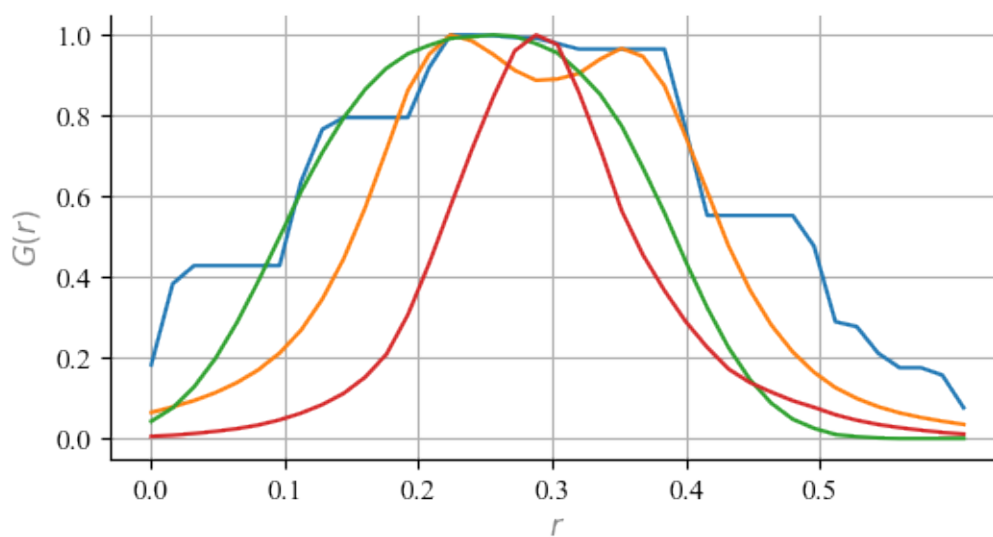


Figure 4.8: Fluorescence Fluctuation-based method: profile. This comparison shows the profile across the the 2 objects highlighted in Figure 4.7. Only MUSICAL was able of distinguishing the two objects.



# Chapter 5

## MUSICAL

The Multiple Signal Algorithm (MUSICAL) [11] is a super-resolution algorithm that leverages the temporal fluctuations in the emission of photons produced by each fluorescent emitter in the sample. It is based on the multiple signal classification method (MUSIC), which was originally published by Schmidt [13] as a technique for estimating the parameters of multiple wavefronts arriving at an array of sensors [64]. Subsequently, it was adopted by the electromagnetic community to detect the positions of point-like scatterers [65], a problem commonly referred to as *reverse scattering*. In the context of fluorescence microscopy, a similar approach can be employed by considering the emitters as the sources which is the key idea of MUSICAL as demonstrated by Agarwal.

The basic working principle was described previously in Section 4.5.2. Here, some of the information is repeated and then extended. However, before proceeding, it is necessary to explain the concept of a *sliding window*

MUSICAL can be explained using the entire field of view given by the measurement matrix  $\mathbf{I}_o$ . However, computers' memory is limited and therefore it is impractical to process the entire field of view at once. In practice, this would mean matrices far too big for conventional computers to handle. As a result, the processing strategy involves taking small sections of the image at a time. Each single section is referred to as a *sliding window* and MUSICAL is sequentially applied to every one of them. The size of the Airy disk of the estimated PSF gives the size of this by estimating the parameter  $N_w$  as:

$$N_w = 1 + 2\mathit{floor}\left(\frac{0.61\lambda}{NAw}\right) \quad (5.1)$$

In this expression,  $\lambda$  corresponds to the the wavelength of fluorescent emission,  $NA$  is the numerical aperture of the system and  $w$  is the pixel size. The units of  $N_w$  are pixels and by design, it is always odd. The function *floor* produces the greater integer that is equal to or lower than its argument.

The sliding window corresponds to an array that belongs to  $\mathbb{R}^{N_w \times N_w \times T}$  that can then be processed using MUSICAL as explained next. This is illustrated in Figure 5.1.

In addition, the algorithm applies a windowing function to avoid issues in the borders and aid in the stitching of the results. The windowing function conventionally corresponds

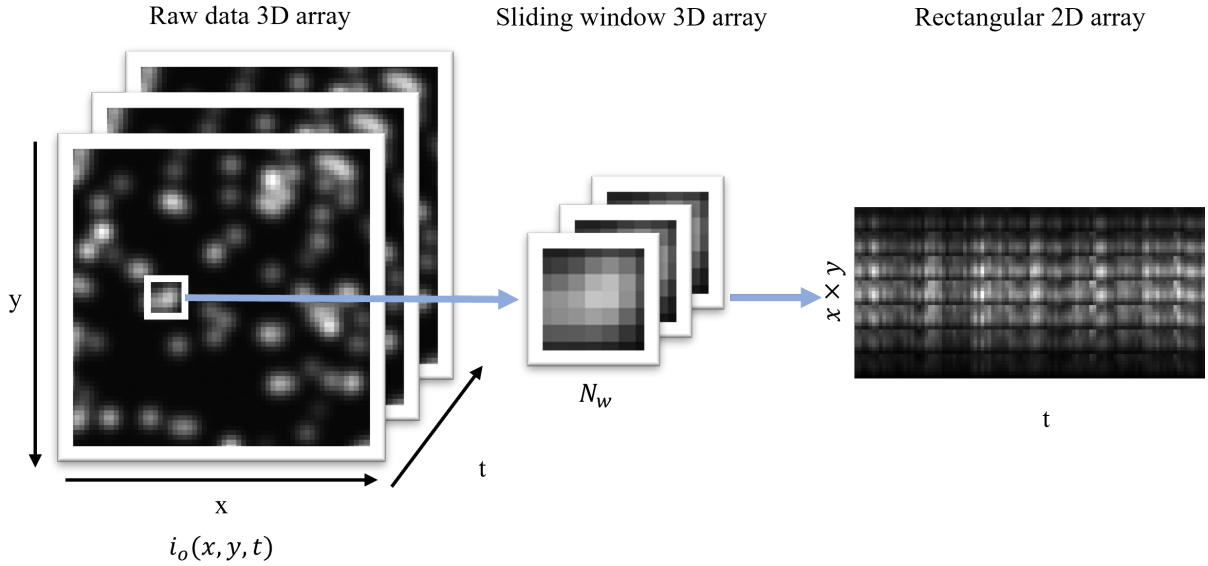


Figure 5.1: Sliding window. This example takes a slice of the raw data of size  $N_w$  by  $N_w$  in the  $x$  and  $y$  (the  $y$  arrows point downward following the order the image is accessed by a computer). The array is converted into a rectangular matrix where each frame has been converted into a column.

to a Gaussian function and it will not be discussed further in this thesis.

The workflow of MUSICAL's processing is given in Figure 5.2.

## 5.1 Imaging model and decomposition

In the previous chapter, it was introduced that the discrete imaging model of a microscopy temporal fluorescent image stack  $\mathbf{I}_o$  can be written as a matrix multiplication between the PSF given by  $\mathbf{G}$  and a known distribution denoted as  $\mathbf{B}$ . At the same time, it can be decomposed using SVD. This can be expressed as follows:

$$\mathbf{I}_o = \mathbf{G}\mathbf{B} = \mathbf{U}\mathbf{S}\mathbf{V}^T \quad (5.2)$$

In this representation it is clear that the columns of  $\mathbf{I}_o$  (the frames) are linear combinations of the PSF columns of  $\mathbf{G}$ . These columns are grouped in the set  $\mathcal{G}$  and therefore it is possible to conclude that every frame belongs to  $\text{span}(\mathcal{G})$ . The fluorescent sample is assumed to be known in the sense that the number of emitters and their respective locations is already encoded in the matrices in the indices used to refer to the column in the case of  $\mathbf{G}$ , and the row in the case of  $\mathbf{B}$ . For a camera with  $M$  sensor elements,  $N$  emitters and  $T$  frames,  $\mathbf{I}_o \in \mathbb{R}^{P \times T}$ ,  $\mathbf{G} \in \mathbb{R}^{P \times N}$  and  $\mathbf{B} \in \mathbb{R}^{N \times T}$ .  $\mathbf{U}$  has  $M$  columns referred to as eigenimages, and the set of all of them is denoted  $\mathcal{U}$ . Similar to the case of  $\mathcal{G}$ , it is possible to conclude that the frames belong to  $\text{span}(\mathcal{U})$ . The diagonal matrix  $\mathbf{S}$  contains their respective singular values that can be interpreted as their respective importance, while  $\mathbf{V}$  contains temporal information that is not currently leveraged by MUSICAL.

In this model, if  $N < M$  then  $\text{rank}(\mathbf{I}_o) = N$  (assuming the vectors in  $\mathcal{G}$  are linear



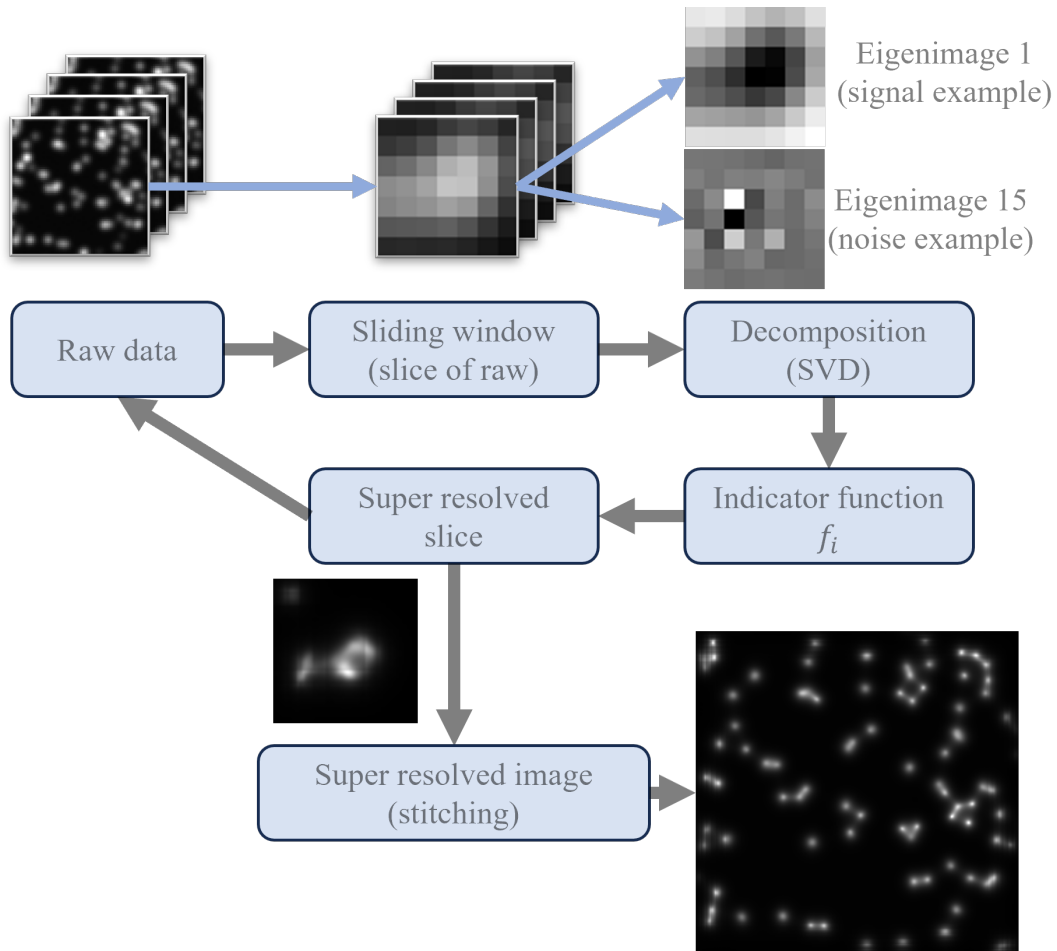


Figure 5.2: MUSICAL processing workflow. The process starts from the raw data, a temporal stack of  $T$  2D images of width  $X$  and height  $Y$ . Then, the algorithm iteratively processes smaller patches of this image. A patch is called a sliding window and has  $T$  images of width and height equal to  $N_w$ . The decomposition allows the separation of the components that are then combined using an indicator function. In this illustration, two different eigenimages are shown corresponding to signal and noise respectively. The super-resolved patch is collected and added to the final MUSICAL result.

independent). In the noiseless case, therefore, the rank of the measurements  $\mathbf{I}_o$  can be used to directly estimate the number of emitters in the sample in case this is unknown. With a fixed number  $N$ , it is possible to separate  $\mathcal{U}$  into two sets, the so-called signal set  $\mathcal{U}_S$  and the noise set  $\mathcal{U}_N$ . Since  $N$  is given, then the number of non-singular values is also  $N$  in which case the groups are defined as follows:

$$\mathcal{U}_S = \{\mathbf{u}_i \in \mathcal{U} \mid s[i] > 0\} \quad \mathcal{U}_N = \{\mathbf{u}_i \in \mathcal{U} \mid s[i] = 0\} \quad (5.3)$$

$\mathcal{U}_S$  corresponds to the eigenimages that enable the construction of the sample, while  $\mathcal{U}_N$  corresponds to its complement. Therefore, it is possible to restrict the frames to only  $\text{span}(\mathcal{U}_S)$  since all the elements of  $\mathcal{U}_N$  have singular values of zero and thus do not contribute to the measurements. Consequently, the frames can be formed by either  $\text{span}(\mathcal{U}_S)$  or  $\text{span}(\mathcal{G})$ , which leads to the conclusion that  $\text{span}(\mathcal{U}_S) = \text{span}(\mathcal{G})$ . More importantly, it can be deduced that  $\text{span}(\mathcal{G})$  is orthogonal to  $\text{span}(\mathcal{U}_S)$ . From this, we can observe that for an emitter  $n$  whose image  $\mathbf{g}_n$  already belongs to  $\mathcal{G}$ , the following is true:

$$\mathbf{g}_n \cdot \mathbf{u}_i = 0 \quad \text{for } \mathbf{u}_i \in \mathcal{U}_N \quad (5.4)$$

Equation 5.4 assumes that  $\mathbf{g}_n$  is known. However, it also suggests a mechanism for detecting the position of an emitter  $k$  by probing different the PSF at different arbitrary points, by accepting the emitters whose dot product is zero for eigenimages in  $\mathcal{U}_N$ . This observation is capitalized by MUSICAL through the definition of an *indicator function* as it will be shown later.

A more realistic scenario though needs to consider the case where  $N > M$  and where the measurements contain noise as well. Under these circumstances,  $\mathbf{I}_o$  is normally fully ranked and therefore it is not possible to classify its eigenimages using Equation 5.3.

This can be partially solved by giving a threshold over the singular values. The assumption is that the signal in the signal is stronger than the noise. This strength is captured by their singular value. In that case, for a threshold given by the user denoted as  $t_u$ , the eigenimages can be separated as:

$$\mathcal{U}_S = \{\mathbf{u}_i \in \mathcal{U} \mid s[i] \geq t_u\} \quad \mathcal{U}_N = \{\mathbf{u}_i \in \mathcal{U} \mid s[i] < t_u\} \quad (5.5)$$

This is the separation mechanism applied originally by MUSICAL [11] and is denoted as applying a *hard threshold*. Later, it will be discussed a different approach that alleviates the strict disjoint separation which is the main contribution of Paper II [66] and Paper III [67]. The  $\text{span}(\mathcal{U}_S)$  corresponds to all the possible images that the most prominent structures of the sample may form, while the  $\mathcal{U}_N$  corresponds to the corruption in the measurements. However, it is up to the user to provide a good threshold that allows such separation. In this more realistic scenario and provided that the separation is possible:

$$\mathbf{g}_n \cdot \mathbf{u}_i \approx 0 \quad \text{for } \mathbf{u}_i \in \mathcal{U}_N \quad (5.6)$$

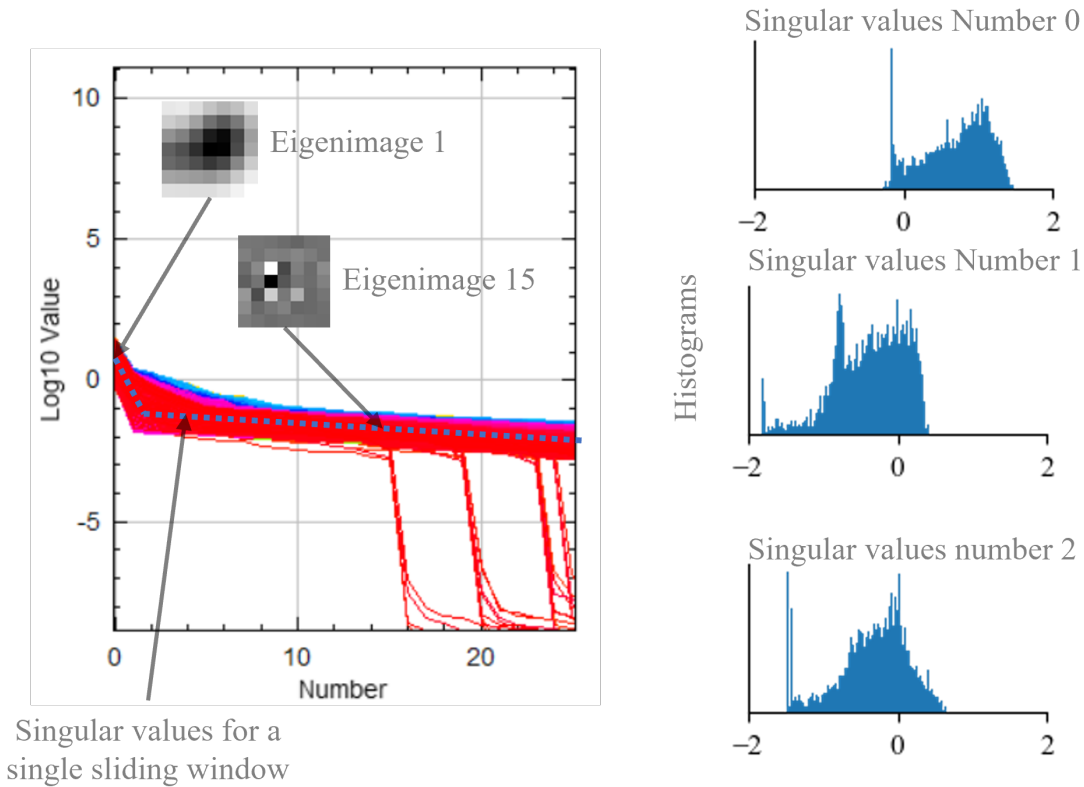


Figure 5.3: Singular values. The plot to the left is used to assess the range for the threshold. Each line corresponds to all the singular values calculated for a single window. The horizontal axis represents the order of the singular value with order zero being the first and largest. This plot was obtained using the graphical tool available in the tool published with Paper I. It also illustrates the eigenimages associated with two of those values in the curve and how the higher the singular value, so lower in order, the more structural information is contained in the eigenimage. Eigenimage 14 in this case is likely to contain more noise. The right part shows a histogram for the three first singular values. The second (Number = 1) shows an inflection point and it is normally used as a guide for the threshold. The convention is to use the inflection point which normally corresponds to the lowest second singular value.

Selecting a threshold often involves heuristics. In [11], a simple procedure was suggested:

- Compute the singular values for all the sliding windows in the sample.
- For each sliding window, create a plot of its singular values on a logarithmic scale.
- Choose a threshold near the inflection or knee-point of the graph.

An example of such a plot is given in Figure 5.3.

It is important to mention that the concept of separating signal from the noise is not particular to MUSICAL. In [68] the authors describe an algorithm based only on the signal set to denoise a temporal stack of images. Similarly, in [69] the authors perform SVD over the covariance matrix of the measurements to achieve super-resolution.

## 5.2 Indicator function

With the signal and noise sets defined, it is possible to apply the knowledge of the PSF of the system and design a function that combines signal and noise to create a single image that exploits the property. This function can be evaluated at arbitrary points as long as the image of an emitter located at such position is known. This condition can be achieved by using the knowledge of the PSF of the system. Then, for a fluorescent emitter  $i$  with an image given by vector  $\mathbf{g}_i$ , the MUSICAL image at a subpixel located at the location of the emitter is given by:

$$f_i = \sqrt{\frac{\sum_{\mathbf{u} \in \mathcal{U}_S} |\mathbf{u} \cdot \mathbf{g}_i|^2}{\sum_{\mathbf{u} \in \mathcal{U}_N} |\mathbf{u} \cdot \mathbf{g}_i|^2}}^\alpha \quad (5.7)$$

The parameter  $\alpha$  as mentioned previously in Section 4.5.2 can be omitted in most of the situation and be left at 4. A nice interpretation of this expression is that the indicator function corresponds to the ratio of the lengths of the *projection* of  $\mathbf{g}_i$  into the hyperplanes formed by the signal (numerator) a noise (denominator) spaces as seen in Figure 5.4.

Therefore, the indicator function can be seen as a measure of closeness to either plane where emitters are assumed to be much closer to the signal space. However, it is important to mention that the denominator's main purpose is to aid in the seamless stitching of the results of applying MUSICAL over every sliding window.

An alternative representation of the indicator function can be given in terms of either the signal or noise part. This can be done since  $span(\mathcal{U}) = \mathbb{R}^P$  and because  $span(\mathcal{U}_S$  and  $\mathcal{U}_N$  form orthogonal spaces.. Therefore, for any  $\mathbf{g}_i$  the following holds:

$$|\mathbf{g}_i|^2 = \sum_{\mathbf{u} \in \mathcal{U}_S} |\mathbf{u} \cdot \mathbf{g}_i|^2 + \sum_{\mathbf{u} \in \mathcal{U}_N} |\mathbf{u} \cdot \mathbf{g}_i|^2 \quad (5.8)$$

Therefore, the indicator function can also be written as:

$$f_i = \sqrt{\frac{\sum_{\mathbf{u} \in \mathcal{U}_S} |\mathbf{u} \cdot \mathbf{g}_i|^2}{\|\mathbf{g}_i\|^2 - \sum_{\mathbf{u} \in \mathcal{U}_N} |\mathbf{u} \cdot \mathbf{g}_i|^2}}^\alpha = \sqrt{\frac{\sum_{\mathbf{u} \in \mathcal{U}_S} |\mathbf{u} \cdot \mathbf{g}_i|^2}{\sum_{\mathbf{u} \in \mathcal{U}_N} |\mathbf{u} \cdot \mathbf{g}_i|^2}}^\alpha \quad (5.9)$$

While simple, this can be exploited for optimizing the processing time of any MUSICAL implementation. An example is MusiJ [70], the implementation presented in Paper I.

The MUSICAL image is formed by taking a large amount of test points across the sample space, normally on a grid finer than the original data. Therefore, the indicator function needs to be calculated for each of these points which allows us to observe structures at a scale that is lower than the diffraction limit. When first published MUSICAL reported a resolution of 50 nm on data with high spatial temporal sparsity. This can be obtained thanks to the design of the indicator function that allows probe points at any arbitrary point in the sample.

It is important to mention that it is the division by the noise that enables the super-

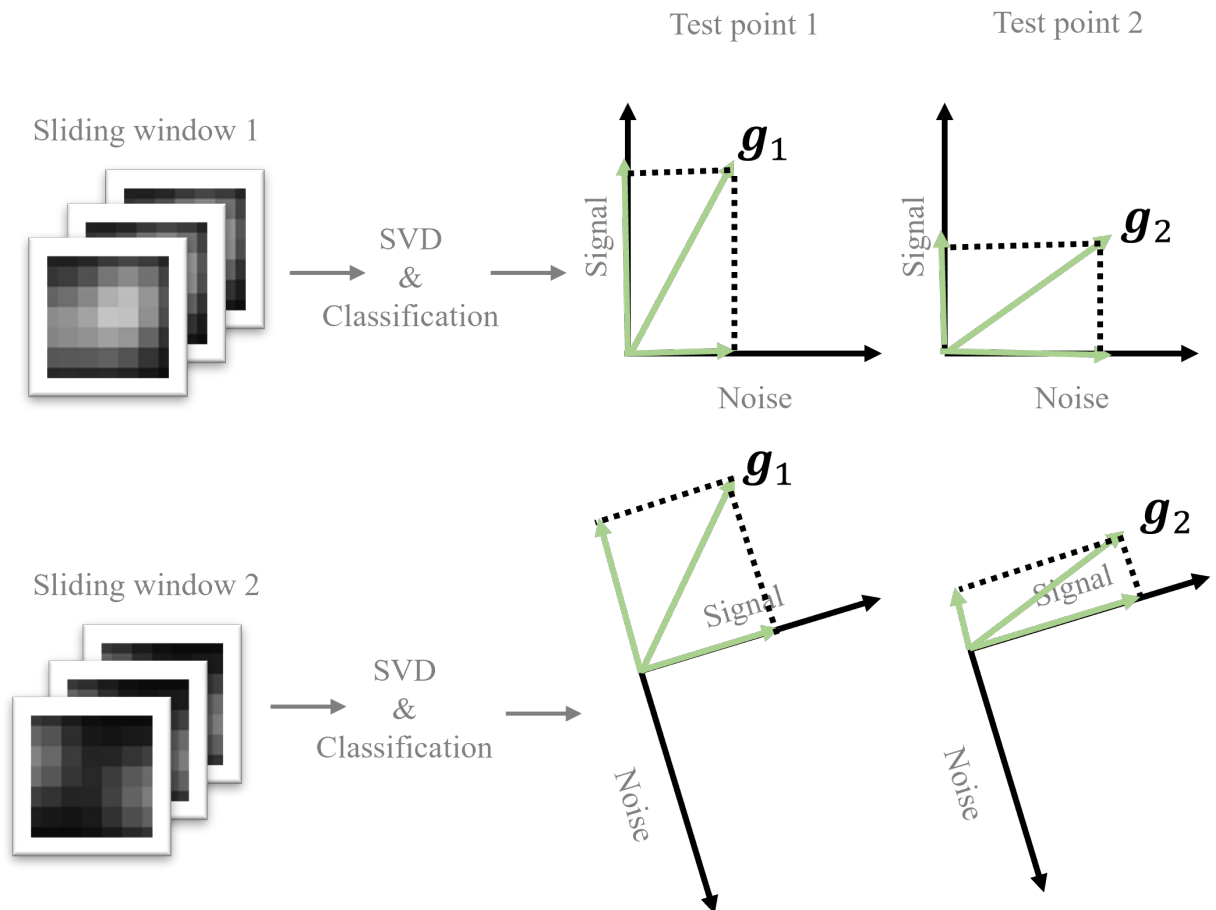


Figure 5.4: Simplified scheme for point versus signal and noise space for a 2-pixel detector. In this system, the test point's image  $\mathbf{g}_i$  and eigenimages are vectors in  $\mathbb{R}^2$ . One of the eigenvectors corresponds to the signal and the other to the noise and their direction depends on the measured data. For the first sliding window,  $\mathbf{g}_1$ 's projection is higher in the signal than in the noise, while  $\mathbf{g}_2$  is higher projection in the noise than in the signal space. This means that this indicator function will have a peak at  $\mathbf{g}_1$ . The second window creates different spaces the projections are affected accordingly.

resolution properties of MUSICAL. Let's consider the eigenimage  $\mathbf{u}_i$  as a sampled vector of a 2D continuous eigenimage  $U_i(x, y)$ . In that case, it is possible to observe that the numerator in the continuous domain is given by the convolution of  $G(x, y)$  by the functions  $U_i(x, y)$  (with  $\alpha = 2$  to simplify the notation):

$$F_{num}(x, y) = \sum_{\mathbf{u} \in \mathcal{U}_S} |U(x, y) * G(x, y)|^2 \quad (5.10)$$

This says that the numerator is given by a sum of diffracted limited images. Therefore, super-resolution must come from the denominator. However, it has not yet been possible to define a quantitative number for the resolution enhancement and more work needs to be done regarding this.

So far, signal and noise components present in the  $\mathcal{U}_S$  and  $\mathcal{U}_N$  are either in the numerator or in the denominator. This changed in 2020 with the publication of Paper II [66]. In this work, the following indicator function was studied (the notation is slightly changed to match the format used in this thesis):

$$f_i = \sqrt{\frac{\sum_{\mathbf{u} \in \mathcal{U}} a(\mathbf{u}) |\mathbf{u} \cdot \mathbf{g}_i|^2}{\sum_{\mathbf{u} \in \mathcal{U}} b(\mathbf{u}) |\mathbf{u} \cdot \mathbf{g}_i|^2}}^\alpha \quad (5.11)$$

There are two main changes concerning the conventional function shown in Equation 5.7. The first is that the numerator and denominator may contain all the eigenimages of  $\mathcal{U}$ . The second is that their contribution is given a function that depends on every eigenimage  $\mathbf{u}$ . For example, the Equation 5.7 can be expressed using coefficients as a function of an eigenimage  $\mathbf{u}$ :

$$a(\mathbf{u}) = \begin{cases} 1 & \text{if } \mathbf{u} \in \mathcal{U}_S \\ 0 & \text{otherwise} \end{cases} \quad b(\mathbf{u}) = \begin{cases} 1 & \text{if } \mathbf{u} \in \mathcal{U}_N \\ 0 & \text{otherwise} \end{cases} \quad (5.12)$$

The only requirement imposed into Equation 5.11 is that  $a(\mathbf{u}) + b(\mathbf{u}) = 1$ . Using this constrain, a more interesting case is obtained when considering  $\mathbf{u}$  and its corresponding singular value  $s$ :

$$a(\mathbf{u}) = \begin{cases} 1 & \text{if } s \geq s_{\max} \\ 0 & \text{if } s \leq s_{\min} \\ \frac{\log_{10} s - \log_{10} s_{\min}}{\log_{10} s_{\max} - \log_{10} s_{\min}} & \text{otherwise} \end{cases} \quad ; \quad b(\mathbf{u}) = 1 - a_i(x) \quad (5.13)$$

MUSICAL with this modified indicator function is referred to as Soft MUSICAL and previously suggested, it allows to weight and include eigenimages that may be ambiguous to classify into both denominator and numerator. The terms  $s_{\max}$  and  $s_{\min}$  in principle depend on the user. However, since their publication, their values correspond to the maximum a lowest second singular value across all the sliding windows being processed. This

means that a preprocessing step is required that computes only the singular values for each window. This value is based on heuristics and from the observation that the first singular value tends to be orders higher than the second.

Another interesting application of this definition of the indicator function was published later in Paper III [67] with the following constants:

$$a(\mathbf{u}) = \begin{cases} 1 & \text{if } s \geq s_{\max} \\ 0 & \text{if } s \leq s_{\min} \\ \frac{\log_{10} s - \log_{10} s_{\min}}{\log_{10} s_{\max} - \log_{10} s_{\min}} & \text{otherwise} \end{cases} ; \quad b(\mathbf{u}) = 1 \quad (5.14)$$

This modifies the constants given by Equation 5.13 by setting  $b(\mathbf{u}) = 1$ , neglecting the effect of the denominator. This method is referred to as *CE* as its main intention is to be applied as a contrast enhancement technique with no expected gain in resolution.

A qualitative comparison can be by looking at the results obtained in Figure 5.5. The analysis of these indicator functions is the main topic of Papers II and III.

### 5.3 Eigenimages

So far the discussion has focused on the indicator function. However, a key component of the algorithm is the concept of the eigenimages. For an optical system with  $P$  sensor elements,  $\mathbf{U} \in \mathbb{R}^{P \times P}$ . This means that each column can be interpreted as an actual image by converting the column vector into a 2D array. One way of looking at the eigenimages is by drawing a parallel with the Fourier transform. With the Fourier transform it is possible to express a 2D image in terms of a sinusoidal pattern. A 2D image  $I(x, y)$  can be expressed in term of its Fourier transform  $I(u, v)$  as:

$$I(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} I(u, v) e^{2\pi i(xu + yv)} du dv \quad (5.15)$$

Here, the basis is given by the exponential components that express a sinusoidal pattern at different spatial frequencies and directions. In this decomposition, the relevant information is contained in the complex function  $I(u, v)$  that not only incorporates the strength of each basis element but also spacial information through its phase.

Similarly, SVD can be used to find a basis for a discrete signal. The main difference is that the basis itself depends on the structure and therefore it is specific to the data being analyzed. Therefore it is not possible to look just at the weights as it is done in the Fourier domain.

It is possible to go further though. In 2017 Agarwal provided a deeper understanding of the eigenimages by associating them with different frequency components in the Fourier domain [71]. This article concluded that eigenimages with lower singular values provide information about higher spatial frequency components in the sample. With this understanding, it is possible to make the following observations:

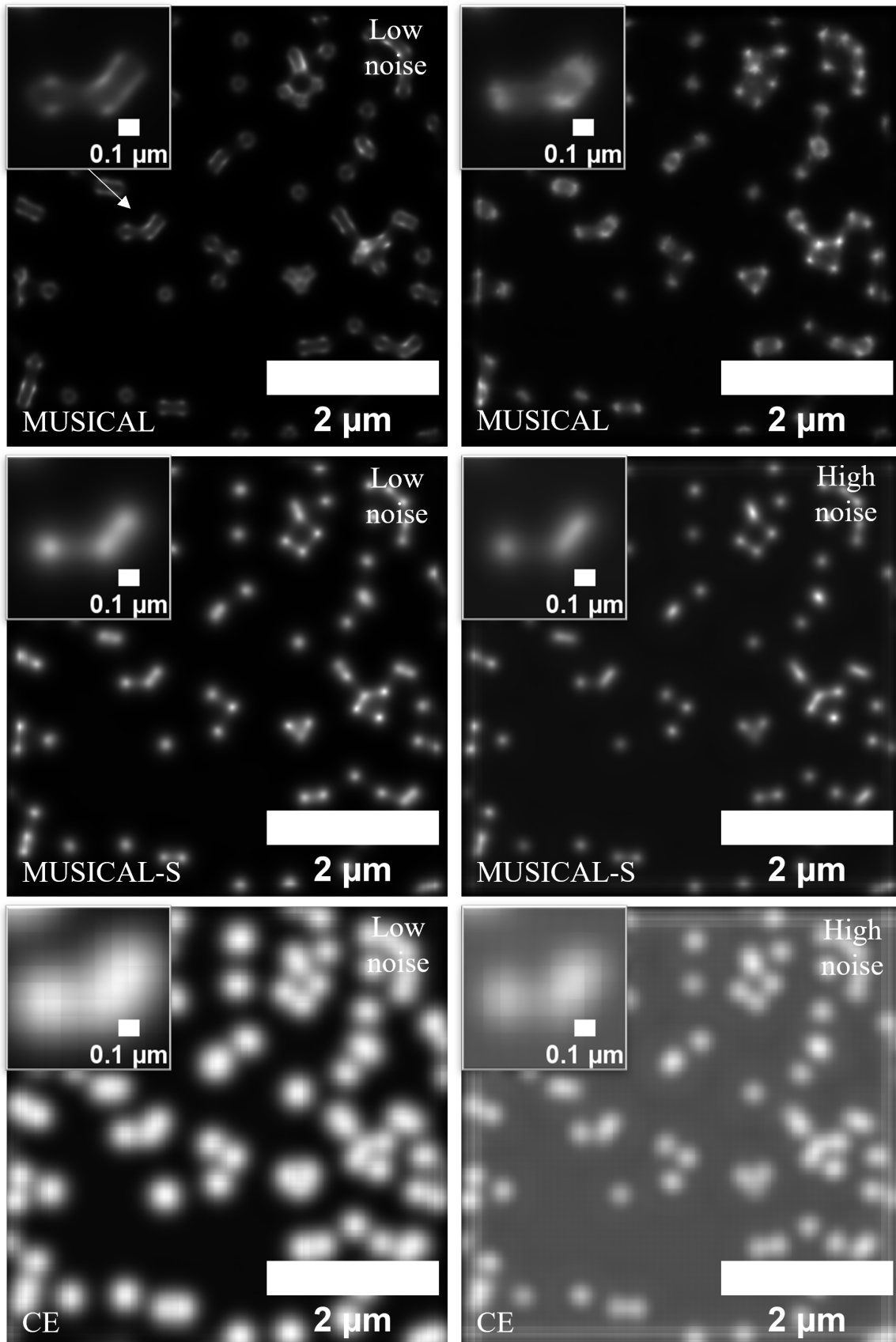


Figure 5.5: Indicator function applied on simulated data for two different noise levels. Left column presents reconstructions for a dataset whose SBR is around 20. The right column displays results of the same image but with SBR closed to 4.



- As more eigenimages are included in the signal set, higher frequency structural information is shown in the reconstruction.
- Noise is mainly observed as high-frequency spatial disturbances. Therefore by including more eigenimages into the signal set the chances of enhancing noise increase.

This matches the observations made in Paper II where several threshold strategies were tested. The main conclusion was that a low threshold translates into more details but also more artifacts. Therefore, there exists a trade-off between resolution gain and artifacts.

Figure 5.6 shows how the eigenimages for a single sliding window look like (upper figure). Each eigenimage has been plotted with its respective singular value. This example also illustrates the difficulty in selecting a threshold as two persons may disagree on where to place the threshold as the limit between signal and noise seems diffuse. This figure also shows the projection of all the test points for every eigenimage. This also illustrates how the threshold affects the structures appearing in the final reconstruction. For example, when considering the eigenimage with singular value 0.531 (second row, second column) as signal, will produce that its projection will be present in the signal part. As a result, the strong peaks in the middle-top and bottom-right will appear as part of the reconstructed structure.

An example of eigenimages can be seen in Figure 5.6.

## 5.4 3D MUSICAL

The articles included in this thesis are focused on 2D images. However, it is also possible to adapt the algorithm for 3D. The following results correspond to unpublished results where MUSICAL was adapted to work with 3D data where the images are 4-dimensional arrays. It is also possible to perform reconstruction over 2D images provided that the PSF encodes the axial information. The results here presented were obtained using raw data generated with a 3D SIM microscope. The raw data is publicly available at [72] as part of [73].

To adapt MUSICAL for 3D temporal data the processing step needed to be modified. This modification is illustrated in Figure 5.7. This process allows to convert 4D data into a 2D array. This procedure is similar to the one performed for MUSICAL in 2D, except that the rectangular matrix contains information from multiple planes simultaneously.

The reconstruction obtained through this method is shown in Figure 5.8. This preliminary result shows an improvement over the average image. However, it is worth mentioning that this result used only 15 frames as the raw data corresponds to a 3D-SIM device. For comparison, an improved version of 3D-SIM called 4B-SIM was included [73]. This result is promising as the reconstruction is done almost blindly and with only a few volumes.

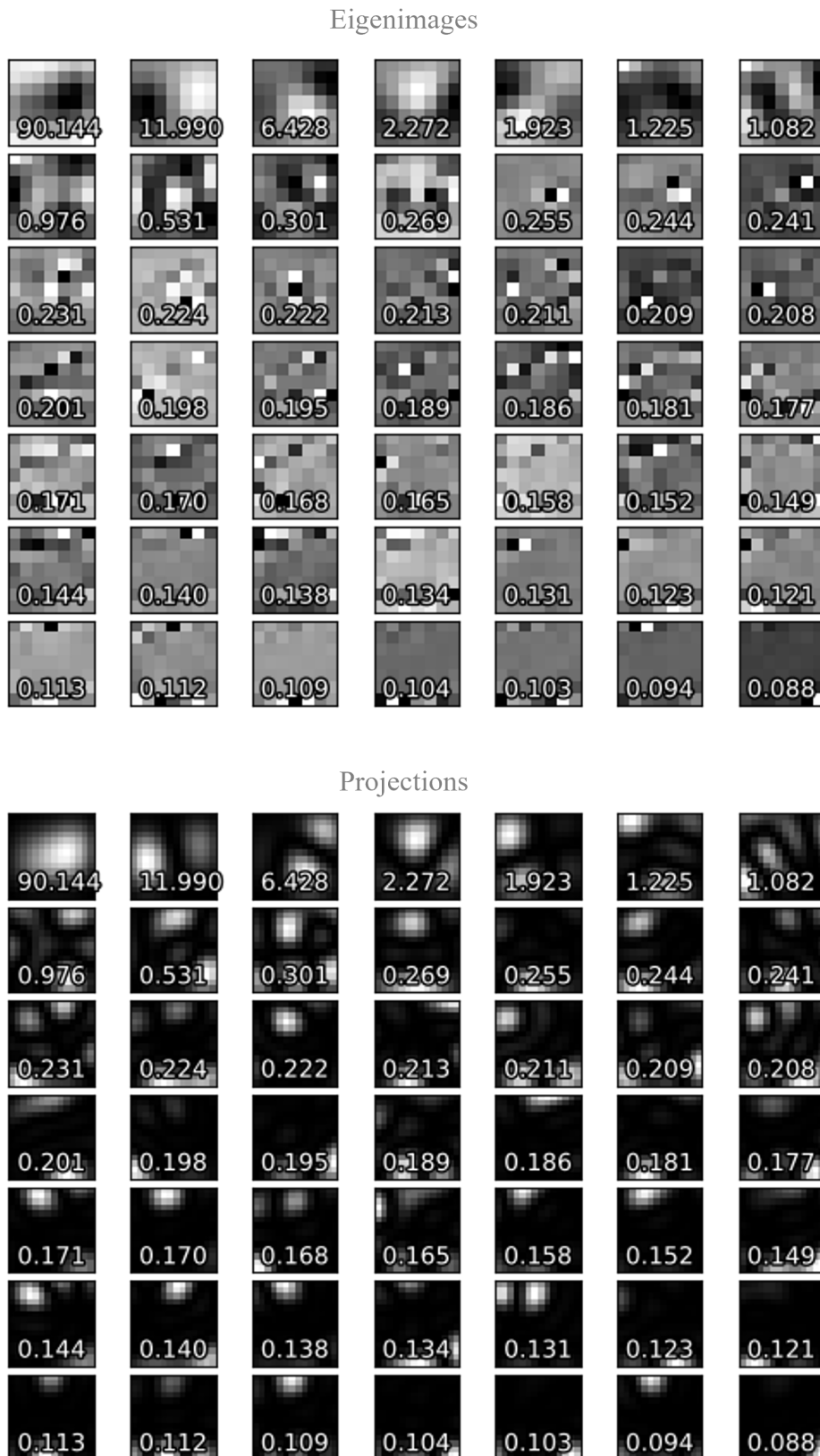


Figure 5.6: Example of eigenimages in decreasing order and their respective projections. The projections are obtained by using test points similar to how MUSICAL images are created, except that only one eigenimage is considered.

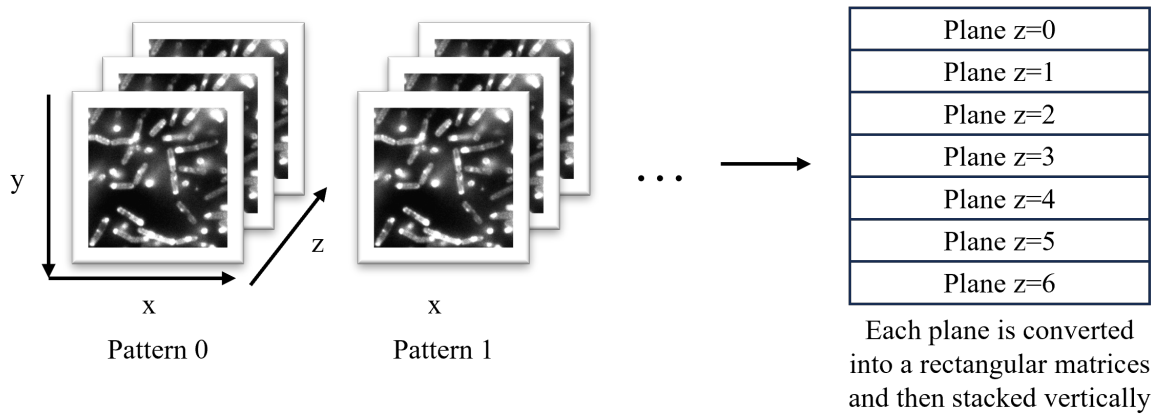


Figure 5.7: 4D array into 2D array for MUSICAL processing. To produce a rectangular matrix every plane (2D array) is converted into a rectangle using the same procedure for 2D MUSICAL. These matrices are then stacked vertically and then decomposed using SVD. From there, the procedure is the same as for MUSICAL for 2D temporal data.

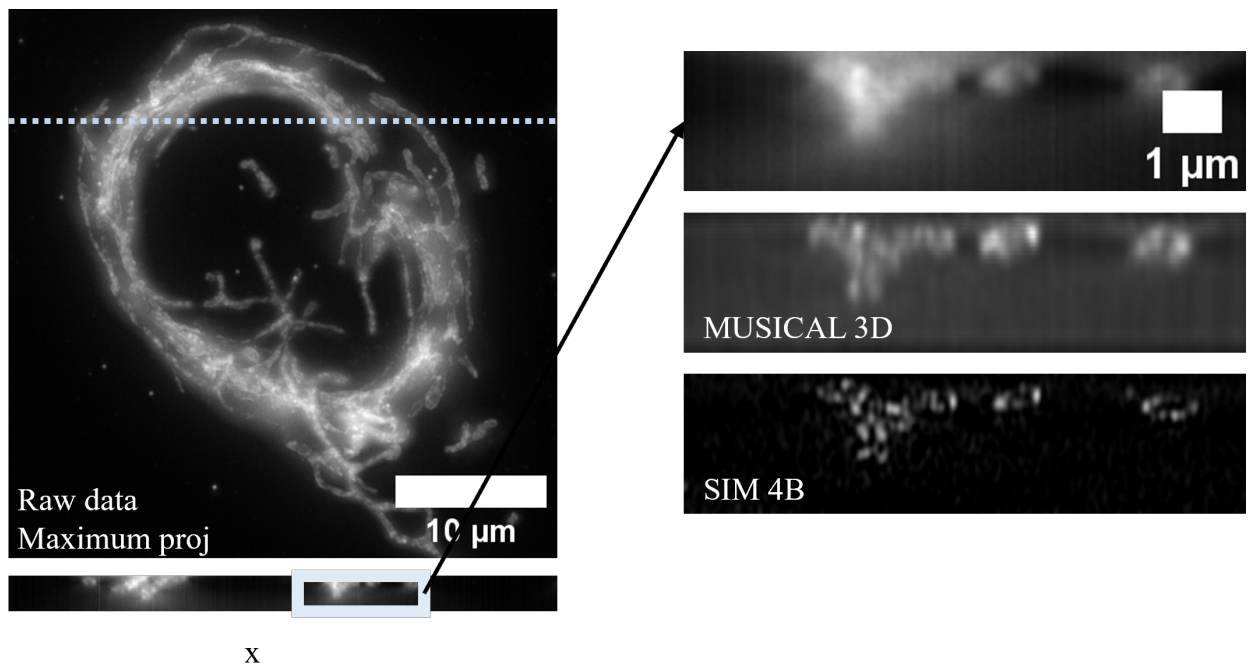


Figure 5.8: 3D reconstruction using MUSICAL



# Chapter 6

## Summary of Research

This chapter is devoted to summarizing the results obtained in the papers selected for this thesis. Since MUSICAL is the main topic developed it relies on the concepts explained in Chapter 5.

### 6.1 Paper I

#### **MusiJ: an ImageJ plugin for video nanoscopy**

The publication of MUSICAL in 2016 [11] was complemented by the release of a MATLAB open-source implementation available on the author's website [74]. While functional, this interface may be challenging to access as MATLAB requires a license and some understanding of programming language. In this work, we proposed an implementation for ImageJ [14] that is easy to use and therefore serves as a bridge between MUSICAL and non-programmers. ImageJ is a free software written in Java that allows interaction with the most common image formats used by commercial microscopes. This is enabled by the vast amount of tools bundled with the initial installation including profiling and analysis tools for large temporal stacks. Its community is active and the software facilitates the integration of new functionality through a central system that allows to installation of custom plugins from 3<sup>rd</sup> party developers. MusiJ [70] is a plugin written in Java that can be installed easily in ImageJ. Its publication fills the gap existing between researchers with no experience in MATLAB and MUSICAL by providing easy access to its capabilities. It additionally presents an extra layer through a macro that allows it to increase its functionality even more by making easier processing of multiple channels and long temporal stacks.

The work presented in this paper aimed to achieve two objectives:

- **User interface:** This interface enables users to apply MUSICAL to any temporal stack they can load in ImageJ. It provides additional functionality by allowing temporal windowing for samples, thus generating temporal stacks of super-resolved images. This is achieved by sequentially reconstructing MUSICAL images from smaller sets of frames. The interface is implemented as a macro.

- Code optimization: JCode optimization: Java is a high-level programming language that does not work with matrices in the same explicit manner as MATLAB. Therefore, this implementation required several modifications that distinguish it from its MATLAB counterpart. There are three main modifications:
  - Enabling multi-threading on the image: MUSICAL can be applied to multiple regions independently, allowing the algorithm to run in parallel. The image is divided into different sections, and the results are later combined.
  - Utilizing the completeness property of the basis obtained through singular decomposition (as demonstrated in Equation 5.9): This simplifies the reconstruction process, as it necessitates only the projection in the signal space.
  - Reducing data precision to single-precision floating-point: This optimization helps enhance computational efficiency.

It is worth noting that these changes could also be applied to MATLAB, but that would still not fix its usability issue.

The entire process is illustrated in Figure 6.1. The success of the reconstructions depends on selecting an appropriate threshold to separate the signal from the noise accurately. To aid in this process, the software provides a graphical analysis tool that allows the plotting of all the singular values of the image (Figure 6.1b). Users can then make an educated guess regarding a suitable threshold value for reconstructing the final MUSICAL image. This process may involve some trial and error, underscoring the importance of optimizing the implementation.

My contributions to this work included developing the algorithm in Java and contributing to the manuscript. This included selecting a suitable library for matrix manipulation, implementing the necessary operations for MUSICAL, and enhancing the code's efficiency through parallelization and optimization of operations based on Linear Algebra.

## 6.2 Paper II

### Soft thresholding schemes for multiple signal classification algorithm

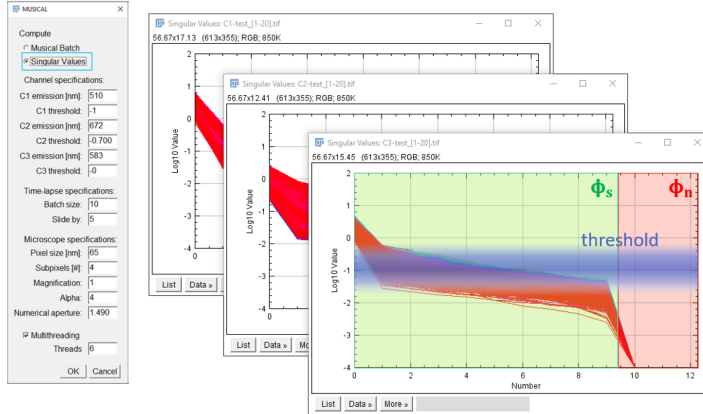
The principle of MUSICAL relies on the computation and separation of two spaces: the signal and the noise space. These two components are extracted from the observed data and then combined using a function called the indicator function. As discussed in Section 5.2, this separation is crucial and is what enables the super-resolution capabilities of the algorithm.

Traditionally, this separation requires user input in the form of a threshold parameter. This value serves as a binary filter applied to every sliding window in the sample. The process of selecting an appropriate threshold places additional responsibility on the user

a) starting the macro from the FIJI toolbar



b) determining a suitable signal/noise threshold



c) computing MUSICAL superresolution images

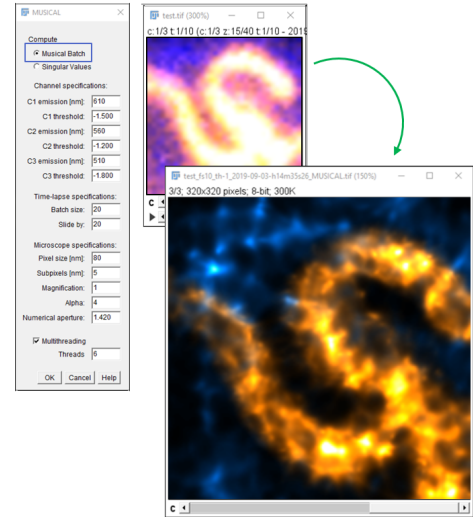


Figure 6.1: MusiJ complete workflow. a) It can be loaded through quick access in case the user needs to. b) Graphical tool for selecting a threshold. c) Example of original and reconstructed image. Adapted from [70]

and may lead to challenges in interpreting the results, as multiple reconstructions must be created by iterating through different threshold values to find a suitable one.

In this context, this article contributes to the following aspects in order of relevance:

- Extension of the indicator function's definition from one based on a hard threshold to one that allows customization.
- Presentation of several alternative indicator functions with different examples of experimental data.
- Conducting a study on the effect of the threshold in the reconstruction based on the simulation of 3D data. This allowed us to analyze both artifacts and the effect of out-of-focus light as a function of the threshold.
- Development of a small simulation environment for fluorescent data that allows the creation of simple 3D structures and the generation of fluorescent raw data based on such structures.

The most intriguing contribution from this list is the extension of the indicator function. This extension allowed for the creation and testing of various functions, not only in this work but also served as an inspiration for the research presented in Paper III.

Among all the tested methods, Soft MUSICAL, also referred to as MUSICAL-S in the article, outperforms the others. It exemplifies a soft separation scheme in which both the numerator and denominator incorporate signal and noise components, with a weighting function that is eigenvalue-dependent. Following its publication, this method has become the standard and has subsequently been integrated into many later works where MUSICAL has been employed (e.g., [75, 76]).

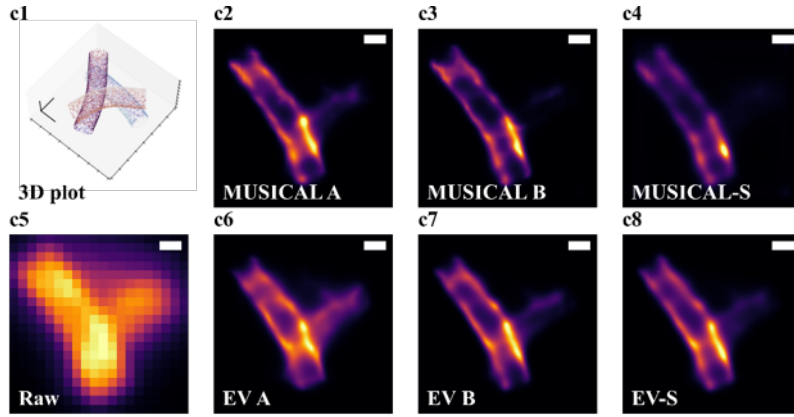


Figure 6.2: Simulated dataset of 3 cylinders with emitters in its perimeter. The cylinders are spatially distributed in 3D dimensions. Scale bar 300 nm. This figure was reproduced and adapted from [66]

By introducing this new indicator function, we formally introduced the concept of MUSICAL's *family* of functions due to the vast number of possibilities for customization. The methods presented in this paper are:

- MUSICAL A: A conventional indicator function with a low threshold. It provides the best results in terms of resolution but is also more prone to including details from the background, regardless of whether it's noise or signal.
- MUSICAL B: A conventional indicator function with a mid threshold. It offers better out-of-focus light rejection but with less resolution gain.
- EV A and EV B: These methods use the singular values of every eigenimage as weights in addition to the conventional classification with a threshold. This approach makes them more prone to artifacts.
- MUSICAL-S and EV-S: The first applications of the soft scheme. Both produce better results in terms of contrast across the entire field of view but come at the cost of a slight reduction in resolution gain.

In terms of the studies performed, they were based on simulations and experimental data. The simulation part allowed us to quantify the results in terms of contrast and resolution, while the experimental part allowed us to observe its effect on challenging but real data.

Figure 6.2 presents an example of a simulated 3D structure with the respective reconstructions. The simulations used a 3D distribution of scattered points, each one with its own temporal intensity. The final step was to create the image of every single emitter using the Gibson and Lanni model [15].

My contribution to this work included the implementation of MUSICAL and its corresponding generalized indicator function, along with contributing to its design and its



different variations. Regarding the simulations, I created them and analyzed all the results associated with them, including the development of metrics to evaluate contrast and resolution for a large set of images. Furthermore, I took the lead in manuscript writing and prepared all the figures in the paper.

## 6.3 Paper III

### Deriving high contrast fluorescence microscopy images through low contrast noisy image stacks

This paper builds upon the concepts presented in Paper II, specifically those related to the generalized indicator function. While MUSICAL was initially designed to leverage the orthogonality of the noise subspace and the point-spread function of the optical system to achieve super-resolution, in this paper, we explore indicator functions that primarily rely on the signal component.

The aim was to test its viability and compare it with similar methods. The methods included in the comparison were Richardson-Lucy deconvolution [52, 53] and automatic correction of sCMOS-related noise [77].

In total, four variations of the indicator function were tested. The best results were obtained with the one named CE1s (short for soft contrast enhancement). This approach uses only the signal with the addition of a weighting component based on the singular values.

An illustrative example of its capabilities is provided in Figure 6.3. This corresponds to pig heart tissue with significant background information. As seen, conventional MUSICAL produces numerous artifacts, which are reduced when applying CE1s. However, it's essential to mention that since the noise part is practically omitted from the indicator function's design, the expected resolution gain is not expected to be beyond the diffraction limit.

My contributions to this work included conceptualizing the indicator function and analyzing the results obtained by the master's student, Mayank Roy. Furthermore, I generated the results for AScN and deconvolution and compiled all the results into the figures presented in the article. Additionally, I took the lead in writing the article.

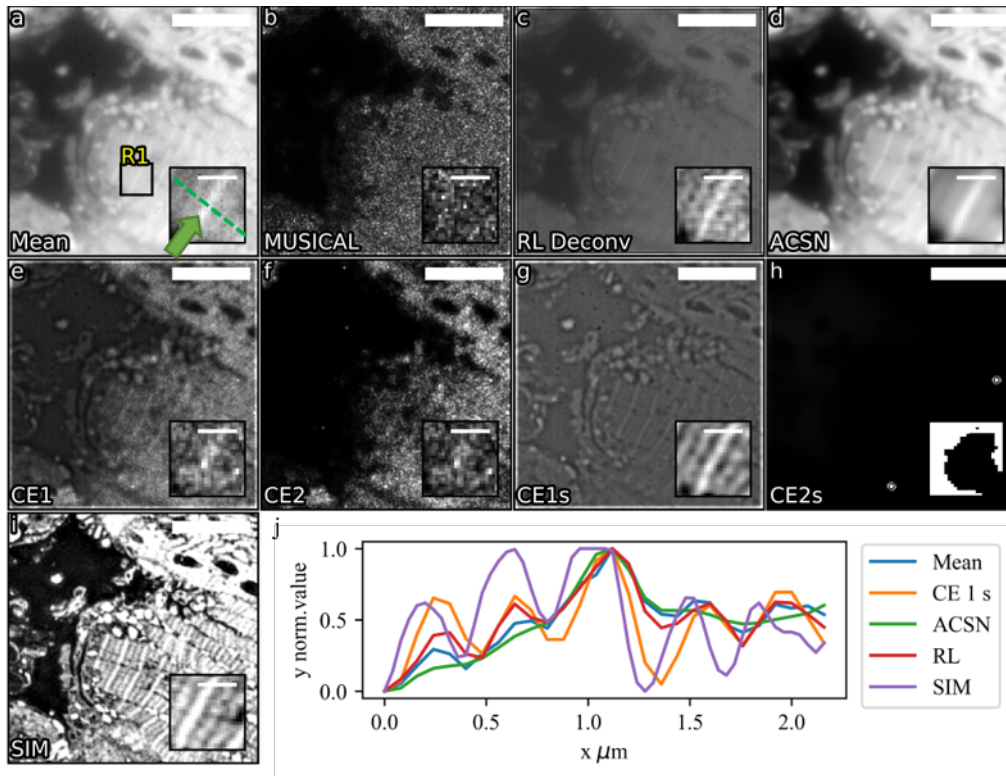


Figure 6.3: Contrast enhancement applied on tissue sample. Several variations of CE were tested. The best results were obtained with CE1s and therefore it is the one that has taken the name. This result illustrates the advantage of using CE vs conventional MUSICAL. The scale bar is  $5 \mu\text{m}$  for the large images and  $1 \mu\text{m}$  for the close view.

## 6.4 Paper IV

### Fluorescence fluctuation-based super-resolution microscopy using multimodal waveguided illumination

Fluctuation-based methods such as the ones described in Chapter 2 rely on the stochastic variation in the fluorescence emission produced by the fluorophores. This can be induced, as described previously, using widefield illumination. An alternative is to create illumination patterns that vary randomly in time. One imaging platform that allows this is total internal reflection microscopy through a photonic chip [78]. This system is able to create a random pattern of illumination inside the waveguide that leaks into the sample due to total internal reflection. The system is designed to quickly change the illumination pattern and therefore it can be used to acquire an image for each single pattern. A schematic that illustrates the concept and the illumination patterns is shown in Figure 6.4 where 3 different patterns are shown to the right. By recording multiple patterns it is possible to obtain an image stack suitable for processing with MUSICAL. In Paper IV we demonstrated the suitability of MUSICAL for illumination engineering systems based on chip-based microscopy. We performed reconstruction using as raw data the sample illuminated with different patterns created sequentially by changing the point at which the illumination laser couples with the waveguide. Figure 6.5 illustrates the framework

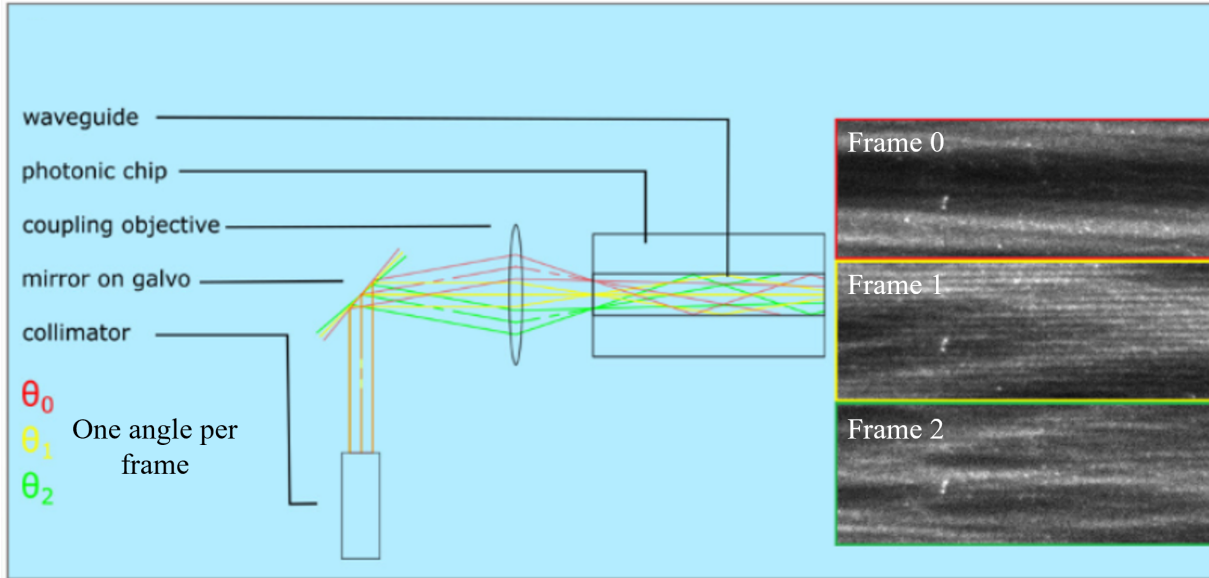


Figure 6.4: Chip-based microscopy system’s illumination patterns. The galvanometer allows the modification of the illumination pattern quickly. The images to the right show 3 different random patterns. Each pattern is captured on a single frame. Adapted from Supplementary of [76].

considered in this work which is relevant for analyzing the number of frames required for a proper reconstruction.

The system was used to acquire sets of images with two distinct magnifications. One set was obtained with an objective lens that allowed 10x magnification with 0.3 NA. As a result, these images correspond to a large field of view of the sample. The second set of images was acquired using a 60x 1.0 NA objective lens. The MUSICAL reconstructions were carried out on these two types of datasets with the advantage that the reconstructions of the 0.3 NA can be compared with the 60x in terms of resolution gain. The raw is available at [79].

My contributions included generating all the reconstructions using different indicator functions. From those results, only the results for soft MUSICAL were reported in the main article. Additionally, I conducted an analysis of the cross-correlations (available in the supplementary material of the article) to study the distance  $\delta x$  (Figure 6.5) to use during the acquisition of actual samples. For this analysis, I used a waveguide with a layer of only fluorophore on top. The distance  $\Delta x$  was set to 100 nm and then 6000 frames were acquired. The results are displayed in Figure 6.7. This plot shows how the correlation is a function of the distance between frames from which is possible to conclude that a distance of 20 reduces the correlation to the minimum. This corresponds to a distance of 2000 nm. For acquiring images of the salmon scale cells,  $\Delta = 400$  nm was used and 1500 frames were acquired. From the analysis performed through correlation, this indicates that using interleaving of 5 is enough for good reconstructions.

Furthermore, I conducted a Fourier analysis of the illumination patterns (also available in the supplementary material of the article).



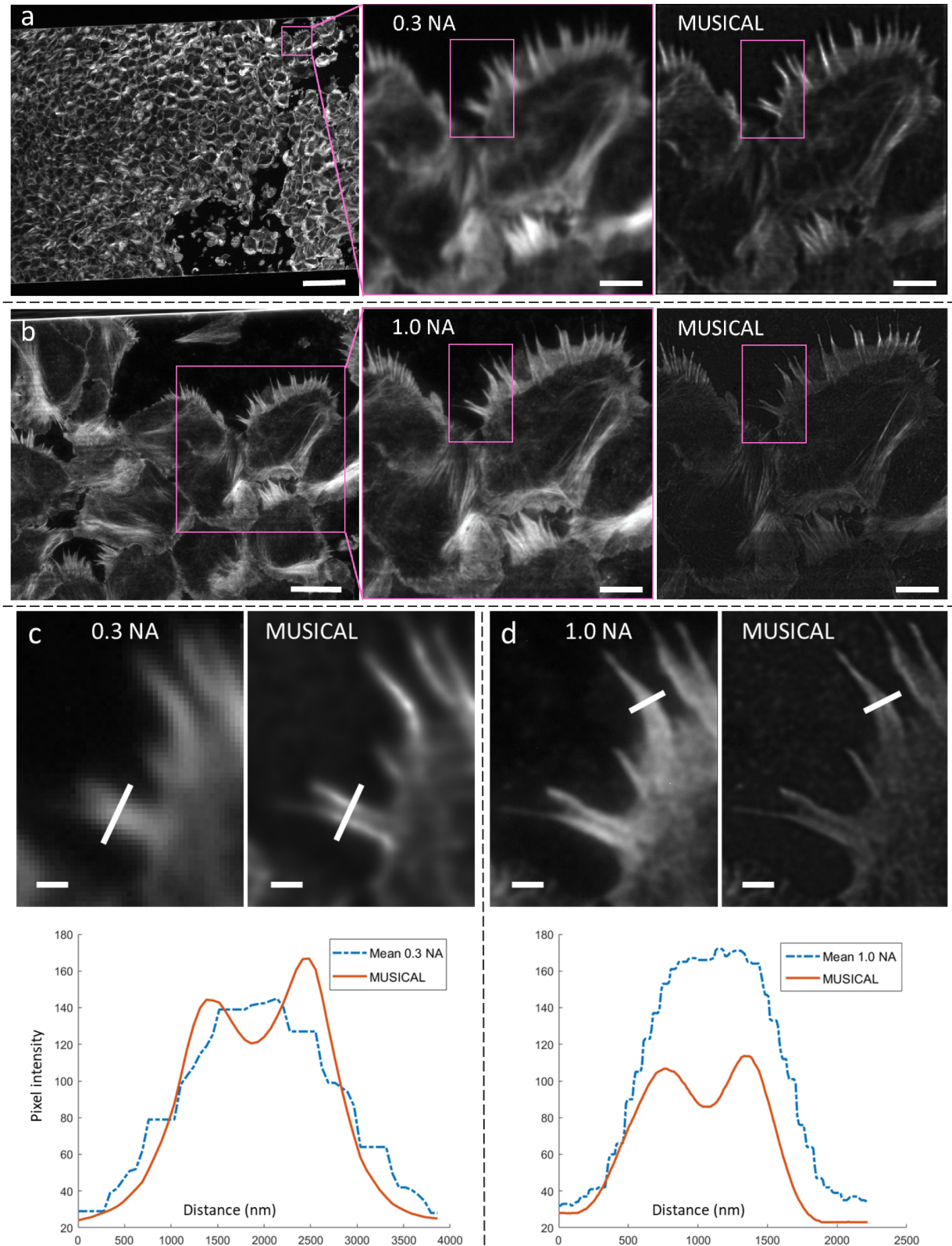


Figure 6.6: MUSICAL applied on chip-based microscopy system for fish scale's cells. MUSICAL allows the reconstructed images of a low NA objective to be similar to the ones obtained through a much higher NA (which also reduces the field of view). Reproduced from [76].



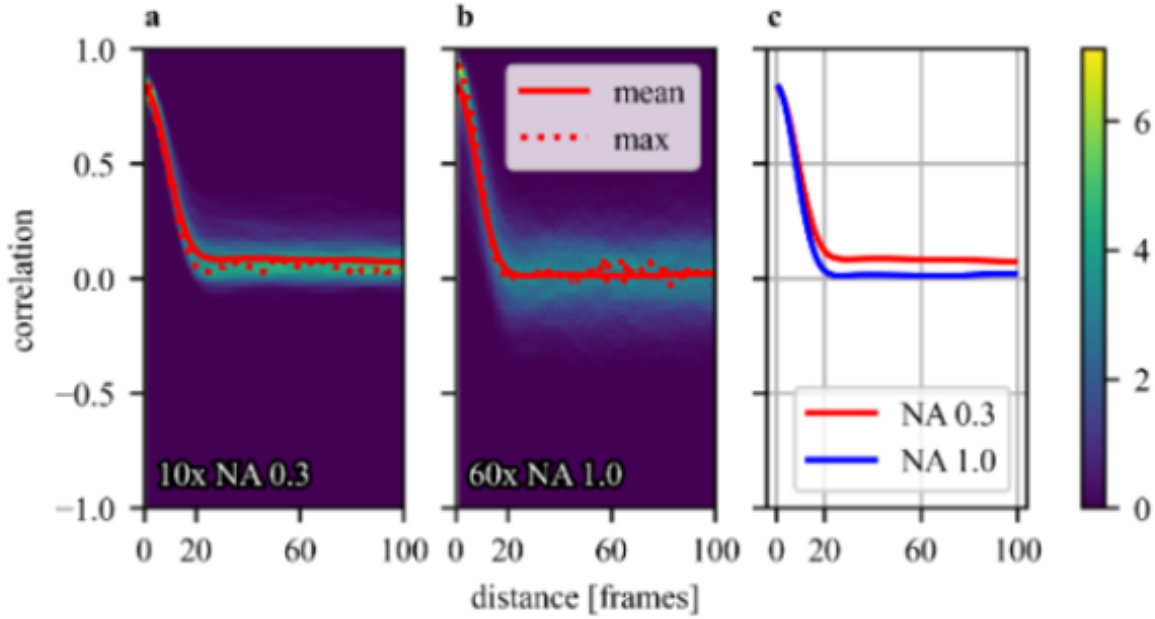


Figure 6.7: Correlation as a function of the distance between frames. The correlation does not decrease after 20 frames. This study was performed with  $\Delta x = 100\text{nm}$  and therefore the distance is 2000 nm. Adapted from [76].

## 6.5 Paper V

### Scalable-resolution structured illumination microscopy

This article [80] investigates the use of MUSICAL as a blind reconstruction method for a structured illumination microscopy-based system. The system designed for this article takes a novel approach to conventional SIM by separating the illumination and detection paths using two different objectives. This enables changes in magnification and field of view in the illumination path while maintaining fixed structured illumination. The system is illustrated in Figure 6.8.

The structured illumination patterns were generated using a spatial light modulator (SLM) with binary patterns, which were subsequently shifted and rotated to create a variety of patterns, allowing the capture of a sequence of pattern-illuminated images. The pattern is depicted in Figure 6.8a in the upper left corner. This pattern is rotated from 0 to 180 degrees using a step of 2 degrees (90 rotations). Then the initial vertical pattern is shifted by one pixel and then rotated again from 0 to 180. The raw data is formed then by a number of frames equal to rotations times shifts.

In conventional SIM [2], the reconstruction demands the clean acquisition of patterns with strong peaks at the harmonics.

In our approach, we employed MUSICAL due to its proficiency in reconstructing super-resolution images from temporal stacks with varying illumination patterns, as demonstrated in Paper IV.

The results demonstrate the system's viability, with a resolution gain of approximately

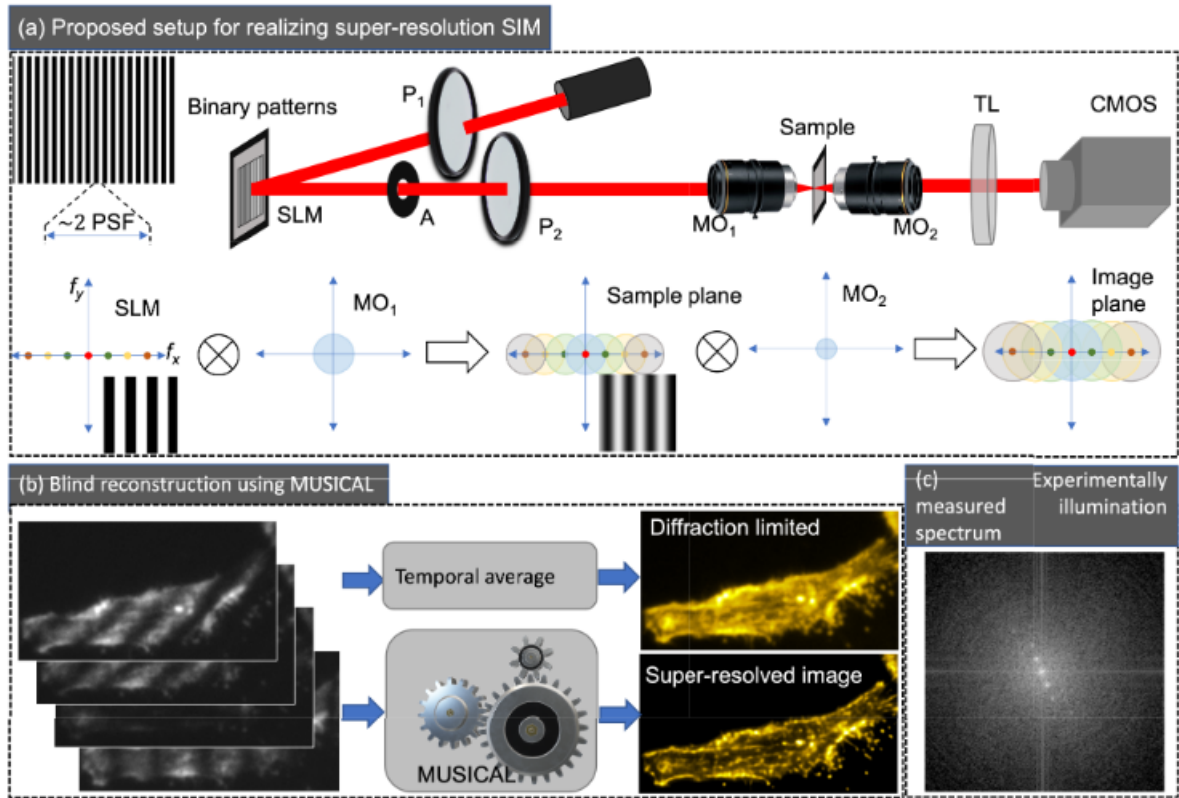


Figure 6.8: Scalable-resolution structured illumination microscopy system. Reproduced from [80]

3 times when compared to the diffraction-limited image.

In this work, I share the first co-authorship. My contributions were the generation of the patterns for the SLM, generating results using MUSICAL, and analyzing the results in the Fourier domain. The latter encompassed an examination of the different eigenimages and their frequency content. The figure that displays that analysis is replicated in Figure 6.9. The rows encode the number of angles used for reconstructing the image while using all the shifts. The columns show the information encoded at every order for eigenimage and were generated using a modified version of the indicator function:

$$f_i^k = |\mathbf{u}_k \cdot \mathbf{g}_i|^2 \quad (6.1)$$

Therefore, each image is obtained using a single eigenimage in the projection with  $\mathbf{g}$ . This allows the visualization of the information present in each one of them. The first eigenimage projection (second column) has mostly the average of the image but with poor contrast between the background and signal. As discussed before, a higher-order eigenimage should contain more structural information but also it may contain more noise. This can be seen when looking at the 10<sup>th</sup> eigenimage projection. This shows that most of the content of this eigenimage is in the background.

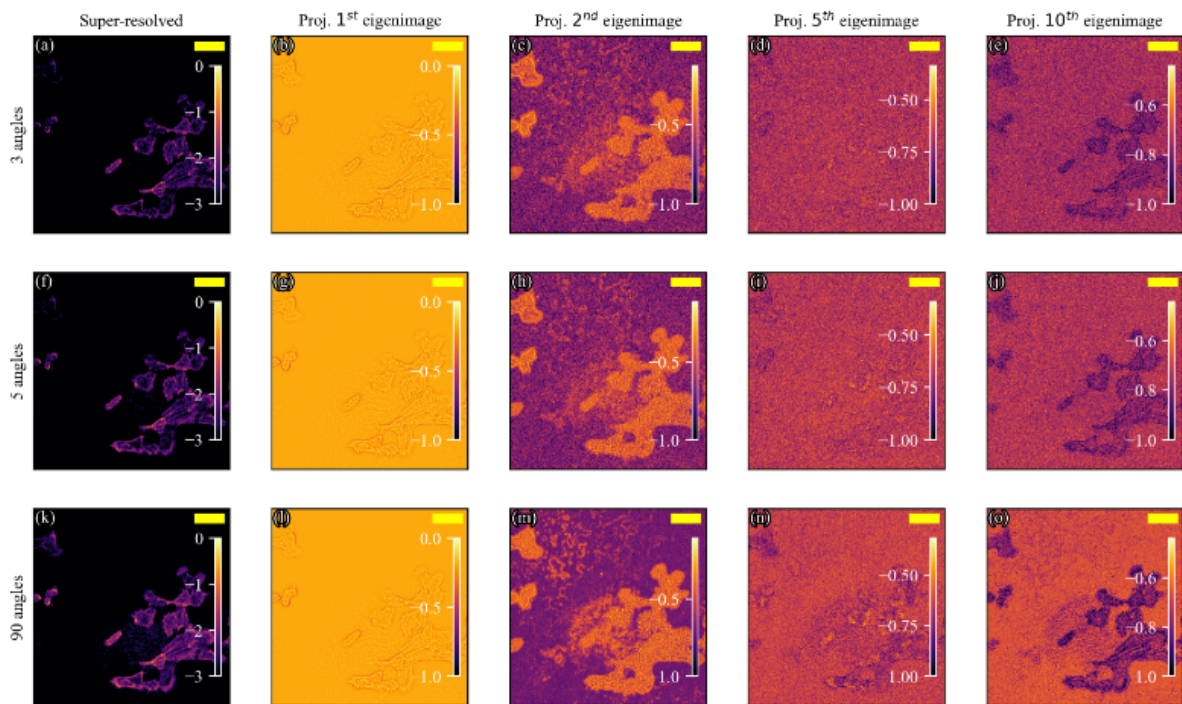


Figure 6.9: Eigenimages of different orders for the entire field of view. The 5th eigenimage seems to contain more information the more frames are used. This is no longer the case for the 10th eigenimage which seems to be noise regardless of the number of frames used. .Reproduced from [80]



# Chapter 7

## Conclusion and future work

Multiple signal classification algorithm (MUSICAL) has come a long way since its conception. For instance, concepts like threshold and singular value plots have slowly fallen into the back seat as they have become less relevant for new users who can just use soft MUSICAL or contrast enhancement. With the advent of the new indicator function, the workflow has been simplified greatly and the algorithm has become almost a one-click solution for a lot of microscopy imaging systems. This thesis and the work presented through the 5 articles included in it are a manifest of this story.

With the results of Paper I, new users found an easy way of experimenting with MUSICAL on their own data. The challenge of using the algorithm then became less about using MATLAB and more about using the algorithm. With MusiJ users can obtain MUSICAL results and immediately assess whether their data is suitable or not. The problem is then, the number of reconstructions that need to be done before finding the suitable parameter of threshold. Then Paper II was published, providing deeper insight into the meaning of the threshold and the eigenimages. But by far its greatest contribution was the establishment of a general indicator function with its more notable member being the one called soft MUSICAL. Then, Paper III was published showing even more potential indicator functions for a different usage. Paper IV and Paper V then came to show the versatility of MUSICAL for different optical systems which may encourage new researchers and new optical systems to be tested with MUSICAL. Last but not least, this thesis includes the first results on 3D MUSICAL which should only grow on applications from this point.

### 7.1 Future work

There are four research fields that could be followed from the point this thesis ends:

- Quantification of the method: MUSICAL is a qualitative method. This means that while it allows the retrieval of super-resolution images, quantification of its resolution it's challenging. Moreover, this also hinders the quantification of its resolution which is the reason why most of the analysis relies on simulations. A different type of indicator function could be designed to make the method quantitative.

- 3D microscopy: It has been tested that MUSICAL may be used in structure illumination data. However, this means that raw data only have 15 different illuminations. Enlarging the number of frames through a different type of illumination could potentially improve the results. Some alternatives could be multi-plane acquisition which has already been proven with SOFI [81].
- Alternative decomposition: MUSICAL uses SVD. This provides a quick method for obtaining an orthogonal basis and posterior separation of components. However, SVD is not the only decomposition method available for matrices. One of its problems is that through SVD, the eigenimages may contain negative values. Since real images only contain non-negative values, the interpretation is not straightforward. An alternative could be Non-negative matrix factorization [82]. A second pitfall is that SVD only works with 2D matrices. This requirement means that 2D data must be restructured to a single 1D vector, which does not use the spatial relation between close pixels. By implementing a new type of decomposition both problems could be tackled.
- Derivation of a lower bound for resolution: Reconstruction's resolution is quantified after the reconstruction is performed and, in most cases, in relation to the noise. However, a more rigorous treatment of the resolution problem is lacking. The indicator function relies on the noise part being almost zero for a test point in order to show a structure. This zero may result in spurious peaks and therefore affect the quantification of resolution as a peak may have infinite components in Fourier and therefore suggest infinite resolution.

## 7.2 Final remarks

This thesis contains most of the relevant advancements in MUSICAL. Nevertheless, this corresponds to just a small fraction of its possibilities. The work presented here should serve as an introduction to MUSICAL's strengths but also what its limitations are. However, these limitations are at the same time opportunities for further development. Developments that I hope to motivate with this thesis. Through the work presented here, I hope to have cemented a prosperous future for the algorithm and to showcase all its possibilities.

# Bibliography

- [1] “Milestones in light microscopy,” *Nature Cell Biology*, vol. 11, pp. 1165–1165, Oct 2009.
- [2] M. G. Gustafsson, “Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy,” *Journal of Microscopy*, vol. 198, no. 2, pp. 82–87, 2000.
- [3] S. W. Hell and J. Wichmann, “Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy,” *Optics Letters*, vol. 19, no. 11, pp. 780–782, 1994.
- [4] M. Lelek, M. T. Gyparaki, G. Beliu, F. Schueder, J. Griffié, S. Manley, R. Jungmann, M. Sauer, M. Lakadamyali, and C. Zimmer, “Single-molecule localization microscopy,” *Nature Reviews Methods Primers*, vol. 1, p. 39, Jun 2021.
- [5] K. C. Gwosch, J. K. Pape, F. Balzarotti, P. Hoess, J. Ellenberg, J. Ries, and S. W. Hell, “MINFLUX nanoscopy delivers 3d multicolor nanometer resolution in cells,” *Nature Methods*, vol. 17, pp. 217–224, Feb 2020.
- [6] T. Dertinger, R. Colyer, G. Iyer, S. Weiss, and J. Enderlein, “Fast, background-free, 3d super-resolution optical fluctuation imaging (sofi),” *Proceedings of the National Academy of Sciences*, vol. 106, no. 52, pp. 22287–22292, 2009.
- [7] I. Yahiatene, S. Hennig, M. Müller, and T. Huser, “Entropy-based super-resolution imaging (ESI): From disorder to fine detail,” *Acs Photonics*, vol. 2, no. 8, pp. 1049–1056, 2015.
- [8] N. Gustafsson, S. Culley, G. Ashdown, D. M. Owen, P. M. Pereira, and R. Henriques, “Fast live-cell conventional fluorophore nanoscopy with imagej through super-resolution radial fluctuations,” *Nature Communications*, vol. 7, p. 12471, 2016.
- [9] O. Solomon, M. Mutzafi, M. Segev, and Y. C. Eldar, “Sparsity-based super-resolution microscopy from correlation information,” *Optics Express*, vol. 26, pp. 18238–18269, Jul 2018.

- [10] S. Cox, E. Rosten, J. Monypenny, T. Jovanovic-Talisman, D. T. Burnette, J. Lippincott-Schwartz, G. E. Jones, and R. Heintzmann, “Bayesian localization microscopy reveals nanoscale podosome dynamics,” *Nature Methods*, vol. 9, no. 2, p. 195, 2012.
- [11] K. Agarwal and R. Macháň, “Multiple signal classification algorithm for super-resolution fluorescence microscopy,” *Nature Communications*, vol. 7, p. 13752, Dec 2016.
- [12] G. H. Golub, “Matrix computations,” 2013.
- [13] R. Schmidt, “Multiple emitter location and signal parameter estimation,” *IEEE Transactions on Antennas and Propagation*, vol. 34, no. 3, pp. 276–280, 1986.
- [14] C. A. Schneider, W. S. Rasband, and K. W. Eliceiri, “NIH Image to ImageJ: 25 years of image analysis,” *Nat Meth*, vol. 9, pp. 671–675, July 2012.
- [15] S. F. Gibson and F. Lanni, “Experimental test of an analytical model of aberration in an oil-immersion objective lens used in three-dimensional light microscopy,” *J. Opt. Soc. Am. A*, vol. 9, pp. 154–166, Jan 1992.
- [16] C. S. Ball, “The early history of the compound microscope,” *Bios*, vol. 37, no. 2, pp. 51–60, 1966.
- [17] A. M. Reigoto, S. A. Andrade, M. C. R. R. Seixas, M. L. Costa, and C. Mermelstein, “A comparative study on the use of microscopy in pharmacology and cell biology research,” *PLOS ONE*, vol. 16, pp. 1–13, 01 2021.
- [18] P. W. Hawkes and J. C. H. Spence, *Springer Handbook of Microscopy*. Springer Handbooks, Cham: Springer International Publishing AG, 2019.
- [19] K. M. Yip, N. Fischer, E. Paknia, A. Chari, and H. Stark, “Atomic-resolution protein structure determination by cryo-em,” *Nature*, vol. 587, pp. 157–161, Nov 2020.
- [20] J. D. Gaskill, “Linear systems, fourier transforms, and optics,” 1978.
- [21] J. W. Goodman, *Introduction to Fourier optics / Joseph W. Goodman*. Englewood, Colo: Roberts & Co., 3rd ed. ed., 2005.
- [22] K. C. Qiang Wu, Fatima Merchant, *Microscope Image Processing*. San Diego: Elsevier Science, 1 ed., 2010.
- [23] J. Li, F. Xue, and T. Blu, “Fast and accurate three-dimensional point spread function computation for fluorescence microscopy,” *J. Opt. Soc. Am. A*, vol. 34, pp. 1029–1034, Jun 2017.

- [24] H. Babcock and K. M. Douglass, “MicroscPSF-Py.” <https://github.com/MicroscPSF/MicroscPSF-Py>, 2018. Accessed: 2023-10-25.
- [25] M. Born and E. Wolf, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*. Saint Louis: Elsevier Science & Technology, 6 ed., 1980.
- [26] P. Török, P. Varga, Z. Laczik, and G. R. Booker, “Electromagnetic diffraction of light focused through a planar interface between materials of mismatched refractive indices: an integral representation,” *J. Opt. Soc. Am. A*, vol. 12, pp. 325–332, Feb 1995.
- [27] E. Abbe Hon., “VII.—on the estimation of aperture in the microscope.,” *Journal of the Royal Microscopical Society*, vol. 1, no. 3, pp. 388–423, 1881.
- [28] Rayleigh, “XXXI. investigations in optics, with special reference to the spectroscope,” *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 8, no. 49, pp. 261–274, 1879.
- [29] B. R. Masters, “Superresolution optical microscopy : The quest for enhanced resolution and contrast,” 2020.
- [30] A. Rizzi, T. Algeri, G. Medeghini, and D. Marini, “A proposal for contrast measure in digital images,” pp. 187–192, 01 2004.
- [31] M. Bertero and P. Boccacci, “Introduction to inverse problems in imaging,” CRC Press, 2020.
- [32] E. Hecht, “Optics,” 2017.
- [33] G. Casella, “Statistical inference,” 2002.
- [34] J. Boulanger, C. Kervrann, P. Bouthemy, P. Elbau, J.-B. Sibarita, and J. Salamero, “Patch-based nonlocal functional for denoising fluorescence microscopy image sequences,” *IEEE Transactions on Medical Imaging*, vol. 29, no. 2, pp. 442–454, 2010.
- [35] I. Ashraf, “Molecular energy levels.” <https://indico.ictp.it/event/8007/session/15/contribution/82/material/slides/0.pdf>, 2017. Accessed: 2023-10-21.
- [36] A. P. Demchenko, “Photobleaching of organic fluorophores: quantitative characterization, mechanisms, protection\*,” *Methods and Applications in Fluorescence*, vol. 8, p. 022001, feb 2020.
- [37] J. W. Verhoeven, “Glossary of terms used in photochemistry (iupac recommendations 1996),” *Pure and Applied Chemistry*, vol. 68, no. 12, pp. 2223–2286, 1996.

- [38] I. S. Opstad, “3DSIM data of mitochondria in the cardiomyoblast cell-line H9c2 adapted to either glucose or galactose,” 2021.
- [39] R. Datta, T. M. Heaster, J. T. Sharick, A. A. Gillette, and M. C. Skala, “Fluorescence lifetime imaging microscopy: fundamentals and advances in instrumentation, analysis, and applications,” *Journal of Biomedical Optics*, vol. 25, no. 7, p. 071203, 2020.
- [40] S. Santos, K. K. Chu, D. Lim, N. Bozinovic, T. N. Ford, C. Hourtoule, A. C. Bartoo, S. K. Singh, and J. Mertz, “Optically sectioned fluorescence endomicroscopy with hybrid-illumination imaging through a flexible fiber bundle,” *Journal of Biomedical Optics*, vol. 14, no. 3, p. 030502, 2009.
- [41] T. Mangeat, S. Labouesse, M. Allain, A. Negash, E. Martin, A. Guénolé, R. Poincloux, C. Estibal, A. Bouissou, S. Cantaloube, E. Vega, T. Li, C. Rouvière, S. Allart, D. Keller, V. Debarnot, X. B. Wang, G. Michaux, M. Pinot, R. Le Borgne, S. Tournier, M. Suzanne, J. Idier, and A. Sentenac, “Super-resolved live-cell imaging using random illumination microscopy,” *Cell Reports Methods*, vol. 1, no. 1, p. 100009, 2021.
- [42] Y. Wan, K. McDole, and P. J. Keller, “Light-sheet microscopy and its potential for understanding developmental processes,” *Annual Review of Cell and Developmental Biology*, vol. 35, no. 1, pp. 655–681, 2019. PMID: 31299171.
- [43] M. G. L. Gustafsson, “Nonlinear structured-illumination microscopy: Wide-field fluorescence imaging with theoretically unlimited resolution,” *Proceedings of the National Academy of Sciences*, vol. 102, no. 37, pp. 13081–13086, 2005.
- [44] Y. C. Eldar, *Sampling Theory: Beyond Bandlimited Systems*. USA: Cambridge University Press, 1st ed., 2015.
- [45] S. Marple, *Digital Spectral Analysis: Second Edition*. Dover Books on Electrical Engineering, Dover Publications, 2019.
- [46] G. Popescu, *Quantitative Phase Imaging of Cells and Tissues*. New York: McGraw-Hill Education, 1st edition ed., 2011.
- [47] G. Zheng, C. Shen, S. Jiang, P. Song, and C. Yang, “Concept, implementations and applications of fourier ptychography,” *Nature Reviews Physics*, vol. 3, pp. 207–223, Mar 2021.
- [48] M. Levoy, R. Ng, A. Adams, M. Footer, and M. Horowitz, “Light field microscopy,” *ACM Trans. Graph.*, vol. 25, p. 924–934, jul 2006.
- [49] A. Girsault, T. Lukes, A. Sharipov, S. Geissbuehler, M. Leutenegger, W. Vandenberg, P. Dedecker, J. Hofkens, and T. Lasser, “Sofi simulation tool: A software package for simulating and testing super-resolution optical fluctuation imaging,” *PLOS ONE*, vol. 11, pp. 1–13, 09 2016.

- [50] T. Novák, T. Gajdos, J. Sinkó, G. Szabó, and M. Erdélyi, “Teststorm: Versatile simulator software for multimodal super-resolution localization fluorescence microscopy,” *Scientific Reports*, vol. 7, p. 951, Apr 2017.
- [51] W. Hossack, “Digital image analysis, lecture notes: Digital reconstruction.” <https://www2.ph.ed.ac.uk/~wjh/teaching/dia/documents/reconstruction.pdf>. Accessed: 2023-10-22.
- [52] W. H. Richardson, “Bayesian-based iterative method of image restoration\*,” *J. Opt. Soc. Am.*, vol. 62, pp. 55–59, Jan 1972.
- [53] L. B. Lucy, “An iterative technique for the rectification of observed distributions,” *The Astronomical Journal*, vol. 79, pp. 745–754, 1974.
- [54] E. Betzig, G. H. Patterson, R. Sougrat, O. W. Lindwasser, S. Olenych, J. S. Bonifacino, M. W. Davidson, J. Lippincott-Schwartz, and H. F. Hess, “Imaging intracellular fluorescent proteins at nanometer resolution,” *Science*, vol. 313, no. 5793, pp. 1642–1645, 2006.
- [55] M. J. Rust, M. Bates, and X. Zhuang, “Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (storm),” *Nature Methods*, vol. 3, pp. 793–796, Oct 2006.
- [56] A. Sharonov and R. M. Hochstrasser, “Wide-field subdiffraction imaging by accumulated binding of diffusing probes,” *Proceedings of the National Academy of Sciences*, vol. 103, no. 50, pp. 18911–18916, 2006.
- [57] M. G. Kendall, “The advanced theory of statistics : Vol. 1 : Distribution theory,” 1969.
- [58] M. Pawlowska, R. Tenne, B. Ghosh, A. Makowski, and R. Lapkiewicz, “Embracing the uncertainty: the evolution of SOFI into a diverse family of fluctuation-based super-resolution microscopy methods,” *Journal of Physics: Photonics*, vol. 4, p. 012002, Jan. 2022.
- [59] T. Dertinger, R. Colyer, R. Vogel, J. Enderlein, and S. Weiss, “Achieving increased resolution and more pixels with superresolution optical fluctuation imaging (sofi),” *Opt. Express*, vol. 18, pp. 18875–18885, Aug 2010.
- [60] S. Geissbuehler, N. L. Bocchio, C. Dellagiacomma, C. Berclaz, M. Leutenegger, and T. Lasser, “Mapping molecular statistics with balanced super-resolution optical fluctuation imaging (bsofi),” *Optical Nanoscopy*, vol. 1, p. 4, Apr 2012.
- [61] S. C. Stein, *Advanced Data Processing in Super-resolution Microscopy*. PhD thesis, the Georg-August-Universität Göttingen, 2017.

- [62] W. Zhao, S. Zhao, Z. Han, X. Ding, G. Hu, L. Qu, Y. Huang, X. Wang, H. Mao, Y. Jiu, Y. Hu, J. Tan, X. Ding, L. Chen, C. Guo, and H. Li, “Enhanced detection of fluorescence fluctuations for high-throughput super-resolution imaging,” *Nature Photonics*, vol. 17, pp. 806–813, Sep 2023.
- [63] I. S. Opstad, S. Acuña, L. E. V. Hernandez, J. Cauzzo, N. Škalko Basnet, B. S. Ahluwalia, and K. Agarwal, “Fluorescence fluctuations-based super-resolution microscopy techniques: an experimental comparative study,” 2020.
- [64] X. Chen, *Computational Methods for Electromagnetic Inverse Scattering*. Wiley - IEEE, Newark: Wiley, 1 ed., 2018.
- [65] J. M. J. Devaney, “Super-resolution Processing of Multi-static Data Using Time Reversal and MUSIC A,” 2000.
- [66] S. Acuña, I. S. Opstad, F. Godtlielsen, B. S. Ahluwalia, and K. Agarwal, “Soft thresholding schemes for multiple signal classification algorithm,” *Opt. Express*, vol. 28, pp. 34434–34449, Nov 2020.
- [67] S. Acuña, M. Roy, L. E. Villegas-Hernández, V. K. Dubey, B. S. Ahluwalia, and K. Agarwal, “Deriving high contrast fluorescence microscopy images through low contrast noisy image stacks,” *Biomed. Opt. Express*, vol. 12, pp. 5529–5543, Sep 2021.
- [68] T. Furnival, R. K. Leary, and P. A. Midgley, “Denoising time-resolved microscopy image sequences with singular value thresholding,” *Ultramicroscopy*, vol. 178, pp. 112–124, 2017. FEMMS 2015.
- [69] Y. Deng, M. Sun, P.-H. Lin, J. Ma, and J. W. Shaevitz, “Spatial covariance reconstructive (score) super-resolution fluorescence microscopy,” *PLOS ONE*, vol. 9, pp. 1–9, 04 2014.
- [70] S. Acuña, F. Ströhl, I. S. Opstad, B. S. Ahluwalia, and K. Agarwal, “Musij: an imagej plugin for video nanoscopy,” *Biomed. Opt. Express*, vol. 11, pp. 2548–2559, May 2020.
- [71] K. Agarwal and D. K. Prasad, “Eigen-analysis reveals components supporting super-resolution imaging of blinking fluorophores,” *Scientific Reports*, vol. 7, p. 4445, Jun 2017.
- [72] X. Li, “Three-dimensional structured illumination microscopy with enhanced axial resolution,” Nov. 2022.
- [73] X. Li, Y. Wu, Y. Su, I. Rey-Suarez, C. Matthaeus, T. B. Updegrave, Z. Wei, L. Zhang, H. Sasaki, Y. Li, M. Guo, J. P. Giannini, H. D. Vishwasrao, J. Chen, S.-J. J. Lee, L. Shao, H. Liu, K. S. Ramamurthi, J. W. Taraska, A. Upadhyaya, P. La Riviere, and H. Shroff, “Three-dimensional structured illumination microscopy with enhanced axial resolution,” *Nature Biotechnology*, vol. 41, pp. 1307–1319, Sep 2023.



- [74] K. Agarwal, “Musical.” <https://sites.google.com/site/uthkrishth/musical>. Accessed: 2023-10-10.
- [75] L. E. Villegas-Hernández, V. Dubey, M. Nystad, J.-C. Tinguely, D. A. Coucheron, F. T. Dullo, A. Priyadarshi, S. Acuña, A. Ahmad, J. M. Mateos, G. Barmettler, U. Ziegler, Å. B. Birgisdottir, A.-M. K. Hovd, K. A. Fenton, G. Acharya, K. Agarwal, and B. S. Ahluwalia, “Chip-based multimodal super-resolution microscopy for histological investigations of cryopreserved tissue sections,” *Light: Science & Applications*, vol. 11, p. 43, Feb 2022.
- [76] I. S. Opstad, D. H. Hansen, S. A. na, F. Ströhl, A. Priyadarshi, J.-C. Tinguely, F. T. Dullo, R. A. Dalmo, T. Seternes, B. S. Ahluwalia, and K. Agarwal, “Fluorescence fluctuation-based super-resolution microscopy using multimodal waveguided illumination,” *Opt. Express*, vol. 29, pp. 23368–23380, Jul 2021.
- [77] B. Mandracchia, X. Hua, C. Guo, J. Son, T. Urner, and S. Jia, “Fast and accurate scmos noise correction for fluorescence microscopy,” *Nature Communications*, vol. 11, p. 94, Jan 2020.
- [78] J.-C. Tinguely, Ø. I. Helle, and B. S. Ahluwalia, “Silicon nitride waveguide platform for fluorescence microscopy of living cells,” *Opt. Express*, vol. 25, pp. 27678–27690, Oct 2017.
- [79] I. S. Opstad, “Replication data for: Fluorescence fluctuation-based super-resolution microscopy using multimodal waveguided illumination,” 2021.
- [80] A. Butola, S. Acuna, D. H. Hansen, and K. Agarwal, “Scalable-resolution structured illumination microscopy,” *Opt. Express*, vol. 30, pp. 43752–43767, Nov 2022.
- [81] S. Geissbuehler, A. Sharipov, A. Godinat, N. L. Bocchio, P. A. Sandoz, A. Huss, N. A. Jensen, S. Jakobs, J. Enderlein, F. Gisou van der Goot, E. A. Dubikovskaya, T. Lasser, and M. Leutenegger, “Live-cell multiplane three-dimensional super-resolution optical fluctuation imaging,” *Nature Communications*, vol. 5, p. 5830, Dec 2014.
- [82] I. S. Dhillon and S. Sra, “Generalized nonnegative matrix approximations with bregman divergences,” in *Proceedings of the 18th International Conference on Neural Information Processing Systems*, NIPS’05, (Cambridge, MA, USA), p. 283–290, MIT Press, 2005.



## Appendix: Included articles

- Paper I: MusiJ: an ImageJ plugin for video nanoscopy
- Paper II: Soft thresholding schemes for multiple signal classification algorithm
- Paper III: Deriving high contrast fluorescence microscopy images through low contrast noisy image stacks
- Paper IV: Fluorescence fluctuation-based super-resolution microscopy using multi-modal waveguided illumination
- Paper V: Scalable-resolution structured illumination microscopy



# Paper I

MusiJ: an ImageJ plugin for video nanoscopy





# MusiJ: an ImageJ plugin for video nanoscopy

SEBASTIAN ACUÑA,<sup>1,3</sup> FLORIAN STRÖHL,<sup>1,3</sup> ID IDA S. OPSTAD,<sup>1</sup>  
BALPREET S. AHLUWALIA,<sup>1,2</sup> ID AND KRISHNA AGARWAL<sup>1,\*</sup>

<sup>1</sup>Department of Physics and Technology, UiT The Arctic University of Norway, NO-9037 Tromsø, Norway

<sup>2</sup>Department of Clinical Science, Intervention and Technology, Karolinska Institute, 17177 Stockholm, Sweden

<sup>3</sup>These authors contributed equally

\*krishna.agarwal@uit.no

**Abstract:** We present an open-source implementation of the fluctuation-based nanoscopy method MUSICAL for ImageJ. This implementation improves the algorithm's computational efficiency and takes advantage of multi-threading to provide orders of magnitude faster reconstructions than the original MATLAB implementation. In addition, the plugin is capable of generating super-resolution videos from large stacks of time-lapse images via an interleaved reconstruction, thus enabling easy-to-use multi-color super-resolution imaging of dynamic systems.

© 2020 Optical Society of America under the terms of the [OSA Open Access Publishing Agreement](#)

## 1. Introduction

The past two decades have witnessed a huge development in nanoscopy techniques that allow to surpass the resolution limit of optical microscopy and thus provide *super-resolution* [1]. One way of classifying the broad range of those techniques is to distinguish all-optical, hybrid, and purely computational approaches. All-optical nanoscopy techniques manage to shrink the effective point spread function before detection and include stimulated emission depletion (STED) microscopy [2] or instant structured illumination microscopy (iSIM) [3]. Hybrid and purely computational techniques, in contrast, make use of temporal changes in the sample's fluorescent emission profile and extract additional information from time-series of raw frames of the same underlying sample structure. Such changes can be induced extrinsically via spatially varying illumination patterns [4,5] or intrinsically by exploiting fluorophore photokinetics that result in fluctuations in the measured fluorescence intensity [6,7].

Despite the often simplified optical setup of computational and hybrid nanoscopy techniques in comparison to all-optical ones, a lack of user-friendly and open-source implementations has often hindered fast integration of nanoscopy into biological research routines. A prime example for such a delay is the case of structured illumination microscopy (SIM). Its complex reconstruction algorithm was published in 2000, just to be implemented anew countless times in microscopy laboratories all over the world until the release of the easy-to-use FairSIM plugin [8], more than 15 years after the original publication. Further, as all nanoscopy techniques vary in their strengths and weaknesses, it is desirable to make as many different techniques available as possible, ideally in a single standard analysis environment. For microscopy, this environment is the image processing toolbox ImageJ [9] or its advanced version Fiji (*Fiji is just ImageJ*) [10].

Akin to the SIM reconstruction plugin FairSIM, single molecule localisation microscopy (SMLM) software has been made freely available by a vibrant community (for a comprehensive list see [11–13]). Similar to SIM, SMLM techniques can be regarded as hybrid nanoscopy methods due to the requirements of multiple high-power lasers and additional optical elements for field flattening [14] or when 3D information is desired [15]. A notable exception are 3D SMLM techniques with purely computational 3D information extraction based on aberrations in the microscope's point spread function (PSF) - these are also available in Fiji [16]. Despite their impressive resolution, SMLM reconstruction algorithms require data sets comprising thousands

of raw frames with sparse single molecule blinking events, which renders live-cell, let alone time-lapse imaging challenging.

Approaches to extract sub-diffraction features from data sets with densely packed emitters and with well below thousand raw frames taken on conventional, non-specialized microscopes, can be grouped as fluctuation-based nanoscopy techniques. Fluctuation-based algorithms exploit small intensity variations in time-series via statistical analyses to generate better resolved images. The actual algorithm depends on the statistical approach, thus giving rise to various flavors of this idea. Examples include SOFI (super-resolution optical fluctuation imaging) [17], 3B (Bayesian analysis of blinking and bleaching) [18], ESI (entropy-based super-resolution imaging) [19], and SRRF (super-resolution radial fluctuations) [20]. All of the above mentioned algorithms (except SOFI) have been translated to ImageJ. The recently developed fluctuation-based method MUSICAL (multiple signal classification algorithm) [21], however, has not been translated to ImageJ until now.

MUSICAL's origin can be traced to MUSIC (multiple signal classification) first developed for direction of arrival measurements [22]. In MUSIC, the number of independent sources, often equal to the number of targets (aeroplanes or ships for example), is determined by the number of non-zero eigenvalues if the number of sources is less than the number of independent measurements taken at multiple time instances over an array of radar/sonar sensors. The multiple time instances provide multiple independent measurements from the sources and the problem of determining the direction of arrival is an inverse source problem. An indicator function is computed for each candidate point (also called test point)  $r_{\text{test}}$  in the target region by taking the reciprocal of a distance function  $d_n$ .

$$f(r_{\text{test}}) = \frac{1}{d_n(r_{\text{test}})} \quad (1)$$

The distance function  $d_n$  is the projection of the expected output vector for a test point onto the eigenvectors with zero eigenvalues. The presence and direction of arrival of the target is then indicated if the indicator function has a large value at a candidate target point. MUSIC has survived the test of times and is constantly being reinvented for modern applications such as cognitive radars [23] and inverse imaging in the microwave domain [24,25]. In the case of inverse imaging in the microwave domain, the full electromagnetic wave model of scattering applies and a multiple input multiple output system is used for taking measurements. Compared to the case of radars, the problem here is an inverse scattering problem and not an inverse source problem. This means the number of non-zero eigenvalues indicates the number of independent dipoles induced on the scatterers, as long as the number of transmitters and receivers is more than the number of induced dipoles. This number is equal to or more than the number of scatterers, depending upon the degrees of freedom for the induced dipole, which in turn depends upon isotropicity of the scatterers as well as the polarization constraints on the incident waves [24].

In the case of optical microscopy or nanoscopy, the measurements are comprised of fluorescent intensity detections on an optical detector array (i.e. the microscope's camera) at multiple time instances. At each time instance, the number of photons emitted by any fluorescent molecule is independent of the other molecules and follows a statistical distribution [26]. Therefore, the underlying problem is an inverse source problem like as in MUSIC. Nonetheless, adapting MUSIC to MUSICAL for optical nanoscopy is a non-trivial task and only a short intuitive understanding shall be given here without resorting to the detailed mathematical formulation.

Consider a small window around a given pixel of the size of the point spread function (PSF) of the microscope as the region of influence for the fluorescent emitters within the region of that pixel. Despite the main region of interest being within the pixel, the structure on which fluorophores are attached may extend beyond it. Therefore, it is important that the MUSIC indicator function is computed over the entire window. This is applicable to a first approximation. Nevertheless, as PSFs has not hard boundary, the PSF of fluorophores never lie completely



within the window but will stretch outside. Similarly, the window may contain data from the trailing part of the PSFs of fluorophores completely outside the window. This non-reliability is suppressed in MUSICAL by using a soft window function on the measurement in the window as well as the PSF of candidate locations of emitters (i.e. the test points). This is equivalent to weighing the indicator function at a test point on the basis of distance from the center pixel of the window. Moreover, for stitching the reconstructions of all the windows, instead of using conventional image-processing techniques, MUSICAL takes a physics-based route in which an additional distance metric  $d_s$  in the numerator is the indicator function.

$$f(r_{\text{test}}) = \frac{d_s(r_{\text{test}})}{d_n(r_{\text{test}})} \quad (2)$$

The distance function  $d_s$  is the projection of the expected output vector for a candidate target point on the eigenvectors with non-zero eigenvalues. Introducing this distance is equivalent to stitching the reconstructed images based on the energy contributed by the test point in the numerical space of measurements.

In essence, MUSICAL identifies spatio-temporal patterns present in the image sequence through patch-wise singular value decomposition. A manually selected threshold then partitions the spatio-temporal patterns (i.e. eigenvectors) into two sets, 'signal' set  $Q_s$  that contains those eigenvectors whose corresponding singular value is larger than the threshold and 'noise' set  $Q_n$  containing the remaining eigenvectors. This is illustrated in Fig. 1). The final MUSICAL values are then computed as the ratio between sub-sampled image patches projected onto signal and noise vectors. A detailed derivation can be found in Appendix A and in the original MUSICAL publication [21].

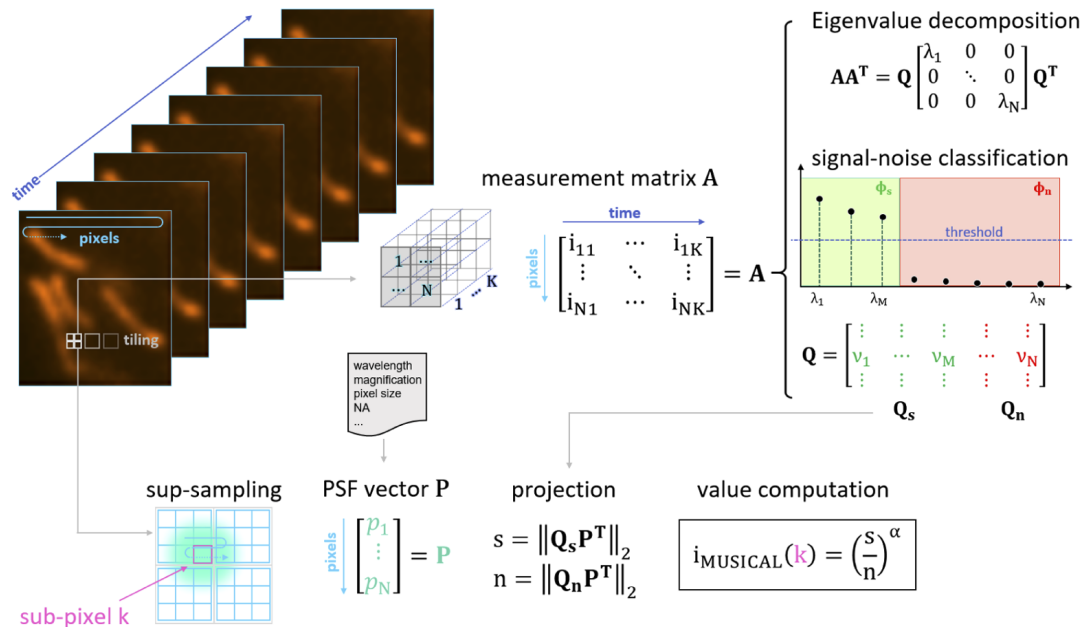


Fig. 1. Summary of the multiple signal classification algorithm (MUSICAL) [21].

MUSICAL was initially implemented in MATLAB with a focus on code-readability with respect to the mathematical background of the technique, rather than computational efficiency. Also, the MATLAB version provided only a rudimentary user interface with no extended capabilities for video generation or multi-color imaging. Hence, for a successful translation into a handy tool, we have developed MusiJ, a plugin for Fiji that improves both on the front end and back end of the original MATLAB implementation in several ways.

## 2. MusiJ: MUSICAL for ImageJ

### 2.1. Back end

MusiJ has three main differences compared to the previously released MATLAB version in terms of back end implementation of the algorithm. The most basic change is the data type. MATLAB uses by default double-precision floating-point format, defined by the IEEE Standard 754 [27] (named as *binary64* from 2008) which means that every value requires 64 bits of memory. In contrast, our implementation works with the single-precision floating-point or *binary32* format which halves the memory usage and speeds up individual computation steps. Although, in principle, this comes at the cost of numerical precision, we found no noticeable difference in image quality between the outputs generated by the two data types in practice. The second change is in the computation of MUSICAL image values  $i_{\text{MUSICAL}}$  (also called the indicator function) during image synthesis (see Fig. 1). In order to compute the values  $\mathbf{s}$  and  $\mathbf{n}$  for the indicator function  $i_{\text{MUSICAL}}$ , we perform eigenvalue decomposition to obtain  $\mathbf{Q}$ . The columns of  $\mathbf{Q}$  correspond to a basis with orthonormal columns. Thanks to the Pythagorean theorem, it is therefore sufficient to compute only one of them since the vector  $\mathbf{P}$  is projected into the subspace spanned by  $\mathbf{Q}_s$  and its orthogonal complement  $\mathbf{Q}_n$ . Hence, the following holds:

$$\|\mathbf{P}\|^2 = \|\mathbf{Q}_s \mathbf{P}^T\|^2 + \|\mathbf{Q}_n \mathbf{P}^T\|^2. \quad (3)$$

In practice, the cardinality of  $\mathbf{Q}_s$  is significantly smaller than that of  $\mathbf{Q}_n$ . Moreover, the PSF vector  $\mathbf{P}$  is purely defined by optical system parameters. Therefore, we redefine the indicator function in its equivalent form given in Eq. (4). This permits a reduction in the number of operations by computing the norm of  $\mathbf{P}$  in advance.

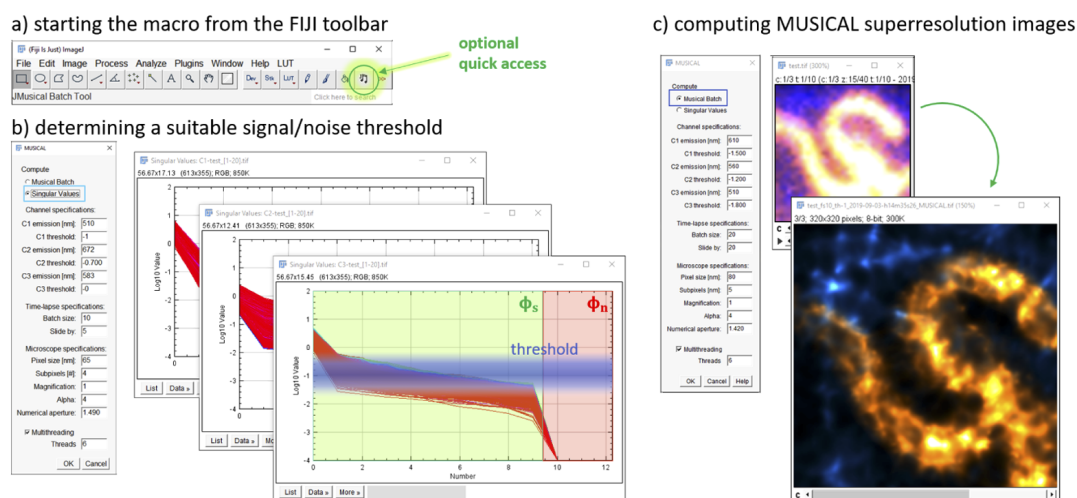
$$i_{\text{MUSICAL}} = \left( \frac{\mathbf{s}^2}{\|\mathbf{P}\|^2 - \mathbf{s}^2} \right)^{\frac{\alpha}{2}}. \quad (4)$$

The final improvement is via multi-threading. As the image contains many non-overlapping regions, it is possible to process them in different threads of execution simultaneously, before merging the results into a single final image. This improvement is available as an option to the user, and the user may specify the number of used threads based on their system configuration and the load that the system may be experiencing due to other applications executing concurrently.

### 2.2. Front end

Along with changes in the computational efficiency in MusiJ, the developed plugin offers a range of features to simplify the usage and adds to MUSICAL's capabilities. The most prominent feature is the graphical user interface, GUI, which is shown in Fig. 2.

It provides easy access to the plugin's two main functions: (1) singular value computation and (2) MUSICAL image computation. Note that eigenvalues are the squares of singular values and thus equivalent for the purpose of signal/noise thresholding. In accordance with the original MUSICAL publication the plugin hence displays singular values. Additionally, a quick-access button can be added to the Fiji toolbar, which is especially convenient for heavy use. In the main GUI window, all necessary parameters can be filled in for MUSICAL image computation. It is possible and recommended to change the values stored as default in the accompanying MusiJ macro when the same parameters are in regular use, for example for repeat-experiments on the same microscope. This is to save time and to avoid typographical errors. The required thresholds to separate 'signal' from 'noise' (one per color channel) is estimated through visual inspection of the singular value plots and normally computed before the MUSICAL image generation. At the top of the GUI, the user may select among two options ('Singular Values' or 'Musical Batch'). The first option allows visual inspection and selection of the threshold value to separate signal



**Fig. 2.** The graphical user interface of MusiJ. (a) The plugin can be found in the regular 'Plugins' tab or directly launched via the optional quick-access button. (b) Using the raw data and microscope parameters as input, singular values can be computed that allow the user to find suitable thresholds for signal/noise classification. The blue region in the image above shows a region in which a suitable threshold is likely to lie. Note that each color channel has its own threshold. (c) With the determined thresholds, a nanoscopy image or a time-series can be computed from the raw data temporal stack. Here, mitochondria (orange) and microtubules (blue) are shown.

and noise. The other option allows batch processing of the entire data set for the pre-selected and specified threshold value. Under 'Time-lapse specifications', parameters for video generation can be set. For instance, if only a single image is to be reconstructed from the first 100 images of a much longer time sequence (e.g to optimize thresholds quickly), use 'Batch size' 100, and 'Slide by' a number larger than the remaining number of frames in the image stack subjected to analysis. To reconstruct super-resolved details and visualize the changes over time, use 'Slide by' equal to or smaller than the batch size. We call this feature *interleaved reconstruction*, and it allows for a time-resolution smaller than dictated by the total acquisition time of all frames used for image reconstruction. The maximum interleave of an image stack (at a significant increase in computation time) is achieved by using 'Slide by' 1. We do not recommend this as a starting point for MUSICAL video analysis. The 'Multithreading' option allows to choose how much of the computers resources to be made accessible for MUSICAL image computation. For fastest multi-thread reconstructions use 'Threads' equal to the number of CPU cores. In practice, if running MusiJ on an office computer, we recommend to not use all cores but spare some processing power for other applications to continue executing in the background. When all parameters are set, clicking on the 'OK' button generates a super-resolved MUSICAL image or time-lapse batch and saves it along with a log file of all parameters.

### 3. Results and discussion

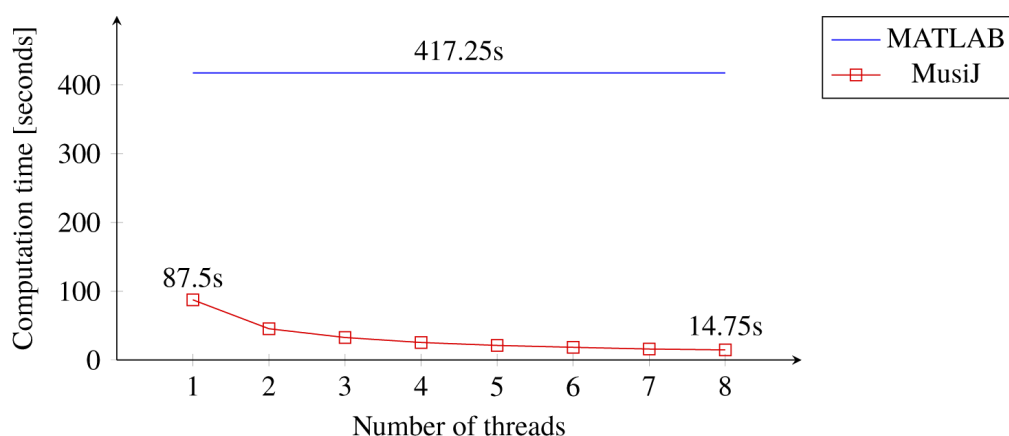
A summary of the improvements upon the MATLAB version and new capabilities only available in MusiJ is provided in Table 1.

We tested both MATLAB and ImageJ implementations on a desktop computer running Windows 10, with an Intel Xeon Gold 5118 processor (12 physical cores) and 128 GB DDR4 RAM. The MATLAB version was obtained from the official MUSICAL website (<https://sites.google.com/site/uthkrishth/musical>) and executed using MATLAB version R2018b. MusiJ was tested using FIJI 2.0.0-rc-69 with ImageJ 1.52b. For algebraic operations, MusiJ relies

**Table 1. Comparison of MUSICAL implementations for MATLAB and MusiJ.**

	MATLAB	MusiJ
processing time	417.25 s	14.75 s
graphical user interface	(✓)	✓
multi-color capability	✗	✓
video capability	✗	✓
interleaved reconstruction	✗	✓

on Nd4j version 1.0.0-beta2 using CPU as back end, with MKL 2019.1 installed. In addition, we set the number of threads used by this library to 1 by setting the *Environment Variable* `OMP_NUM_THREADS` in our system. To test the speed of both implementations, we used a  $256 \times 256$  pixel image stack with 50 frames and set the subpixel parameter to 10. For MusiJ we used an increasing number of threads from 1 to 8. The results are shown in the plot of Fig. 3 and a visual comparison of the generated MusiJ reconstruction to the MATLAB reconstruction is provided in Supplementary Figure S3.

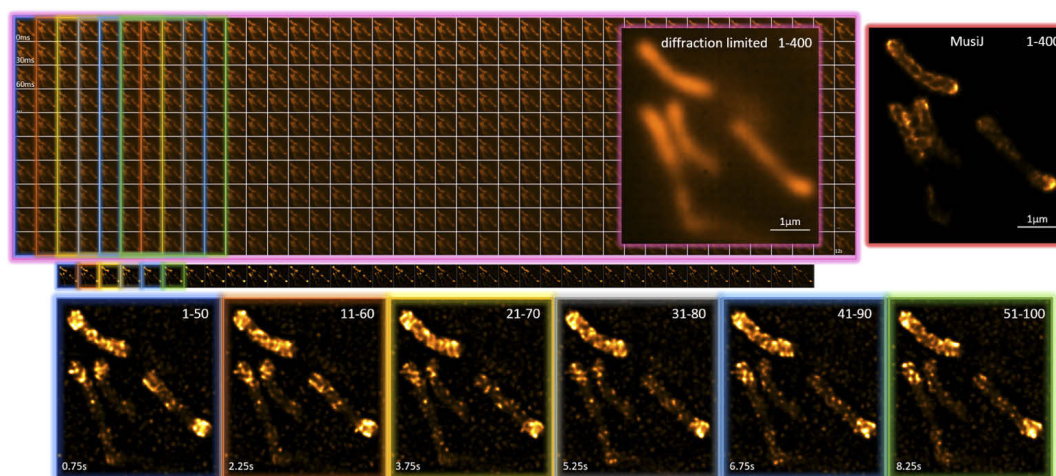


**Fig. 3.** Time used for a single reconstruction of an image of size  $256 \times 256$  pixels, 50 frames, and subpixelation of 10. The results for MATLAB were replicated to make this comparison.

Due to the increased speed of the reconstruction process, a multitude of MUSICAL frames can be computed from a long time-series with overlapping raw-frames, termed as interleaved reconstruction. This is beneficial to enhance time-resolution when the imaged objects are changing their morphology or moving fast compared to the capture time of the entire raw frame series to determine the onset of events. The exact number of raw frames used for each MUSICAL time-point has to be adapted individually to the system dynamics. As can be seen in Fig. 4, interleaved reconstruction presents a trade-off between time and spatial resolution. Figure 4 illustrates the principle of interleaved reconstruction on exemplary time-lapse image data of mitochondria.

Many sources of signal fluctuations arise in living cells in addition to the intrinsic photokinetic fluctuations of fluorescent molecules that MUSICAL relies on. This is a challenge for threshold selection and when interpreting the results. Objects moving in and out of the imaged focal plane, or any other motion of the fluorescent emitters at nanometer scales, create signal fluctuations that are picked up by the algorithm. These different sources of signal fluctuations can be a potential source of misinterpretation. Trying different thresholds and cross-checking the reconstruction results with the system dynamics visible in the raw data is thus helpful and necessary to reach interpretations consistent with both raw data and MUSICAL reconstruction. On our data, a





**Fig. 4.** Using MusiJ's interleaved reconstruction feature, dynamics of sub-cellular organelles like mitochondria can be visualised with super-resolution. In this example, a total of 400 time frames (30ms raw frame exposure time) were reconstructed into the MUSICAL image in the rightmost panel, while batches of 50 images with a 10 frame 'Slide-by' were converted into the time series displayed underneath. In the bottom six panels, zooms of the first few frames are shown. Labelled is the inner membrane of mitochondria in a cardiomyoblast cell line using dsRed. The imaging set-up was a commercial OMX microscope in widefield mode with LED illumination. A  $60\times 1.42\text{NA}$  oil immersion objective was equipped.

suitable threshold for most samples was found to be in the mid-range of the 2nd singular values (around the first *elbow* visible in the singular value plot of Fig. 2). For data with strong signal and low background fluctuations, the threshold can be set even lower to include more information in the computation for enhanced resolution. Figures S1 and S2 contain a visual comparison of images generated with different thresholds.

#### 4. Conclusion

We have presented a user-friendly implementation of the fluctuation-based super-resolution algorithm MUSICAL for ImageJ/Fiji with a significant speed-up by a factor of almost 30 compared to the previous MATLAB version. The plugin can be kept up-to-date automatically via Fiji's update site. A step-by-step tutorial for installation and usage can be found at [github.com/sebsacuna/MusiJ](https://github.com/sebsacuna/MusiJ). Fluctuation-based video nanoscopy is an advancing field, but requires further experimentation and computational speed-up for increased understanding and usability of these techniques. Hence, this plugin was created with the objective of advancing the availability and usability of computational live-cell friendly super-resolution methods.

#### Appendix A: Mathematical background of MUSICAL

MUSICAL is an algorithm that allows to obtain super-resolution from a short ( $<100$ ) sequence of frames. Here, a brief mathematical background is presented.

For a sensor with  $M$  pixels and a sample composed of  $N$  emitters, and under the assumption that emitters' locations do not change with time, the imaging model can be approximated as the matrix-vector multiplication shown in Eq. (5). This model is generalizable to moving emitters by making a hypothetical list of emitters, which take unique positions along the motion trajectory of the emitter. One hypothetical emitter is then modeled as having zero emissions at all other times

except at the time when the real emitter is at the location of the hypothetical emitter.

$$\bar{I}(t) = \begin{bmatrix} G(\vec{r}_{em}^{(1)}, \vec{r}_{im}^{(1)}) & \dots & G(\vec{r}_{em}^{(N)}, \vec{r}_{im}^{(1)}) \\ \vdots & \ddots & \vdots \\ G(\vec{r}_{em}^{(1)}, \vec{r}_{im}^{(M)}) & \dots & G(\vec{r}_{em}^{(N)}, \vec{r}_{im}^{(M)}) \end{bmatrix} \begin{bmatrix} e_1(t) \\ \vdots \\ e_N(t) \end{bmatrix} \quad (5)$$

This model defines the acquired image in time  $\bar{I}(t)$  as a column vector where each element correspond to the intensity value for every pixel. The matrix that contains the values obtained from the mapping function  $G(\vec{r}_{em}, \vec{r}_{im})$  will be referred to as  $(G)$ . The function  $G(\vec{r}_{em}, \vec{r}_{im})$  maps the intensity produced by an emitter located at  $\vec{r}_{em}$  to the pixel located in  $\vec{r}_{im}$  using the known point spread function (PSF) of the system. Finally,  $e_i(t)$  corresponds to the brightness of emitter  $i$  during time  $t$ . Note that each image is obtained then as a linear combination of the columns of  $\mathbf{G}$  which is not time dependent.

Let's consider now a sequence of  $K$  image vectors to form the matrix  $\mathbf{I}$  and corresponding Singular Value Decomposition (SVD) shown in Eq. (6). This allow us to generate an orthonormal basis for  $M$ - dimensional space of real numbers  $\mathfrak{R}^M$  given by the columns of  $\mathbf{U}$ .

$$\mathbf{I} = \mathbf{U}\mathbf{S}\mathbf{V}^T \quad (6)$$

Equation (5) and Eq (6) are two fundamental relations used by MUSICAL. The simplest case is when the number of emitters is less than the number of pixels ( $N < M$ ), and assuming  $M < K$ . In this case,  $\mathbf{G}$  has  $N$  columns, meaning that its rank can be at most  $N$ . These columns span a subspace of  $\mathfrak{R}^M$ , and this is what we will call the signal space, corresponding to all the images that a set of  $N$  emitters can produce. Another implication is that the rank of  $\mathbf{I}$  is equal to the rank of  $\mathbf{G}$ , which means that there must be  $N$  non-zero singular values. The vectors associated to these singular values then, must span the same subspace as  $\mathbf{G}$ . Alternatively, the subspace associated to the vectors with singular value zero, referred to as the null space, is orthogonal to the signal space. In this scenario, we can test if a point  $\vec{r}_s$  belongs to the set of emitters by evaluating the expression shown in Eq. (7).

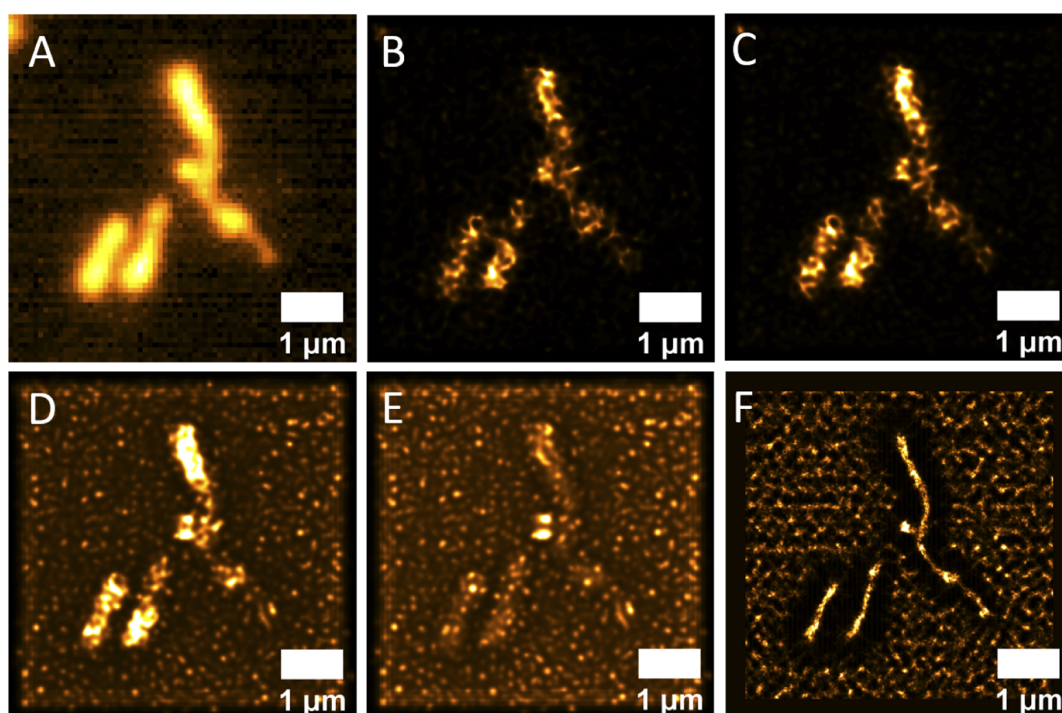
$$\bar{G}(\vec{r}_s) \cdot \bar{u}_{\sigma=0} = \begin{cases} 0 & \text{if an emitter is present at } \vec{r}_s \\ \text{non-zero} & \text{if no emitter is present at } \vec{r}_s \end{cases} \quad (7)$$

In reality, noise coming from undesired emission in the sample, shot-noise, and electronics is present in the images. Due to these factors, the singular values are unlikely to be zero. In order to split the space into signal and null space, a threshold  $\sigma_0$  is given by the user. The final function used by MUSICAL is given by Eq. (8).

$$f(\vec{r}_{test}) = \left( \frac{\sqrt{\sum_{\sigma < \sigma_0} \|G(\vec{r}_{test}) \cdot \bar{u}_i\|^2}}{\sqrt{\sum_{\sigma \geq \sigma_0} \|G(\vec{r}_{test}) \cdot \bar{u}_i\|^2}} \right)^\alpha \quad (8)$$

## Appendix B: Effect of hyper-parameters

MusiJ works over a stack of images, following a workflow similar to SRRF. Figure 5 presents a series of reconstructed images from the same source file comparing MUSICAL and SRRF. The sample used as example was cardiomyoblast cells with labelled mitochondria and is available in [28]. Since the sample presented significant motion only 50 frames were used for reconstructions. The parameters relevant for a MUSICAL reconstruction were as follows:

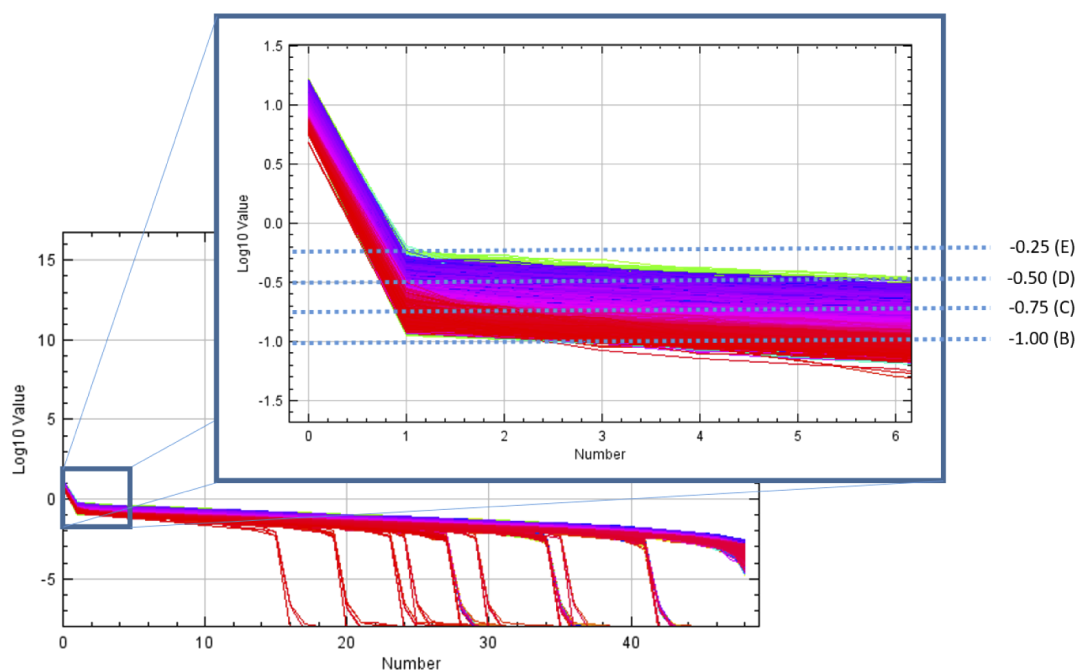


**Fig. 5.** A comparison between fluctuation images of the same raw data generated by MusiJ and SRRF. (A) average of 50 frames. (B-E) are MUSICAL reconstructions using MusiJ with thresholds -1.0, -0.75, -0.5 and -0.25. (F) correspond to a reconstruction made with SRRF.

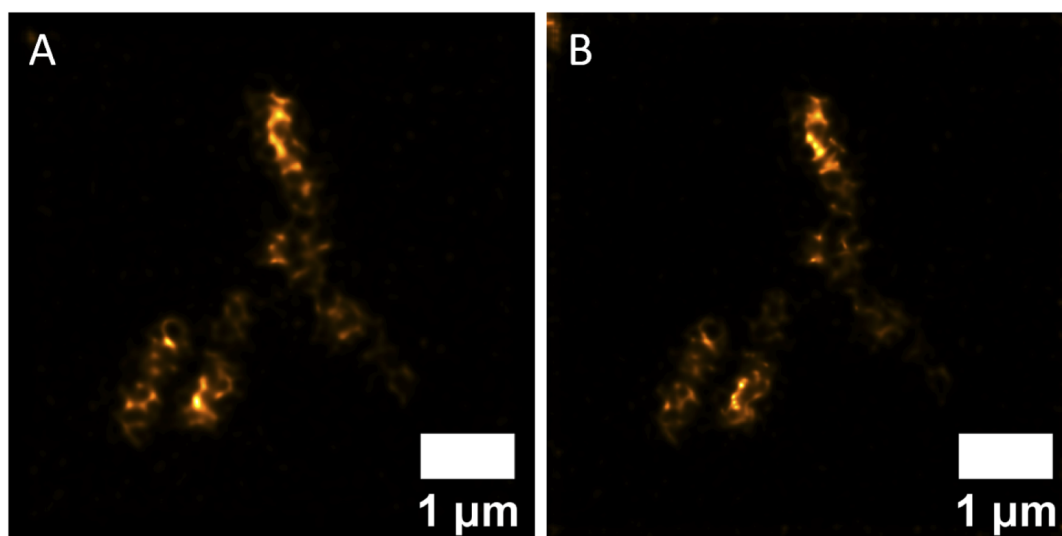
- Emission wavelength of the fluorophore: 525 nm
- Numerical aperture: 1.42
- Magnification: 1 $\times$  (as pixel size was scaled directly)
- Pixel size: 80 nm
- Exposure time per frame: 30 ms
- Number of raw image frames: 50

The parameter for alpha was set to 4, and the subpixelation to 10. In the case of SRRF the method picked was TRAC with all parameters set to default. Figure 5 address the importance of the threshold in the quality of the reconstruction. Note that as the threshold increases the relation between background and foreground get diffused. The corresponding SRRF image shows as characteristic property that all features are presented as uniform lines. This discrepancy between fluctuation techniques is known and a current topic of research. Figure 6 shows the parameters used for MUSICAL reconstructions in Fig. 5. This figure uses the plotting tool included in MusiJ.

Fig. 7 presents a comparison between the previous MATLAB implementation and MusiJ. Note that, by default, MATLAB uses 64 bits as data type, but the final printed result is presented as a PNG image of 8bits. It is hence less than the resolution used internally. Another minor difference in the implementation is the size of the sliding window. In MATLAB, the window size (in pixels) is always matched to exactly an airy disk, while in MusiJ the minimum size is 7 pixels. Hence, whenever the computed size is less than 7, the sliding window adds additional content. Nevertheless, no significant difference can be seen between the implementations.



**Fig. 6.** The singular values for each sliding window in the cardiomyoblast sample shown in Fig. 5. This was obtained using the tool included in MusiJ. The horizontal axis corresponds to the index of the singular value. The dashed lines correspond to the threshold used for every reconstruction.



**Fig. 7.** A reconstructed nanoscopy image as produced by the previous (A) Matlab implementation and (B) MusiJ. A threshold of -1.0 was used in both versions of MUSICAL. The other parameters were the same as used in Fig. 5.

### Funding

H2020 Marie Skłodowska-Curie Actions (836355, SEP-210382872); Norges Forskningsråd (285571); H2020 European Research Council (336716, 804233).



## Acknowledgements

FS acknowledges funding from a Horizon 2020 Marie Skłodowska-Curie Action (836355). KA acknowledges funding from a Horizon 2020 Marie Skłodowska-Curie Action (SEP-210382872) and a Horizon 2020 ERC Starting Grant (804233). SA also acknowledges funding from the Horizon 2020 Marie Skłodowska-Curie Action (SEP-210382872). BSA acknowledges funding from a FP7 ERC Starting Grant (336716) and RCN BioTek2021 (285571). The authors thank Åsa Birgisdóttir for kind provision of cultured cells.

KA conceived the project. SA developed the eigenvalue-based version of MUSCIAL images and implemented the back end of the ImageJ plugin. ISO prepared samples for microscopy, acquired data for MUSICAL and tested the plugin on a broad range of multi-color live-cell data for different reconstruction parameters. ISO provided feedback on necessary functionalities of the plugin and ease of use. FS and ISO designed the front end of the plugin. FS wrote the front end wrapper macro. FS, SA, ISO, and KA wrote the manuscript. BSA provided guidance and research tools. All authors commented on the manuscript.

## Disclosures

The computational enhancements in the back end are a part of the invention disclosure 2019/3595 claimed by UiT The Arctic University of Norway. This part of the invention disclosure is not intended for commercialization and does not conflict with the commercialization of the remaining part. The authors declare that there are no conflicts of interest related to this article.

## References

1. L. Schermelleh, A. Ferrand, T. Huser, C. Eggeling, M. Sauer, O. Biehlmaier, and G. P. Drummen, "Super-resolution microscopy demystified," *Nat. Cell Biol.* **21**(1), 72–84 (2019).
2. S. W. Hell and J. Wichmann, "Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy," *Opt. Lett.* **19**(11), 780–782 (1994).
3. A. G. York, P. Chandris, D. Dalle Nogare, J. Head, P. Wawrzusin, R. S. Fischer, A. Chitnis, and H. Shroff, "Instant super-resolution imaging in live cells and embryos via analog image processing," *Nat. Methods* **10**(11), 1122–1126 (2013).
4. M. G. Gustafsson, "Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy," *J. Microsc.* **198**(2), 82–87 (2000).
5. E. Mudry, K. Belkebir, J. Girard, J. Savatier, E. Le Moal, C. Nicoletti, M. Allain, and A. Sentenac, "Structured illumination microscopy using unknown speckle patterns," *Nat. Photonics* **6**(5), 312–315 (2012).
6. M. J. Rust, M. Bates, and X. Zhuang, "Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM)," *Nat. Methods* **3**(10), 793–796 (2006).
7. T. Dertinger, R. Colyer, G. Iyer, S. Weiss, and J. Enderlein, "Fast, background-free, 3D super-resolution optical fluctuation imaging (SOFI)," *Proc. Natl. Acad. Sci.* **106**(52), 22287–22292 (2009).
8. M. Müller, V. Mönkemöller, S. Hennig, W. Hübner, and T. Huser, "Open-source image reconstruction of super-resolution structured illumination microscopy data in ImageJ," *Nat. Commun.* **7**(1), 10980 (2016).
9. C. A. Schneider, W. S. Rasband, and K. W. Eliceiri, "NIH image to ImageJ: 25 years of image analysis," *Nat. Methods* **9**(7), 671–675 (2012).
10. J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J.-Y. Tinevez, D. J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, and A. Cardona, "Fiji: an open-source platform for biological-image analysis," *Nat. Methods* **9**(7), 676–682 (2012).
11. D. Sage, H. Kirshner, T. Pengo, N. Stuurman, J. Min, S. Manley, and M. Unser, "Quantitative evaluation of software packages for single-molecule localization microscopy," *Nat. Methods* **12**(8), 717–724 (2015).
12. D. Sage, T. A. Pham, H. Babcock, T. Lukes, T. Pengo, J. Chao, R. Velmurugan, A. Herbert, A. Agrawal, S. Colabrese, A. Wheeler, A. Archetti, B. Rieger, R. Ober, G. M. Hagen, J. B. Sibarita, J. Ries, R. Henriques, M. Unser, and S. Holden, "Super-resolution fight club: assessment of 2D and 3D single-molecule localization microscopy software," *Nat. Methods* **16**(5), 387–395 (2019).
13. S. van de Linde, "Single-molecule localization microscopy analysis with ImageJ," *J. Phys. D: Appl. Phys.* **52**(20), 203002 (2019).
14. C. J. Rowlands, F. Ströhl, P. P. V. Ramirez, K. M. Scherer, and C. F. Kaminski, "Flat-field super-resolution localization microscopy with a low-cost refractive beam-shaping element," *Sci. Rep.* **8**(1), 5630 (2018).
15. H. Deschout, F. C. Zanicchi, M. Młodzianoski, A. Diaspro, J. Bewersdorf, S. T. Hess, and K. Braeckmans, "Precisely and accurately localizing single emitters in fluorescence microscopy," *Nat. Methods* **11**(3), 253–266 (2014).

16. C. Franke, M. Sauer, and S. van de Linde, "Photometry unlocks 3D information from 2D localization microscopy data," *Nat. Methods* **14**(1), 41–44 (2017).
17. T. Dertinger, R. Colyer, R. Vogel, M. Heilemann, M. Sauer, J. Enderlein, and S. Weiss, "Superresolution optical fluctuation imaging (SOFI)," in *Nano-Biotechnology for Biomedical and Diagnostic Research*, (Springer, 2012), pp. 17–21.
18. S. Cox, E. Rosten, J. Monypenny, T. Jovanovic-Talisman, D. T. Burnette, J. Lippincott-Schwartz, G. E. Jones, and R. Heintzmann, "Bayesian localization microscopy reveals nanoscale podosome dynamics," *Nat. Methods* **9**(2), 195–200 (2012).
19. I. Yahiatene, S. Hennig, M. Müller, and T. Huser, "Entropy-based super-resolution imaging (ESI): From disorder to fine detail," *ACS Photonics* **2**(8), 1049–1056 (2015).
20. N. Gustafsson, S. Culley, G. Ashdown, D. M. Owen, P. M. Pereira, and R. Henriques, "Fast live-cell conventional fluorophore nanoscopy with ImageJ through super-resolution radial fluctuations," *Nat. Commun.* **7**(1), 12471 (2016).
21. K. Agarwal and R. Macháň, "Multiple signal classification algorithm for super-resolution fluorescence microscopy," *Nat. Commun.* **7**(1), 13752 (2016).
22. R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.* **34**(3), 276–280 (1986).
23. A. M. Elbir, "Deepmusic: Multiple signal classification via deep learning," arXiv pp. arXiv–1912 (2019).
24. K. Agarwal and X. Chen, "Applicability of music-type imaging in two-dimensional electromagnetic inverse problems," *IEEE Trans. Antennas Propag.* **56**(10), 3217–3223 (2008).
25. K. Agarwal, L. Pan, Y. K. Leong, M. Han, O. Y. Chan, X. Chen, and S. P. Yeo, "Practical applications of multiple signal classification," *Int. J. RF Microw. Comput. Eng.* **22**(3), 359–369 (2012).
26. G. T. Dempsey, J. C. Vaughan, K. H. Chen, M. Bates, and X. Zhuang, "Evaluation of fluorophores for optimal performance in localization-based super-resolution imaging," *Nat. Methods* **8**(12), 1027–1036 (2011).
27. D. Zuras, M. Cowlshaw, A. Aiken, M. Applegate, D. Bailey, S. Bass, D. Bhandarkar, M. Bhat, D. Bindel, and S. Boldo, "IEEE standard for floating-point arithmetic," *IEEE Std* **754**, 1–70 (2008).
28. S. Acuña and I. Opstad, "MusIJ: example of results," <https://www.doi.org/10.6084/m9.figshare.12006615.v3> (2020). Accessed: 2020-03-20.

## Paper II

Soft thresholding schemes for multiple signal  
classification algorithm





# Soft thresholding schemes for multiple signal classification algorithm

SEBASTIAN ACUÑA,<sup>1,\*</sup>  IDA S. OPSTAD,<sup>1</sup> FRED GODTLIEBSEN,<sup>1</sup>  
BALPREET SINGH AHLUWALIA,<sup>1,2</sup>  ID AND KRISHNA AGARWAL<sup>1</sup> 

<sup>1</sup>Department of Physics and Technology, UiT The Arctic University of Norway, NO-9037 Tromsø, Norway

<sup>2</sup>Department of Clinical Science, Intervention & Technology, Karolinska Institute, 17177 Stockholm, Sweden

\*sebastian.acuna@uit.no

**Abstract:** Multiple signal classification algorithm (MUSICAL) exploits temporal fluctuations in fluorescence intensity to perform super-resolution microscopy by computing the value of a super-resolving indicator function across a fine sample grid. A key step in the algorithm is the separation of the measurements into signal and noise subspaces, based on a single user-specified parameter called the threshold. The resulting image is strongly sensitive to this parameter and the subjectivity arising from multiple practical factors makes it difficult to determine the right rule of selection. We address this issue by proposing soft thresholding schemes derived from a new generalized framework for indicator function design. We show that the new schemes significantly alleviate the subjectivity and sensitivity of hard thresholding while retaining the super-resolution ability. We also evaluate the trade-off between resolution and contrast and the out-of-focus light rejection using the various indicator functions. Through this, we create significant new insights into the use and further optimization of MUSICAL for a wide range of practical scenarios.

© 2020 Optical Society of America under the terms of the [OSA Open Access Publishing Agreement](#)

## 1. Introduction

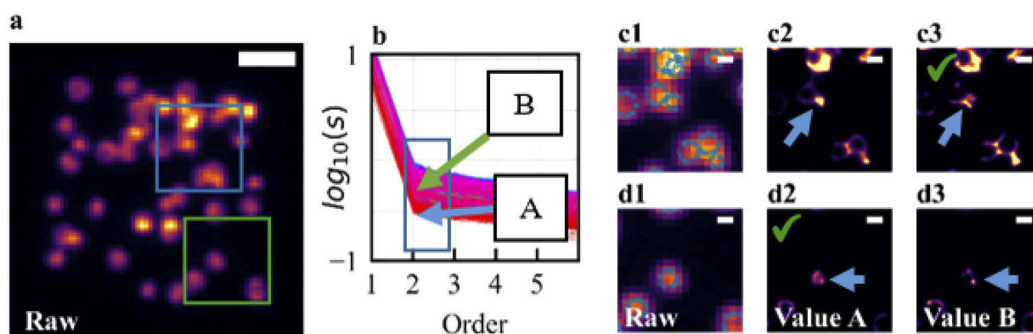
Conventional optical microscopy is limited in resolution due to diffraction of light. The need to overcome this limit has given rise to super-resolution microscopy techniques, also called optical nanoscopy. Among these techniques, structured illumination microscopy (SIM) [1] allows a lateral resolution enhancement by a factor of 2 over the optical resolution limit, stimulated emission depletion (STED) microscopy [2] and single molecule localization (SML) [3–5] can achieve resolutions close to 20 nm, and MINIFLUX [6] which combines concepts of SML and STED to achieve even 2 nm resolution. However, these techniques require expensive and complex setups as in SIM and STED, or a high amount of light dose and long acquisition time in SML.

A new family of purely computational nanoscopy techniques has also emerged, where statistical analysis of spatio-temporal fluctuations of fluorescence intensity arising from the photokinetic properties of fluorescent molecules is used for super-resolution. Examples include super-resolution optical fluctuation imaging (SOFI) [7], super-resolution radial fluctuations (SRRF) [8], multiple signal classification algorithm (MUSICAL) [9], entropy-based super-resolution imaging (ESI) [10], 3B [11], spatial covariance reconstructive (SCORE) super-resolution fluorescence microscopy [12], super-resolution with auto-correlation two-step deconvolution (SACD) [13] and sparsity-based super-resolution correlation microscopy (SPARCOM) [14].

Most of these methods assume that the emitters are stationary and that the photokinetic properties do not change during the image acquisition. MUSICAL, on the contrary, simply exploits the spatio-temporal variations as acquired on a temporal stack of fluorescence microscopy images, irrespective of whether they arise from photokinetics or movement of fluorophores. It decomposes the stack into a set of vectors called eigenimages (details in section 2.2) and separates them into two orthogonal spaces called the signal and the noise subspaces. These are used into a special function called the *indicator function* that exploits the fact that the spatial-temporal

distribution of fluorophores is encoded in eigenimages through the point spread function (PSF) of the microscope. It is designed to have a high value at an emitter location and lower otherwise, which enables super-resolution when applied to a grid finer than the original microscopy image.

MUSICAL uses three user-specified control parameters that have a bearing on the reconstructed nanoscopy image: a parameter  $\alpha$  determines the effective contrast, a sub-pixelation parameter determines the size of the finer grid, and a threshold parameter determines how the eigenimages are split between the signal and noise subspaces. While  $\alpha$  and sub-pixelation prominently determine the aesthetic of the nanoscopy image, the threshold parameter has a significant role in the nature and scale of details that get reconstructed. The problem is that its selection depends on multiple factors such as fluorophore's density and photokinetics, sample's geometry, out of focus light, signal to noise ratio, etc. in a complicated and non-obvious manner. Even though some rules of thumb for threshold selection have been reported [9,15,16], it still remains as the most non-intuitive parameter of MUSICAL as its value is subjective and depends mostly on the user's experience. The problem of threshold selection is illustrated in Fig. 1, where for the same sample (Fig. 1(a)), a single value does not produce the best possible reconstruction for the entire sample. This figure also illustrates the process of threshold selection, which starts with a plot similar to Fig. 1(b) from where the user selects a value such as  $-0.5$  (point A) or  $-0.3$  (point B). The name of the points come from the rules used to pick them (more details in section 2.4) and as it can be seen from the results on region 1 (Fig. 1(c)) and 2 (Fig. 1(d)), the results vary between regions. For example, if we consider the indicated ring in each case, for region 1 the ring is only recovered for value B, while is reconstructed better for value A in region 2 as the ring appears clearer.



**Fig. 1. Effect of MUSICAL's threshold on image reconstruction.** A sample consisting of rings in the focal plane allows to see the effect of two different thresholds in the final reconstruction (a green tick marks the best one). The sample consists of 500 frames. (a) Mean image with 2 regions marked with rectangles. Scale bar  $1\mu\text{m}$ . (b) Plot of the singular values [15] used by the user to pick a threshold. (c),(d) MUSICAL results in regions 1 and 2, respectively. Scale bars 200 nm. (c1),(d1) show the mean images with the actual emitters on top in blue color.

This work addresses the problem of the subjectivity of threshold selection and the sensitivity of MUSICAL's reconstructions to it. This is achieved in the following manner:

- We have scrutinized the effect of threshold selection and identified the root cause of threshold sensitivity.
- We have proposed a generalized form of indicator function design that carry two new families of indicator functions.
- We have evaluated quantitatively and qualitatively the comparative advantages of the new indicator functions while generating new insights into practically useful properties such as resolution-contrast trade-offs, out-of-focus light rejection, and dynamic range utilization.

The outline of the paper is as follows. section 2 presents the theory behind MUSICAL and the role of the threshold in image reconstruction. The new generalized framework for the MUSICAL indicator function design is presented in section 3 section 4 presents the results and insights on a variety of simulated and experimental data. The conclusions are summarized in section 5.

## 2. Background

### 2.1. Imaging model

Let's consider a sample composed of point-like blinking photon emitters, i.e. individual fluorophore molecules. In this case, blinking means that the number of photons emitted by each particle varies over time, owing to the photokinetics of the fluorophores [17], with no assumption on temporal sparsity. A stack of images taken of such sample from a diffraction-limited system over  $T$  time-steps is expressed in a matrix form as  $\mathbf{A} = [\mathbf{a}(1) \dots \mathbf{a}(T)]$ , where each column vector  $\mathbf{a}(t)$  contains the intensity measured by a set of sensing elements (i.e. camera pixels) at time step  $t$ . For  $N$  emitters, a single image  $\mathbf{a}(t)$  can be modelled as follows:

$$\mathbf{a}(t) = \sum_{n=1}^N \mathbf{g}(\mathbf{r}_n(t))s_n(t). \quad (1)$$

In this model,  $\mathbf{g}(\mathbf{r})$  represents the PSF of the microscope and corresponds to the intensity produced by an emitter located at  $\mathbf{r}$  as measured on the sensor. Finally,  $s_n(t)$  is the number of photons produced by the  $n^{\text{th}}$  emitter during the time step  $t$ , which is a random variable resulting from the photokinetics of the fluorophore. For simplicity, we here only consider stationary emitters,  $r_n(t) = r_n$ . Then, Eq. (1) can be written as the matrix equation  $\mathbf{a}(t) = \mathbf{G}\mathbf{s}(t)$  by constructing two new matrices  $\mathbf{G} = [\mathbf{g}(\mathbf{r}_1) \dots \mathbf{g}(\mathbf{r}_N)]$  and  $\mathbf{s}(t) = [s_1(t) \dots s_N(t)]^T$ . Therefore, every single image is a linear combination of the columns of  $\mathbf{G}$ , weighted by the photon emissions of emitters.

### 2.2. Key concept of MUSICAL

MUSICAL involves a sliding window operation, with window defined as a crop of the image stack. Processing a single window returns a super-resolved version of it, and the final MUSICAL reconstruction is built by overlaying and stitching all super-resolved windows together. The size of the window corresponds to the approximate size of the main lobe of the PSF, which is estimated from the wavelength of emission and numerical aperture of the imaging system. While performing nanoscopy using MUSICAL on a single window, the observed data is decomposed into two orthogonal subspaces. In this paper, we use notation associated with singular value decomposition such as used in [9]. Accordingly, the matrix  $\mathbf{A}$  is decomposed as  $\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^T$ , where  $\mathbf{U}$  contains the basis vectors  $\mathbf{u}_i$ . These vectors are called eigenimages and contain the spatial information of the sample, while their corresponding singular values  $\sigma_i$  in the diagonal of matrix  $\mathbf{S}$  is a measure of its statistical significance. Since matrix  $\mathbf{V}$  is not used, MUSICAL can benefit of the relation between SVD and Eigenvalue decomposition as  $\mathbf{A}\mathbf{A}^T = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T$  (as shown in [18] for computational efficiency), where  $\mathbf{\Lambda}$  contains the eigenvalues  $\lambda_i$  instead of the singular values. Since both values are related ( $\lambda_i = \sigma_i^2$ ), the discussion below can be generally applicable irrespective of whether singular value or eigenvalue decomposition is used. The two key concepts of MUSICAL are presented below.

**Key concept 1 (KC1):** MUSICAL separates the basis  $\mathbf{U}$  into two subspaces, namely the signal subspace  $\mathcal{S}$  and the orthogonal complementary noise subspace  $\mathcal{N}$  using a threshold  $\sigma_0$  such that:

$$\mathbf{u}_i \in \begin{cases} \mathcal{S} & \text{if } \sigma_i \geq \sigma_0 \\ \mathcal{N} & \text{otherwise.} \end{cases} \quad (2)$$

**Key concept 2 (KC2):** If the threshold is chosen such that the subspace  $\mathcal{N}$  indeed contains contribution from only the noise, then the subspace  $\mathcal{S}$  contains  $\mathbf{g}(\mathbf{r}_n)$  for every emitter and  $\mathcal{N}$  is



devoid of them. Consequently, using the orthogonality of  $\mathcal{S}$  and  $\mathcal{N}$ ,  $\mathbf{g}(\mathbf{r}_n)$  are also orthogonal to  $\mathcal{N}$ . This property is used to design the indicator function  $f(\mathbf{r})$ , discussed in detail in section 2.3.

In the ideal case, the matrix  $\mathbf{A}$  is rank deficient (i.e., it contains some zero singular values), which happens when the number of emitters is smaller than the number of sensor elements. In such case, the threshold is simply  $\sigma_0 = 0$  and the signal subspace  $\mathcal{S}$  is formed by the eigenimages with non-zero singular values while the noise subspace  $\mathcal{N}$  by the ones with value zero. However, real samples are composed of a large number of emitters such that the matrix  $\mathbf{A}$  is full ranked. In addition, real microscopy data contains noise, such that none of the singular values are strictly zero. Therefore,  $\sigma_0$  is chosen on a case-by-case basis. Intuitively, the eigenvectors associated with large eigenvalues represent the more statistically prominent structures in the stack as compared to the other eigenvectors.

**The key concept 2:** Since the space spanned by  $\mathbf{g}(\mathbf{r}_i)$  and the eigenimages of the signal space are the same, the noise space is also orthogonal to  $\mathbf{g}(\mathbf{r}_i)$ , i.e. when the PSF is evaluated at an emitter's location. This property is exploited in designing indicator function that indicates the presence of emitters.

### 2.3. Indicator function

Let us consider an arbitrary test point  $\mathbf{r}_{test}$  in the sample region. Its image is given by the vector  $\mathbf{g}(\mathbf{r})$ , which can be represented as  $\mathbf{g}(\mathbf{r}) = \mathbf{g}_S(\mathbf{r}) + \mathbf{g}_N(\mathbf{r})$ , where  $\mathbf{g}_S(\mathbf{r})$  and  $\mathbf{g}_N(\mathbf{r})$  are the projections of  $\mathbf{g}(\mathbf{r})$  onto the signal and noise spaces. The magnitudes of these projections are computed as:

$$\|\mathbf{g}_S(\mathbf{r}_{test})\| = \sqrt{\sum_{\mathbf{u}_i \in \mathcal{S}} g_i^2} \quad ; \quad \|\mathbf{g}_N(\mathbf{r}_{test})\| = \sqrt{\sum_{\mathbf{u}_i \in \mathcal{N}} g_i^2} \quad \text{where} \quad g_i = |\mathbf{g}(\mathbf{r}_{test}) \cdot \mathbf{u}_i|. \quad (3)$$

Using KC2 and the projections in Eq. (3), MUSICAL constructs the following indicator function:

$$f(\mathbf{r}_{test}) = \left( \frac{\|\mathbf{g}_S(\mathbf{r}_{test})\|}{\|\mathbf{g}_N(\mathbf{r}_{test})\|} \right)^\alpha. \quad (4)$$

This indicator function generates a large value when  $\mathbf{r}_{test}$  is at the location of the emitters, since  $\|\mathbf{g}_N(\mathbf{r}_{test})\|$  is zero (KC2). For this it is important that the threshold value  $\sigma_0$  for defining the signal and noise subspaces is chosen appropriately.

### 2.4. Threshold selection and its associated problem

The current practice of threshold selection starts with a plot of the singular values of the microscopy image stack, such as shown in Fig. 1(b). This plot is made by computing the singular values for each window and then plotting them as lines in logarithmic scale, with the x-axis showing the order when sorted decreasingly. Then, the user specifies a threshold  $\sigma_0$  based on observations derived from this plot with particular interest in the inflection or knee point (indicated with a blue rectangle) observed at the second singular value of every window. The rule of thumb reported in the original MUSICAL article [9], referred to as Rule A here, involves choosing roughly the threshold as the smallest second singular value across all the windows, i.e.

$$\text{Rule A: } \sigma_0 = \min(\sigma_2; \forall \text{windows}). \quad (5)$$

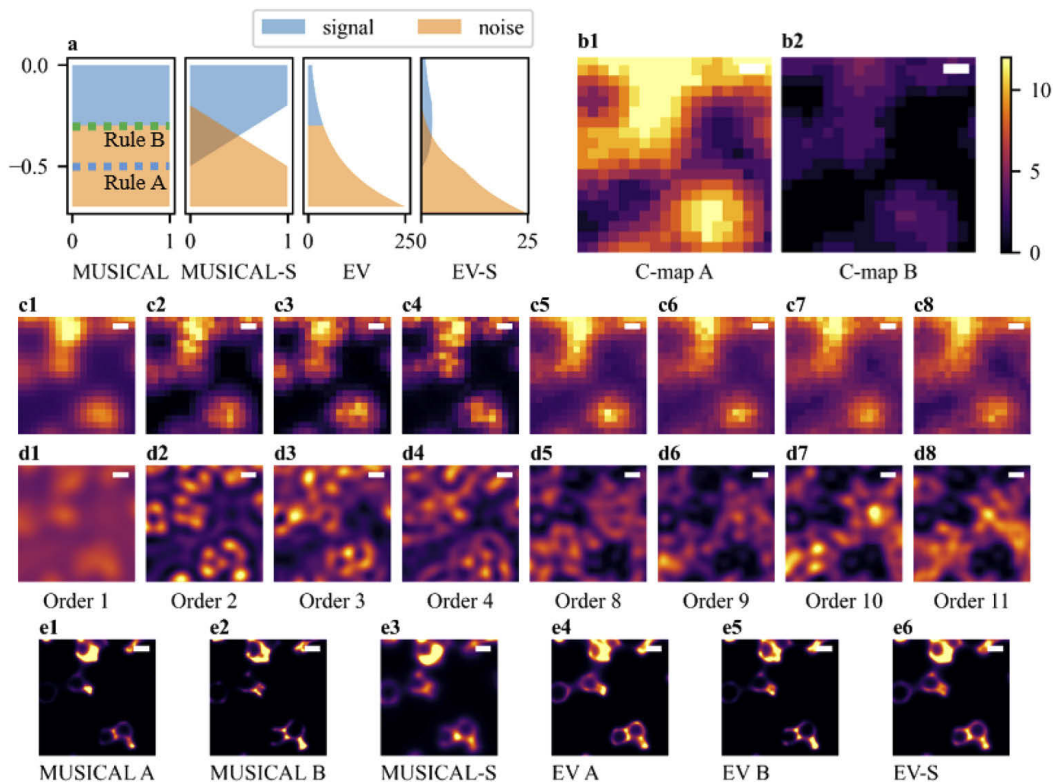
According to another rule reported in [16], referred here as Rule B, the threshold is roughly selected as the center of the span of the second singular values for all the windows, i.e.

$$\text{Rule B: } \sigma_0 = \frac{\min(\sigma_2; \forall \text{windows}) + \max(\sigma_2; \forall \text{windows})}{2}. \quad (6)$$

Both rules assume that the eigenimages can be clearly separated and therefore, they constitute hard thresholding schemes. An illustrative example is shown in Fig. 2(a) where the classification



and corresponding weighing is shown for Rule B in comparison to Rule A. In more practical terms, a higher value of threshold, decreases the cardinality of the signal subspace across the entire images as shown in the cardinality maps (c-map) of Fig. 2(b). These maps encode the number of eigenimages in the signal subspace for each window as an intensity value, which is reason why the the c-map B (corresponding to Rule B) looks dimmer. For even further insight, we show in Fig. 2(c) and (d), the actual singular values and eigenimages projection respectively for several orders. Something interesting to note, is that foreground regions generally have higher eigenvalues than the background for any given order. However, the projections of  $g_i$  shown Fig. 2(d) follow two different trends. First, the value of  $g_i$  observe an inversion in the pattern as the order increases. The lower orders have high values of  $g_i$  in the foreground relative to the background, and vice versa for the higher orders. This reversal of trend for higher orders (low singular values) gets exploited in the indicator function as their contribution to the denominator is small. Second, generally speaking, eigenimages with lower eigenvalues (i.e. higher orders) are associated with higher spatial frequency components [19], an effect that is more prominent for the first few orders. These observations will be further discussed later in relation with the new indicator functions.



**Fig. 2. MUSICAL's singular values and eigenimages.** (a) Weighting coefficients for the different methods. The rules are labeled on top to show how the weights change depending on the threshold. In this example, rule A means a threshold of  $-0.5$  while rule B means  $-0.3$ . For soft methods, the noise and signal are used both in numerator and denominator of the indicator function. (b) Cardinality maps that indicate the number of eigenimages picked as signal for the two rules of selection for every windows (each pixel corresponds to one window). (c), (d) Heatmap of singular values (c) and projection of different eigenimages (d) sorted by increasing order. (e) Results of using different indicator functions. The scale bars are 200 nm.

Since Rule B decreases the cardinality of the signal space, it also discards noise much more effectively. However, information related to the actual structure may also be relegated as noise, which increases the value of denominator in the foreground and compromises resolution. On the other hand, and following the example of Fig. 2, Rule A includes eigenimages up to order 11 in the foreground. Yet, as noted in Fig. 2(d), the  $g_i$  of the foreground is smaller than the background in orders 10 and 11. Therefore, their exclusion from the noise subspace is not useful for the foreground. Even more: the inclusion of higher  $g_i$  corresponding to the background due to these orders may increase the background artifacts. Therefore, a trade-off is involved in either rule and the manifestation of this trade-off varies from case-to-case.

In practice, even the most experienced bioimaging user of MUSICAL may not know what to expect from the sample being imaged. Moreover, one could argue that more candidate rules can be designed from analysis of the histogram of the second singular values, such as explored in [18]. However, in all these cases, a fundamental limitation is that they all imply hard thresholding: eigenimages included in the signal subspace are considered in a hard solely as representing the structure and the eigenimages included in the noise subspace are considered in solely as representing the noise. In reality, the presence of noise implies that each eigenimage is corrupted. Therefore, a perfect separation of the eigenimages into a signal or noise subspace is not possible.

### 3. New indicator function design

Here, we consider two solutions to the problem mentioned previously:

- Eigenvalue (EV) weighing: the magnitude of the eigenvalue is included in the indicator function. It keeps the hard thresholding but softens the effect of first few eigenimages that are classified as noise but may have structural information, as shown in Fig. 2(a).
- Soft-thresholding: hard thresholding is removed using a weighing function for each eigenimage. It can be added to MUSICAL and EV, with the new methods abbreviated MUSICAL-S and EV-S, respectively.

We begin with a generalized form of indicator function, which allows for more flexibility in its design and that paves the path for developing the indicator functions for the above identified solutions:

$$f(\mathbf{r}_{test}) = \left( \sqrt{\frac{\sum_{i=1}^N a_i g_i^2}{\sum_{i=1}^N b_i g_i^2}} \right)^\alpha. \quad (7)$$

Here,  $a_i$  and  $b_i$  are the design parameters of the indicator function. This generalization can be readily adapted to the original MUSICAL indicator function given by Eq. (4) by using the following assignments for  $a_i$  and  $b_i$ , with  $a_i + b_i = 1$ :

$$a_i = \begin{cases} 1 & \text{if } \mathbf{u}_i \in \mathcal{S} \\ 0 & \text{otherwise} \end{cases} \quad b_i = \begin{cases} 1 & \text{if } \mathbf{u}_i \in \mathcal{N} \\ 0 & \text{otherwise} \end{cases}. \quad (8)$$

#### 3.1. Indicator function with eigenvalue (EV) weighing

This indicator function follows  $a_i + b_i = \lambda^{-1}$  and is defined as:

$$a_i = \begin{cases} \lambda_i^{-1} & \text{if } \mathbf{u}_i \in \mathcal{S} \\ 0 & \text{otherwise} \end{cases} \quad b_i = \begin{cases} \lambda_i^{-1} & \text{if } \mathbf{u}_i \in \mathcal{N} \\ 0 & \text{otherwise} \end{cases}. \quad (9)$$

This indicator function design retains both KC1 and KC2, but does not use the Euclidean projections  $\|\mathbf{g}_{\mathcal{S}}(\mathbf{r}_{test})\|$  and  $\|\mathbf{g}_{\mathcal{N}}(\mathbf{r}_{test})\|$  on the signal and noise subspaces. It instead weighs the

projections on individual eigenimages  $g_i$  according to the inverse of its singular value. This is graphically illustrated in Fig. 2(a) for rule B.

We present further insight into the EV indicator function using Fig. 2(c-d). Due to multiplication with the inverse of eigenvalues,  $g_i$  for any order gets amplified for background regions and attenuated for the foreground regions. The foreground attenuation helps the higher order eigenimages to better support the resolution and the lower orders in reducing the dynamic range of the nanoscopy image on the higher side. The background amplification helps the higher order eigenimages to better suppress the background artifacts and the lower orders in reducing the dynamic range on the lower side. Thereon, the effect of hard thresholding is still present. Nonetheless, a significant softening is achieved as described next. Consider the orders that which are assigned to the signal subspace using Rule A but to the noise subspace using Rule B, however when treated using EV indicator function. When included in the signal subspace, they reduce the dynamic range of the original version of MUSICAL. This is significant because when using rule A, the original indicator function of MUSICAL generally supports better resolution but has extremely high indicator function values for few foreground pixels in nanoscopy image. This results into some pixels being highly saturated in the MUSICAL images and dynamic range of the image is not well-utilized, as reported in the supplement of the original MUSICAL paper [9]. On the other hand, when included in the noise subspace they help in improving the resolution, which may have been compromised in rule B as discussed before in section 2.4.

The proposed indicator function is inspired from the EV formulation reported previously for inverse source problems, for example in [20,21]. The similarity between these previously reported EV formulations and the one proposed here is limited to the denominator component of Eq. (7) when combined with  $b_i$  defined in Eq. (9). The use of the signal subspace in the numerator and the application of the indicator function on one sliding window at a time are unique to the MUSICAL architecture, first reported in [9] while incorporation of EV weighing in  $a_i$  in Eq. (9) is proposed for the first time here.

### 3.2. Indicator function with soft threshold (MUSICAL-S)

An alternative approach is to use continuous functions for  $a_i$  and  $b_i$ . Our proposed function is defined in Eq. (10) and graphically illustrated in Fig. 2(a).

$$a_i(x) = \begin{cases} 1 & \text{if } \sigma_i \geq \sigma_{\max} \\ 0 & \text{if } \sigma_i \leq \sigma_{\min} \\ \frac{\log_{10} \sigma_i - \log_{10} \sigma_{\min}}{\log_{10} \sigma_{\max} - \log_{10} \sigma_{\min}} & \text{otherwise} \end{cases} ; \quad b_i(x) = 1 - a_i(x). \quad (10)$$

We enforce that  $a_i + b_i = 1$ . In this equation  $\sigma_{\max}$  and  $\sigma_{\min}$  must be given. We pick those to be the maximum and minimum of the second singular values across all the windows in the image, since the change in the trend is strongly evident in the second singular values. We note that the choice also obviates the need of user specification of  $\sigma_{\min}$  and  $\sigma_{\max}$ . Thus, this approach includes both soft and automatic thresholding properties.

The design of indicator function as suggested above implies that the signal and noise subspaces are no longer orthogonal. There are some eigenimages ( $\mathbf{u}_i; \sigma_i \geq \sigma_{\max}$ ) that are relegated to the signal subspace  $\mathcal{S}$  with full confidence ( $a_i = 1, b_i = 0$ ). Similarly, there are some eigenimages ( $\mathbf{u}_i; \sigma_i \leq \sigma_{\min}$ ) that are relegated to the noise subspace  $\mathcal{N}$  with full confidence ( $a_i = 0, b_i = 1$ ). For the remaining eigenimages, it is understood that these eigenimages may have non-negligible signal information even while being significantly affected by noise. Therefore, these eigenimages are relegated to signal and noise space with reduced confidence (indicated by non-extreme values of  $a_i$  and  $b_i$ ) while the behaviour and roles of the  $g_i$  for lowest and the highest orders are unambiguous. The role of some intermediate orders such as order 8 is not explicit. When included in the signal space, it does not contribute resolution in foreground but may help in

pushing the lower limit of the dynamic range up by enhancing the background. On the other hand, when included in the noise space, they may degrade the resolution but also pull the upper limit of the dynamic range down. By including them in reduced proportions in both denominators and numerators, we expect to strike a balance between the positive and negative aspects of their inclusion in the signal and noise subspaces and thereby achieve a softening effect of the threshold.

We note that this design is a significant deviation from the key concepts of MUSICAL. Since the signal and noise spaces now share some eigenimages, the KC1 defined in section 2.2 does not apply. Furthermore, the KC2 has to be redefined as follows:

**Redefined KC2:** If the signal and noise subspaces are suitably defined, then the  $b_i$  weighted projection of  $\mathbf{g}(\mathbf{r}_n)\forall n$  on the noise subspace  $\mathcal{N}$  is small, which allows the denominator in Eq. (7) to be small and the overall indicator function to be high at the emitter locations.

### 3.3. Indicator function for EV with soft-threshold (EV-S)

The concept of soft-thresholding can be integrated in EV as well, as shown below:

$$a_i(x) = \begin{cases} \lambda_i^{-1} & \text{if } \sigma_i \geq \sigma_{max} \\ 0 & \text{if } \sigma_i \leq \sigma_{min} \\ \lambda_i^{-1} \left( \frac{\log_{10} \sigma_i - \log_{10} \sigma_{min}}{\log_{10} \sigma_{max} - \log_{10} \sigma_{min}} \right) & \text{otherwise} \end{cases} ; \quad b_i(x) = \lambda_i^{-1} - a_i(x). \quad (11)$$

As before,  $a_i + b_i = \lambda_i^{-1}$ . Its illustration in Fig. 2(a) shows that EV-S is the softest indicator function, allowing smoother transition in  $a_i$  and  $b_i$ . While EV alleviates the sensitivity to the threshold, MUSICAL-S removes the need for user-specified threshold and reduces the sensitivity to noise in the overlapping region in Fig. 2(a). Since EV-S combines traits from both, it is expected to demonstrate reduced sensitivity and soft thresholding.

### 3.4. Discussion of the proposed generalized framework

Through the generalized indicator function, we have allowed for the creation of families of MUSICAL algorithms based on some design rules. Specifically, two families have been created.

**Family based on coefficient relationship:** These are defined based on the relationship between  $a_i$  and  $b_i$ . The family  $a_i + b_i = 1$  may be considered the original MUSICAL family while the family  $a_i + b_i = \lambda_i^{-1}$  as the EV family. Other families may also be designed similarly.

**Family of coefficient continuity:** These are defined based on the individual characteristics of  $a_i$  and  $b_i$ . For example, in the family of hard threshold, the curve corresponding to  $a_i$  experiences an abrupt jump at the threshold  $\sigma_0$ , and the signal and noise spaces are disjoint. In the family of soft threshold, the signal and noise subspaces are no longer orthogonal (KC1 does not apply) but the eigenimages in the overlapping space of signal and noise subspaces are weighted log-linearly. Other approaches may be designed for choosing the overlapping region or designing the weights, leading to other families of algorithms.

Two final notes on the newly defined indicator functions:

**No resolution enhancement expected:** The aim of these new indicator functions is not resolution enhancement. In fact, no indicator function is expected to enhance or deteriorate resolution in particular. The only expected effect is a minor trade-off in resolution and contrast arising from different treatments of eigenimages in the signal and noise subspaces, as seen in Fig. 2(e).

**Removing subjectivity through automatic thresholding:** An important implication of the soft automatic indicator functions is that the subjectivity in threshold selection as well as the dependence on heuristics is removed. Outlook for further customization by an advanced user is possible, for example through a different choice of  $\sigma_{min}$  and  $\sigma_{max}$ , or  $a_i$  and  $b_i$ .

## 4. Results

We performed the following studies to compare the performances of the newly proposed indicator functions with the original indicator functions of MUSICAL:

- **Quantitative analysis:** We compare resolution and contrast for the indicator functions. We consider the effect of intensity fluctuations (determined by the photon emission on/off time) and the signal to background ratio (SBR). 2D samples are used so that other effects such as out-of-focus light do not affect the quantitative results.
- **Qualitative analysis:** Structures in biological samples are simulated to study how the different indicator functions deal with realistic 3D structures and out-of-focus light.
- **Results on experimental data:** We show comparative results using experimental data of a diverse set of samples. The dynamic range coverage is also investigated.

We consider a total of six indicator functions referred to as MUSICAL A, MUSICAL B, MUSICAL-S, EV A, EV B and EV-S. The value of  $\alpha = 4$  is used as recommended in [9] and 10 subpixels per pixel are considered sufficient for the investigations presented in this article.

**Simulation methods for studies 1-3:** The simulation involves the following steps: simulating the geometry of the structures, the emitters distribution over the structures, and the photokinetics of each emitter using the photoemission model of [22], applying the Gibson-Lanni PSF [23,24] of the microscope to simulate the raw noise-free image stacks, and then simulating the noise in raw image stack using the noise model in [9]. We have simulated images for an epifluorescence microscope of numerical aperture 1.42, pixel size 80 nm, and 665 nm emission wavelength. Thereby, the theoretical resolution limit is  $\sim 285$  nm using Rayleigh criterion [25]. For each sample, 500 frames were generated with 10 ms of exposure time per frame. Duty cycle is the ratio of the average time a fluorophore emits light ( $\tau_{\text{on}}$ ) to the total cycle ( $\tau_{\text{on}} + \tau_{\text{off}}$ ). Photobleaching is not considered. Where not mentioned explicitly, SBR of 4 and duty cycle of 5% are utilized.

### 4.1. Quantitative analysis

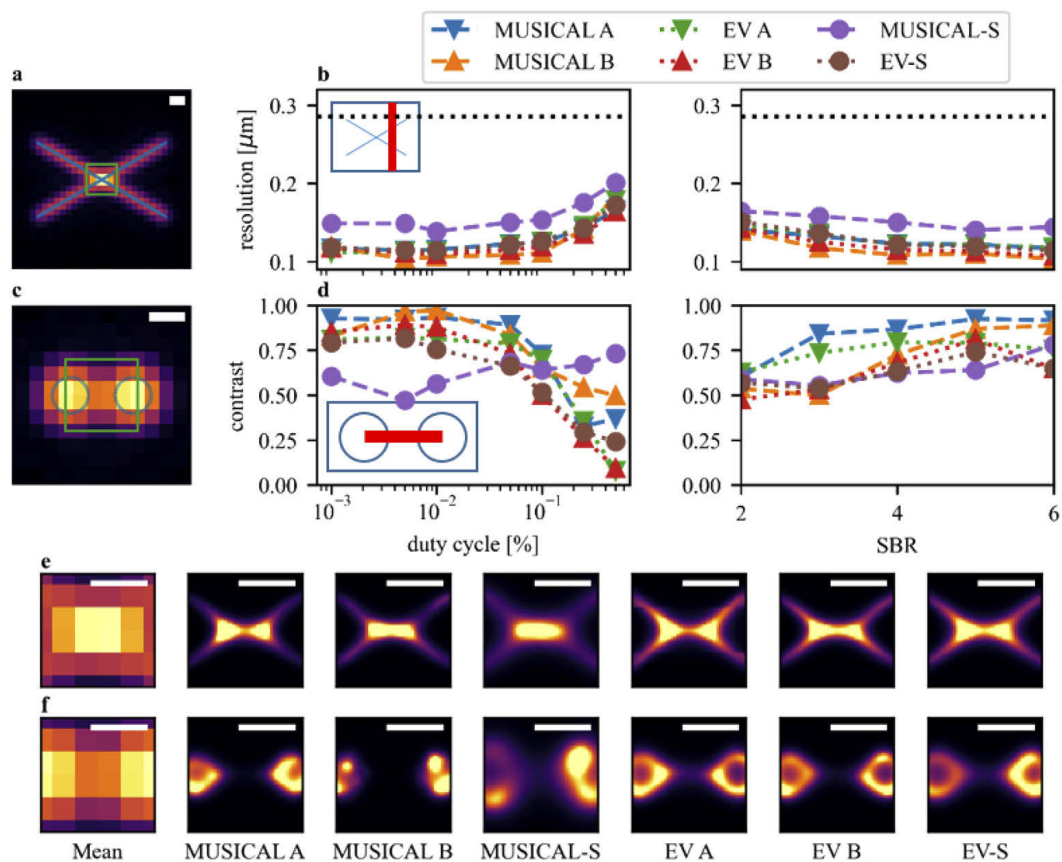
**Study of resolution:** This study is performed on a sample comprising two crossing lines with an angle of  $60^\circ$  as shown in Fig. 3(a), where the emitters are uniformly distributed at a density of 500 per  $\mu\text{m}$ . The results of Fig. 3(b) show the distance from the crossing point at which the 2 lines can be differentiated according to the Rayleigh criterion [25] for different values of the duty cycle and SBR. The diffraction limited resolution is shown as a blue line for reference.

The resolution is estimated by computing the ratio between valley and peaks across a moving vertical section (shown in the inset of Fig. 3(b)), starting from the intersection point ( $x = 0$ ). Let  $\mathbf{I}(x)$  denote the image's intensities across a vertical line passing by  $x$ . We compute the ratio:

$$r(x) = \frac{v(\mathbf{I}(x))}{\min(p_1(\mathbf{I}(x)), p_2(\mathbf{I}(x)))}. \quad (12)$$

Here,  $v$  and  $p_i$  are functions that return the minimum and maximum intensity value respectively value around the expected valley and peaks  $i$  (bottom and top peaks). The reported resolution is given as the minimum value  $x$  at which  $r(x) \leq 0.835$ . The range considered for duty cycle goes from 0.1 % (comparable to single molecule localization microscopy data) up to 50%. This last case corresponds to a highly dense spatio-temporal emission situation, which is a challenging situation for most fluctuations based techniques [26]. Further, we consider SBR from 2 to 6, where SBR 2 is considered quite poor. From Fig. 3, the first observation is that all the methods provide resolution enhancement by a factor between 2 and 3. When comparing rules, we find that the best results are obtained using rule B for both MUSICAL and EV. Considering hard thresholding and comparing the MUSICAL and EV families, MUSICAL takes the lead in general.





**Fig. 3. Resolution and contrast study.** (a), (c) Studied samples with the groundtruth in blue and green square shown magnified in panels (e), (f) for qualitative illustration. (b), (d) Resolution and contrast, respectively, for different methods at different duty cycles and SBR. Each data point in Fig. (b), (d) corresponds to the mean value over 100 independent simulations. Dashed line in (b) shows the resolution limit. Scale bar is 200 nm.

Among soft-threshold methods, EV-S performs comparable to MUSICAL and EV while but MUSICAL-S displays the worst performance among all of them. The poorer resolution of MUSICAL-S is also observed in Fig. 3(c) where the borders of the central cross appear diffused compared to the other methods, producing a lower resolution. Lastly, we note that duty cycles below 10% and SBR higher than 4 do not provide significant improvement in resolution.

**Study of contrast:** We here consider a sample containing emitters distributed uniformly across the perimeter of two circles of diameter 200 nm (density of 500 per  $\mu\text{m}$ ) at a distance of 150 nm between edges. The mean image and groundtruth are shown in Fig. 3(c). The contrast is defined as  $c = 1 - r(\mathbf{l}(0))$ , where  $r(x)$  is defined in Eq. (12) and  $\mathbf{l}(0)$  represents the intensities across the horizontal section between circles (see inset of Fig. 3(d)). Higher values of  $c$  indicate better contrast. The results in Fig. 3(d) indicate that the contrast for all the indicator functions deteriorates as the duty cycle increases, with a larger slope after 10%. The only exception is MUSICAL-S which shows low sensitivity to duty cycle and has relatively flat contrast for all duty cycles. In terms of SBR, similar to the line sample, the curves become flat for  $\text{SBR} \geq 4$ . When comparing rules, rule A achieves marginally better contrast both in MUSICAL and EV, with MUSICAL performing the best. The soft-threshold methods showed all similar and poorer results. An additional study on the effect of different noise distributions is included in Supplement 1.

**Summary of the quantitative study:** The soft-thresholding schemes provide resolution and contrast similar to the other hard thresholding schemes. In general, EV-S outperformed MUSICAL-S in terms of resolution and contrast due to the weighting scheme by including singular values. Additionally, from the results we observe that above a SBR of 4 and a duty cycle of 10%, no improvement in resolution or contrast were achieved.

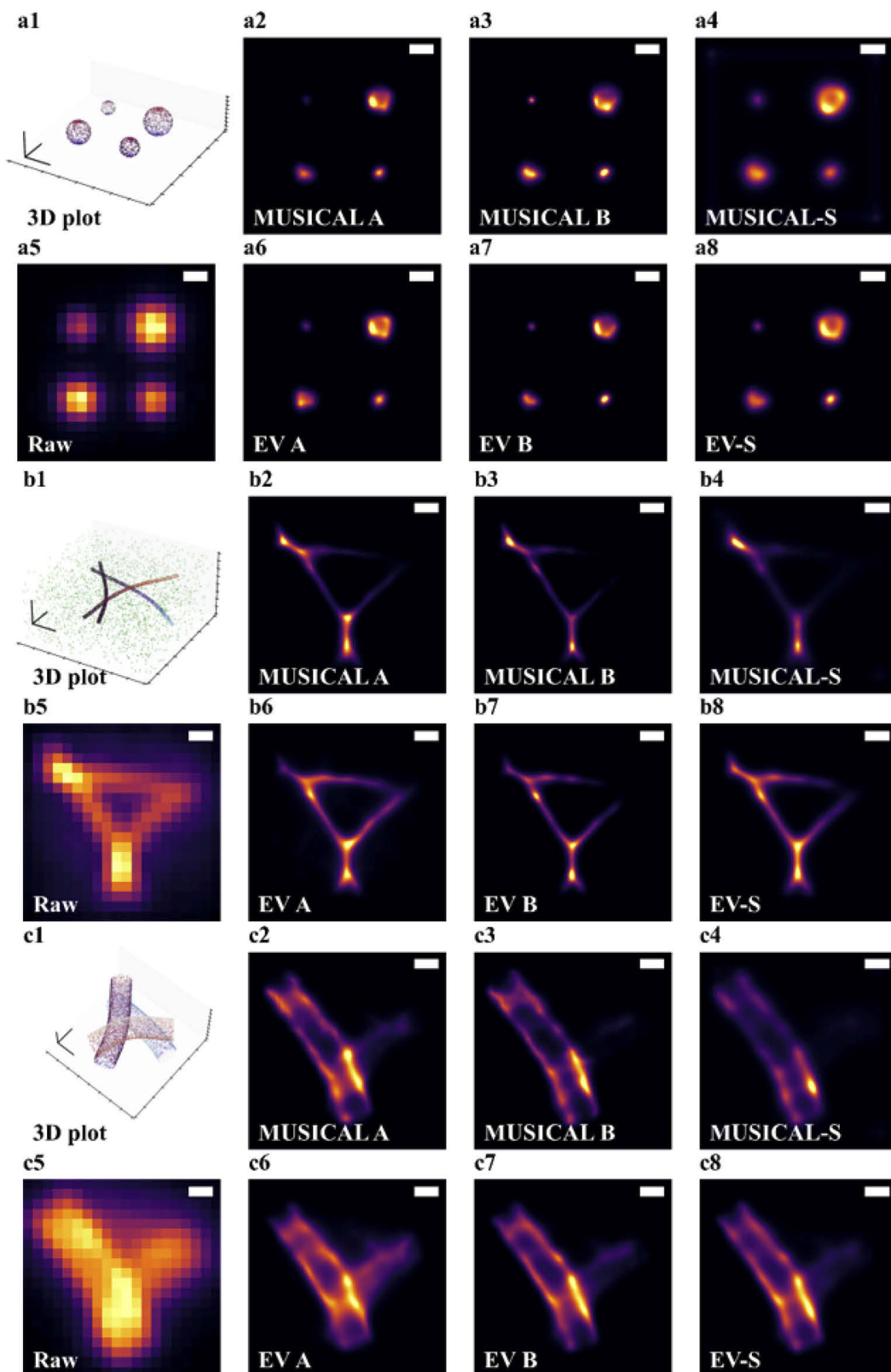
#### 4.2. Qualitative examples

**Vesicles with surface labeling (Fig. 4(a1-a8)):** Four vesicles of different sizes are simulated as spheres of diameters 150, 200, 250 and 300 nm. Membrane labeling is simulated by distributing emitters on surface with density 800 per  $\mu\text{m}^2$ . The vesicles are placed such that their centers are at the focal plane. Larger vesicles result with certain portions out of focus. In MUSICAL A (Fig. 4(a3)), the smallest and therefore dimmer structure (top-left vesicle) is almost invisible. This is not the case for MUSICAL B (Fig. 4(a4)) which is explained by a lower cardinality of the signal space and therefore better discarding of noise. In the case of MUSICAL-S (Fig. 4(a5)) the structure is even more visible, and the remaining vesicles look more uniform. No clear difference is observed across EV reconstructions (Fig. 4(a6, a7) and (a8)). It is of interest to observe that for the largest vesicle (300 nm in diameter), the out of focus light is rejected by all the methods, producing a hollow in the middle that is not visible for the microscopy image. Below that size, the entire vesicle can be considered to be in focus. We observe that even if the size of the circle reconstructed is proportional to the original radius, this does not match the correct size when using full width at half maximum (FWHM).

**Microtubules with background debris (Fig. 4(b1-b8)):** Microtubules are fiber-like polymers of tubulin monomers. Fluorescent dyes label the monomers, which may be present as freely dispersing in addition to microtubule fiber [27]. This results in fluorescent debris, which is generally unwelcome in reconstruction. For this sample, we simulated fibers of 30 nm in diameter, with dyes distributed randomly across their surface at a linear density of 800 emitters per  $\mu\text{m}$ . Additionally, debris is added as single emitters distributed randomly with a volumetric density of 1000 emitters per  $\mu\text{m}^3$ . Both, microtubule-bound and free emitters, are assumed to have the same photo-kinetics. The geometry consists of three crossing lines forming an inverted triangle when seen from the top (Fig. 4(b2)). In the top-left and bottom corner, the structures meet in the focal plane, while the microtubules meeting at the top-right corner are both out of focus and separated by 500 nm in axial direction. The spatial distribution can be seen in Fig. 4(b1).

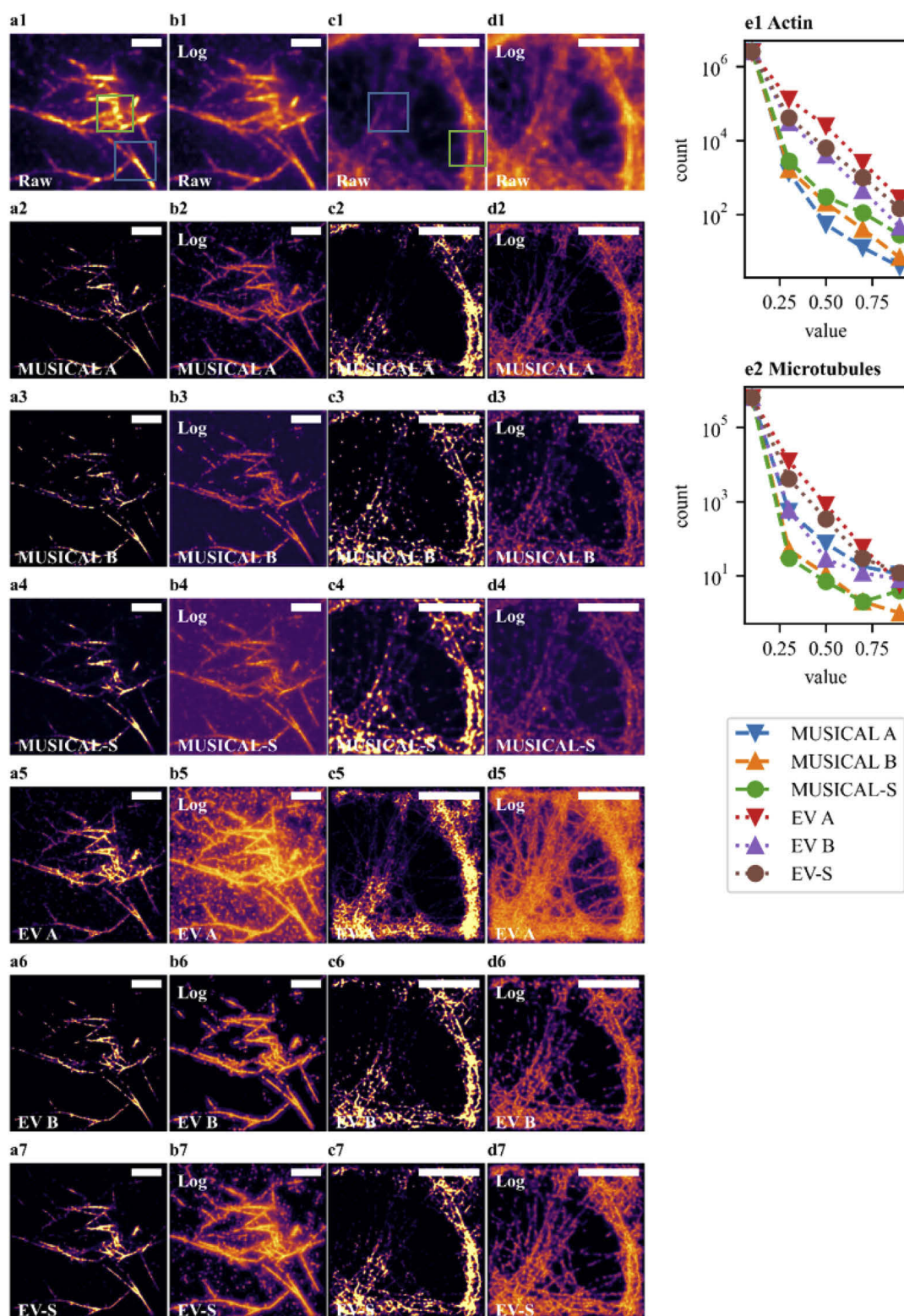
The results show that debris is absent in reconstructions regardless of the method while the overall structure is well reconstructed. A difference in the reconstruction of regions with out-of-focus structures is noticeable: MUSICAL A, B and S (Fig. 4(b3, b4) and (b5) respectively) achieve a better rejection than their EV counterparts (Fig. 4(b6, b7) and (b8)). However, the corners of the triangle which are the brightest points do not allow a clearer reconstruction of the strands due to the dynamic range problem of MUSICAL reported in [9] and discussed in section 3. In this sense, a better reconstruction is obtained by EV. It shows that a desired level of rejection of out-of-focus structures could determine the choice between MUSICAL and EV.

**Mitochondrion (Fig. 4(c1-c8)):** Mitochondria are tubular structures with diameters close to or larger than the diffraction limit. Even if such structures are in focus, the sometimes large diameters combined with the dynamic nature of these organelles in living samples, causes large portions of the samples to be out of focus in realistic microscopy experiments. Here, we consider an example of three mitochondria with diameter 300 nm and a density of emitters of 3000 per  $\mu\text{m}^2$  on the outer membrane. Each mitochondrion is in a different plane, with the left most mitochondrion in the focal plane. The top one is in the plane 300 nm above (closer to the coverslip) and the last one is 300 nm below the focal plane (further from the coverslip). As in the case of microtubules, the out of focus rejection is prominent for all methods with MUSICAL obtaining the strongest rejection of out of focus signal. Only the portions in the focal plane are



**Fig. 4. Qualitative study on synthetic samples.** Three different samples are shown. (a) 4 vesicles of different size. (b) Three crossing microtubules with debris. (c) The crossing mitochondria. All scale bars are 300 nm.





**Fig. 5. Results on experimental data.** The results are displayed in two scales: linear (a) and (c) and logarithmic with base 10 (b) and (d). Blue and green boxes show regions of interest. The scale bar is  $2\ \mu\text{m}$  in every image. (a), (b). Results for actin (invitro). (c), (d) Results on microtubules. (e) Comparative plot of the intensities obtained from the histogram of each normalized image by its maximum.

reconstructed with good contrast. Further, the structure away from the coverslip is rejected less effectively than the structure above, which can be explained by the asymmetry of the PSF.

**Summary of this study:** The MUSICAL indicator functions were found to perform stronger out-of-focus rejection than the EV indicator functions.

### 4.3. Results on experimental data

The methods were here tested on real microscopy data of filamentous actin and microtubules in fixed cells. The samples used for testing the methods correspond to invitro actin filaments [9] and microtubules in fixed cells [28]. 500 frames were used for each reconstruction.

**Invitro actin filaments (Fig. 5(a,b)):** This sample is thin in the sense that that all the structures can be considered in focus. We marked two regions of interest (blue and green boxes) where super-resolution can give better insight of the structure than conventional microscopy. In the blue box (close to bottom), the bifurcation is clearly visible using the different indicator functions, and almost no difference is observed in terms of structure. However, the region in the green box is reconstructed better in EV as they show a better distribution of the intensities. This allows to observe the entire network of strands without saturating the colors in other regions and is mainly attributed to the contrast enhancement due to the softening effect of EV. On the other hand in the images in logarithmic scale (Fig. 5(b)), we observe how EV introduces artifacts in the background, which is reduced when using a larger threshold such as the one used with rule B. In the case of the soft thresholding methods, EV-S looks crisper and more defined than MUSICAL-S, while providing better contrast between foreground and background. The pixel distribution is plotted in log scale in Fig. 5(a8), where it can be clearly observed how EV produces an intensity distribution that is higher in the middle tones, making better use of the dynamic range.

**Microtubules in fixed cells (Fig. 5(c,d)):** This sample of fixed cell is different from the actin sample in the sense that some structures are expected to be out-of-focus. For images in linear scale (Fig. 5(c1-c7)), the methods display similar performance, except for EV A. In particular, the blue region (left box) presents a high degree of artifacts, where it is difficult to visualize individual strands due to high saturation. The same occurs in the green region, where EV A recovers just one single structure, while all the remaining methods manage to recover 2 strands. This illustrates an example of the minor trade-off between resolution and contrast. The better contrast and visibility, as well as poorer rejection of out-of-focus structures is strongly evident in the log scale (Fig. 5(d)). Between MUSICAL-S and EV-S, the latter produces a better result by achieving better contrast and definition. Both methods recover the dim structures in the right-most box.

## 5. Conclusion

Through a generalized framework for MUSICAL indicator function design, we have proposed new indicator function families and specific indicator function designs to address problem of hard threshold in MUSICAL. The EV family of indicator functions soften the effect of threshold and provide better utilization of the dynamic range of MUSICAL. This is generally achieved at no cost of resolution but poorer rejection of out-of-focus light as compared to the MUSICAL family of indicator functions. Further, a soft threshold family is proposed that does not define signal and noise subspaces of MUSICAL as strictly orthogonal, and allows an overlap between them. Therefore, it removes the concept of hard thresholding, however, altering the key concepts of MUSICAL. While MUSICAL-S indicate poorer resolution, the EV-S indicator function which combines the traits of both EV and soft-threshold families shows consistently good results across a wide variety of quantitative, qualitative, and experimental studies. Through this work, we widen the horizon for MUSICAL in two important aspects. First, the sensitivity and subjective choice of threshold is removed which makes it easier to use. Second, it opens exciting avenues for further development of fluctuations based super-resolution algorithms in general, and MUSICAL

in particular. In the future, it will be interesting to design customized indicator functions for challenging scenarios such as dynamic samples where the different methods can be tested.

## Funding

H2020 Marie Skłodowska-Curie Actions (749666); Norges Forskningsråd (285571); Universitetet i Tromsø (Strategic funding for PhD); European Research Council (804233).

## Disclosures

The authors declare that there are no conflicts of interest related to this article.

See [Supplement 1](#) for supporting content.

## References

1. M. G. Gustafsson, "Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy," *J. Microsc.* **198**(2), 82–87 (2000).
2. S. W. Hell and J. Wichmann, "Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy," *Opt. Lett.* **19**(11), 780–782 (1994).
3. E. Betzig, G. H. Patterson, R. Sougrat, O. W. Lindwasser, S. Olenych, J. S. Bonifacino, M. W. Davidson, J. Lippincott-Schwartz, and H. F. Hess, "Imaging intracellular fluorescent proteins at nanometer resolution," *Science* **313**(5793), 1642–1645 (2006).
4. M. J. Rust, M. Bates, and X. Zhuang, "Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM)," *Nat. Methods* **3**(10), 793–796 (2006).
5. J. Schnitzbauer, M. T. Strauss, T. Schlichthaerle, F. Schueder, and R. Jungmann, "Super-resolution microscopy with DNA-PAINT," *Nat. Protoc.* **12**(6), 1198–1228 (2017).
6. K. C. Gwosch, J. K. Pape, F. Balzarotti, P. Hoess, J. Ellenberg, J. Ries, and S. W. Hell, "MINIFLUX nanoscopy delivers 3d multicolor nanometer resolution in cells," *Nat. Methods* **17**, 217–224 (2020).
7. T. Dertinger, R. Colyer, G. Iyer, S. Weiss, and J. Enderlein, "Fast, background-free, 3d super-resolution optical fluctuation imaging (SOFI)," *Proc. Natl. Acad. Sci.* **106**(52), 22287–22292 (2009).
8. N. Gustafsson, S. Culley, G. Ashdown, D. M. Owen, P. M. Pereira, and R. Henriques, "Fast live-cell conventional fluorophore nanoscopy with imagej through super-resolution radial fluctuations," *Nat. Commun.* **7**(1), 12471 (2016).
9. K. Agarwal and R. Machán, "Multiple signal classification algorithm for super-resolution fluorescence microscopy," *Nat. Commun.* **7**(1), 13752 (2016).
10. I. Yahiatene, S. Hennig, M. Müller, and T. Huser, "Entropy-based super-resolution imaging (ESI): From disorder to fine detail," *ACS Photonics* **2**(8), 1049–1056 (2015).
11. Y. S. Hu, X. Nan, P. Sengupta, J. Lippincott-Schwartz, and H. Cang, "Accelerating 3B single-molecule super-resolution microscopy with cloud computing," *Nat. Methods* **10**(2), 96–97 (2013).
12. Y. Deng, M. Sun, P.-H. Lin, J. Ma, and J. W. Shaevitz, "Spatial covariance reconstructive (SCORE) super-resolution fluorescence microscopy," *PLoS One* **9**(4), e94807 (2014).
13. W. Zhao, J. W. Liu, C. Kong, Y. Zhao, C. Guo, C. Liu, X. Ding, X. Ding, J. Tan, and H. Li, "Faster super-resolution imaging with auto-correlation two-step deconvolution," arXiv preprint arXiv:1809.07410 (2018).
14. O. Solomon, M. Mutzafi, M. Segev, and Y. C. Eldar, "Sparsity-based super-resolution microscopy from correlation information," *Opt. Express* **26**(14), 18238–18269 (2018).
15. S. Acuña, F. Ströhl, I. S. Opstad, B. S. Ahluwalia, and K. Agarwal, "MusiJ: an ImageJ plugin for video nanoscopy," *Biomed. Opt. Express* **11**(5), 2548–2559 (2020).
16. I. S. Opstad, F. Ströhl, Å. B. Birgisdottir, S. Maldonado, T. Kalstad, T. Myrnel, K. Agarwal, and B. S. Ahluwalia, "Adaptive fluctuation imaging captures rapid subcellular dynamics," in *Advances in Microscopic Imaging II*, vol. 11076 (SPIE - International Society for Optics and Photonics, USA, 2019), p. 110761W.
17. G. T. Dempsey, J. C. Vaughan, K. H. Chen, M. Bates, and X. Zhuang, "Evaluation of fluorophores for optimal performance in localization-based super-resolution imaging," *Nat. Methods* **8**(12), 1027–1036 (2011).
18. S. A. Acuña-Maldonado, "Multiple Signal Classification Algorithm: computational time reduction and pattern recognition applications," Master's thesis, UiT The Arctic University of Norway, Tromsø, Norway (2019).
19. K. Agarwal and D. K. Prasad, "Eigen-analysis reveals components supporting super-resolution imaging of blinking fluorophores," *Sci. Rep.* **7**(1), 4445 (2017).
20. V. Madisetti, *The Digital Signal Processing Handbook* (CRC, USA, 2009), 2nd ed.
21. D. H. Johnson and D. E. Dudgeon, *Array Signal Processing: Concepts and Techniques* (Simon & Schuster, USA, 1992).

22. A. Girsault, T. Lukes, A. Sharipov, S. Geissbuehler, M. Leutenegger, W. Vandenberg, P. Dedecker, J. Hofkens, and T. Lasser, "SOFI simulation tool: A software package for simulating and testing super-resolution optical fluctuation imaging," *PLoS One* **11**(9), e0161602 (2016).
23. S. F. Gibson and F. Lanni, "Experimental test of an analytical model of aberration in an oil-immersion objective lens used in three-dimensional light microscopy," *J. Opt. Soc. Am. A* **9**(1), 154–166 (1992).
24. J. Li, F. Xue, and T. Blu, "Fast and accurate three-dimensional point spread function computation for fluorescence microscopy," *J. Opt. Soc. Am. A* **34**(6), 1029–1034 (2017).
25. L. Novotny and B. Hecht, *Principles of Nano-Optics* (Cambridge University, 2012), 2nd ed.
26. I. S. Opstad, S. Acuña, L. E. V. Hernandez, J. Cauzzo, N. Škalko Basnet, B. S. Ahluwalia, and K. Agarwal, "Fluorescence fluctuations-based super-resolution microscopy techniques: an experimental comparative study," arXiv preprint arXiv:2008.09195 (2020).
27. E. Nogales, "Structural insights into microtubule function," *Annu. Rev. Biophys. Biomol. Struct.* **30**(1), 397–420 (2001).
28. K. Agarwal, R. Macháň, and D. K. Prasad, "Non-heuristic automatic techniques for overcoming low signal-to-noise-ratio bias of localization microscopy and multiple signal classification algorithm," *Sci. Rep.* **8**(1), 4988 (2018).

## Soft thresholding schemes for multiple signal classification algorithm: supplement

SEBASTIAN ACUÑA,<sup>1,\*</sup>  IDA S. OPSTAD,<sup>1</sup> FRED GODTLIEBSEN,<sup>1</sup>  
BALPREET SINGH AHLUWALIA,<sup>1,2</sup>  AND KRISHNA AGARWAL<sup>1</sup> 

<sup>1</sup>*Department of Physics and Technology, UiT The Arctic University of Norway, NO-9037 Tromsø, Norway*

<sup>2</sup>*Department of Clinical Science, Intervention & Technology, Karolinska Institute, 17177 Stockholm, Sweden*

\**sebastian.acuna@uit.no*

---

This supplement published with The Optical Society on 28 October 2020 by The Authors under the terms of the [Creative Commons Attribution 4.0 License](https://creativecommons.org/licenses/by/4.0/) in the format provided by the authors and unedited. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

Supplement DOI: <https://doi.org/10.6084/m9.figshare.13076807>

Parent Article DOI: <https://doi.org/10.1364/OE.409363>



## Soft thresholding schemes for multiple signal classification algorithm - Supplementary document

SEBASTIAN ACUÑA<sup>1,\*</sup>, IDA S. OPSTAD<sup>1</sup>, FRED GODTLIEBSEN<sup>1</sup>,  
BALPREET SINGH AHLUWALIA<sup>1,2</sup> AND KRISHNA AGARWAL<sup>1</sup>

<sup>1</sup>Department of Physics and Technology, UiT The Arctic University of Norway, NO-9037 Tromsø, Norway

<sup>2</sup>Department of Clinical Science, Intervention & Technology, Karolinska Institute, 17177 Stockholm, Sweden

\*sebastian.acuna@uit.no

### Note S1. A study of the effect of noise sources on the results of different indicator functions

We first explain the noise model simulated for the result in the main paper. This noise model includes the effect of the following noise sources [1]:

- Shot noise (or photon noise): The photon arrival at the detector array follows a Poisson process. Assuming uniform quantum efficiency over all the pixels, the noise in creation of photo-electron can be modeled as Poisson noise arising from photon arrival statistics. This noise model is applied on top of an image formed just by the mean photon count multiplied with the quantum efficiency, and represented as pixel intensity values.
- Dark current and clock induced charge: These sources of noise are associated with detector electronics and can also be modelled with a Poisson its emission rate.

We implement these noise models together as a single Poisson distribution as described by Eq. (10) of [1]. Accordingly, we normalize the simulated image by its maximum so that the values are in the range [0,1]. This result is referred to as  $\tilde{\mathbf{a}}$  and is the input of the noise model. The intensity at a pixel  $\mathbf{r}$  is represented as  $\tilde{\mathbf{a}}_{\mathbf{r}}$ . The output is controlled by two parameters referred to as  $s$  and  $b$ , such that  $b$  represents the background noise sources and  $(s - b)$  represents the mean photon count for the maximum signal coming from the photons. The stochastic model is then defined in terms of probability density function  $p$ :

$$p(\mathbf{a}_{\mathbf{r}}|\tilde{\mathbf{a}}_{\mathbf{r}}) = \text{Poisson}(\mathbf{a}_{\mathbf{r}}|(s - b)\tilde{\mathbf{a}}_{\mathbf{r}} + b), \quad (\text{S1})$$

where  $\mathbf{a}_{\mathbf{r}}$  is the intensity in the noisy output image  $\mathbf{a}$ . For the simulations presented in this work,  $b$  was set to 100, while  $s$  was varied in order to obtain different SBR values, defined as  $s/b$ . The value of  $b$  was taken in accordance with the measured mean intensity when the camera shutter was closed and images were taken in this configuration, as reported in [2]. Now, we consider three different cases of noise model variations.

The first case considers the extreme case where  $s$  is 400 and  $b$  is 10 which exemplifies a very high quality sensor with low background noise. The results are shown in S1a, where it is observed that EV produces stitching artifacts, especially for rule A. This results from having very low singular values for the windows that represent background, which in turn render the numerator and the denominator numerically large and unstable. The resulting indicator function is also unstable as a consequence.

The second case considers the joint effect of photon noise and background, varying both  $b$  and  $s$ . Considering  $b$  equal to 100 as the reference for background [2], we used three values for  $b$ : 10, 50, and 100, while keeping  $s/b$  equal to 4. As a result, we have 3 three different situations. In the first case,  $b$  equals 10 indicates a very high quality sensor while  $s$  equals 40 indicates

either very poor quantum efficiency or very low photon count. The second scenario is the one in which the camera's background noise is realistically small but the photon count is also not too prominent, neither too small. This is a scenario in which realistically good sensors provide advantage. The third scenario is the one simulated in the main manuscript, and representative of the most common practical scenario regarding the camera and the photon count. It is seen that very low photon counts indeed affects all the indicator functions adversely. However, a good balance between the background and the photon count is observed in the last two scenarios, where the results are not significantly different in qualitative sense for any given indicator function.

The third case considers a non-Poisson distribution of noise, namely additive Gaussian noise. Therefore, this noise model is modeled as follows, instead of Eq. (S1):

$$p(\mathbf{a}_r|\tilde{\mathbf{a}}_r) = (s - b)\tilde{\mathbf{a}}_r + \text{Gaussian}(\epsilon|b, \text{sd}), \quad (\text{S2})$$

where  $\epsilon$  is the noise value, sd is the standard deviation of the Gaussian distribution and  $b$  is the mean. We have deliberately avoided the use of  $\sigma$  as a symbol of standard deviation to avoid confusion regarding the notation of singular values in the main paper. The results are presented in Figure S1c for three different values of standard deviations, while  $s$  is equal to 400 and  $b$  to 100 are used for all of them.

In the case of the background pixels, using a Gaussian distribution with sd equals 10 is comparable to using a Poisson distribution as both have the same standard deviation and the mean is large enough. This is based on the fact that Poisson distribution with mean  $\mu$  can be approximated by a Gaussian distribution with mean  $\mu$  and standard deviation  $\sqrt{\mu}$ , provided  $\mu$  is larger than 22. Therefore, it can be compared with the third row of Figure S1b in this regard. The additive Gaussian noise case indeed generates sharper reconstructions for all the indicator functions as compared to the corresponding Poisson noise case.

Further, it is noted that the reconstructions for all the indicator functions deteriorate fast as the standard deviation in Gaussian noise model increases.

## References

1. M. Hirsch, R. J. Wareham, M. L. Martin-Fernandez, M. P. Hobson, and D. J. Rolfe, "A stochastic model for electron multiplication charge-coupled devices—from theory to practice," *PloS one* **8**, e53671 (2013).
2. K. Agarwal and R. Macháň, "Multiple signal classification algorithm for super-resolution fluorescence microscopy," *Nat. Commun.* **7**, 13752 (2016).

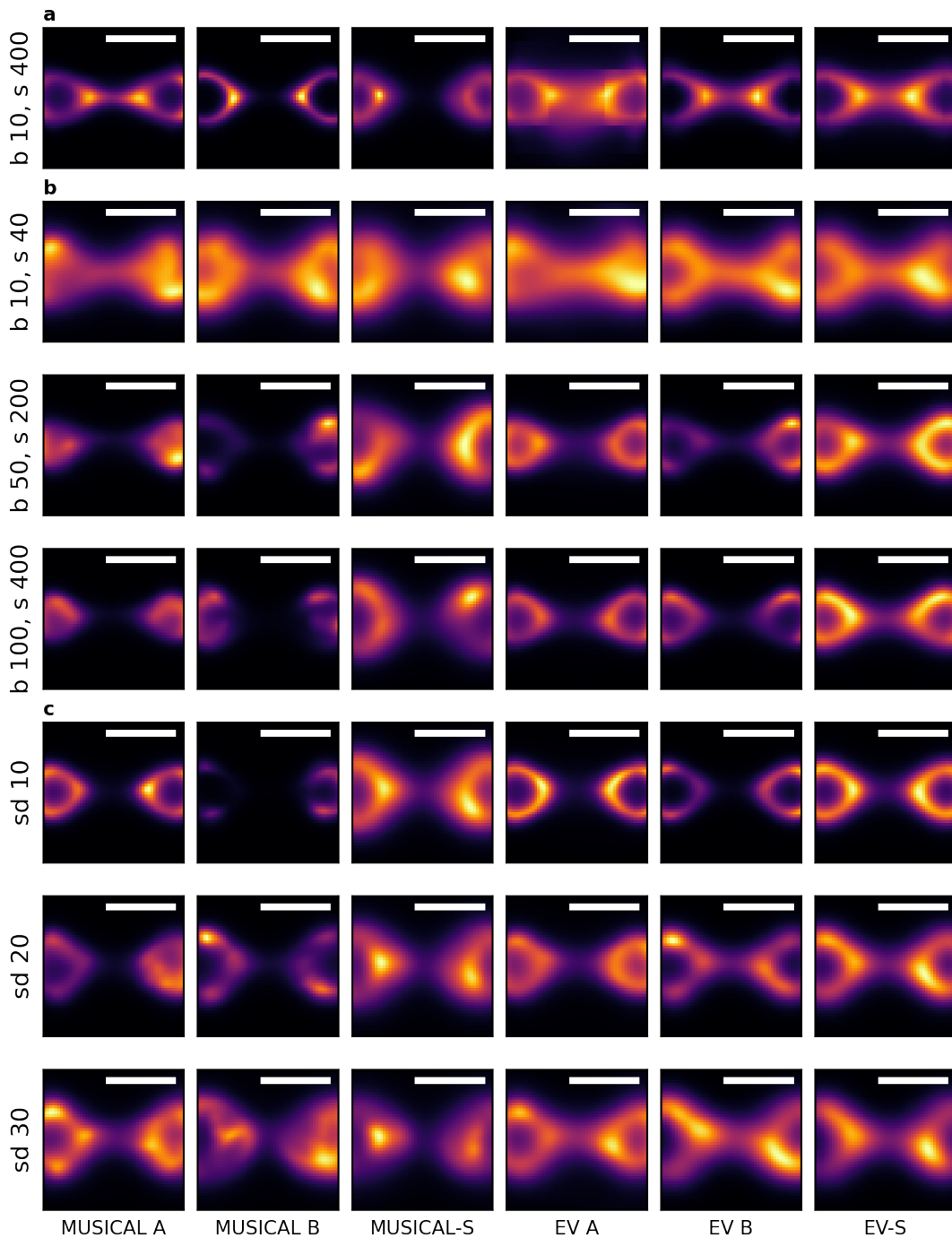


Fig. S1. **Results for different noise models and noise conditions.** Scale bars 200 nm. **a.** Example of extreme low background value. **b.** Variation in the mean background value  $b$  while keeping the  $s/b = 4$ . The values for background ( $b$ ) and signal ( $s$ ) are indicated to the left of each row. **c.** Additive Gaussian noise model is used instead of Poisson noise model. The standard deviation (sd) is given to the left of each row. Mean of Gaussian is  $b = 100$ , and  $s/b = 4$  is used for all examples.



## Paper III

Deriving high contrast fluorescence microscopy images  
through low contrast noisy image stacks





# Deriving high contrast fluorescence microscopy images through low contrast noisy image stacks

SEBASTIAN ACUÑA,<sup>1,3,\*</sup>  MAYANK ROY,<sup>2,3</sup> LUIS E. VILLEGAS-HERNÁNDEZ,<sup>1</sup> VISHESH K. DUBEY,<sup>1</sup>  BALPREET SINGH AHLUWALIA,<sup>1</sup>  AND KRISHNA AGARWAL<sup>1</sup>

<sup>1</sup>Department of Physics and Technology, UiT The Arctic University of Norway, 9010 Tromsø, Norway

<sup>2</sup>Indian Institute of Technology (Indian School of Mines), Dhanbad 826004, India

<sup>3</sup>Shared co-authors

\*sebastian.acuna@uit.no

**Abstract:** Contrast in fluorescence microscopy images allows for the differentiation between different structures by their difference in intensities. However, factors such as point-spread function and noise may reduce it, affecting its interpretability. We identified that fluctuation of emitters in a stack of images can be exploited to achieve increased contrast when compared to the average and Richardson-Lucy deconvolution. We tested our methods on four increasingly challenging samples including tissue, in which case results were comparable to the ones obtained by structured illumination microscopy in terms of contrast.

© 2021 Optical Society of America under the terms of the [OSA Open Access Publishing Agreement](#)

## 1. Introduction

Microscopy allows biologists to study life by observing at a magnified scale the objects of interest such as a cell and its inner components. In particular, fluorescent optical microscopy allows the visualization of structures of interest such as microtubules or mitochondria through special molecules called fluorophores that are attached to the sample during its preparation. However, the ability to resolve between different intensities in the image, the so-called contrast, may be affected by several factors such as blurriness due to point-spread-function of the system (PSF), Poisson noise due to photon counting, electronic noise produced by spurious charges and dark noise. While it is possible to reduce noise during acquisition by increasing the excitation laser power and/or increasing the dwell time, this can produce saturation in the sensor and photobleaching on the sample and therefore it is not always possible nor desirable. The alternative is stacking which consist on taking several images of the sample at low power and low dwell time, and then average them. However, small structures can be affected as their signal becomes closer to the noise level.

Blurriness is present in the image due to diffraction of light [1] which can also be seen as the PSF being applied to the sample by the optical system through a convolution operation. As the PSF depends also on the axial position of emitters, its effect is worse for structures in out-of-focus regions. As a result, off-focus light may introduce background and therefore worsen contrast. In this context, Koho et al. have mentioned that "a good fluorescence image should have good contrast, bright details and dark background" [2], and good efforts have been made in order to improve contrast. Different optical microscopy methods use different optical configurations to enhance the contrast by reducing the out-of-focus noise. This is obtained by performing optical sectioning such that the light arriving to the camera is predominately from the imaging plane. In the case of total internal reflection fluorescence microscopy, optical sectioning is performed by evanescent field that exponentially decays away from the surface, thereby limiting most of the illumination intensity within a few hundred nanometers. Thus, only the part of the sample within the evanescent field gets illuminated, thereby enhancing the contrast. Confocal microscopy achieves optical sectioning by focusing the illumination to a spot such that only the portion within

the spot gets illumination, thereby avoiding the possibility of illumination out-of-focus structures. Similar to confocal microscopy, light sheet microscopy also uses illumination focusing systems such that only a thin sheet within the sample is illuminated, and other out-of-focus regions remain dark. Multi-photon microscopy uses non-linear excitation of the fluorescence signal, that is generated only from the high-intensity focal region. This provides default optical sectioning as the region above and below the focal spot does not emit two-photon fluorescence signal.

All these approaches use specific optical instrumentation and are therefore not versatile, although they present an advantage of no post-processing specifically for the purpose of removing out-of-focus light. Some approaches, such as HiLo [3], combine optical configuration redesign and computational approach. It might also be argued that super-resolution techniques such as structured illumination microscopy [4] also provide indirect contrast enhancement. However, deconvolution is the most versatile and widely used computation-only contrast enhancement approach. An example of such algorithm is the Richardson-Lucy deconvolution [5,6]. Mathematical instability, convergence, and sensitivity to noise present challenges to deconvolution based contrast enhancement. Nonetheless, the popularity of deconvolution approaches clearly indicate that computational-only approaches which do not require specific instrument design and have better stability and noise-robustness are quite valuable.

On the other hand, noise reduction of fluorescence images can be done through a different set of techniques: filtering high spatial frequencies in Fourier space, apply algorithms specially designed for camera noise such as automatic correction of sCMOS-related noise (ACsN) [7], or even deep learning techniques such as noise2void (n2v) [8]. Stacking and averaging is also a good solution for denoising fluorescence image,

Stacking has also being used in several techniques for improving the spatial resolution in fluorescence microscopy without modifying the instrument. In particular, a whole family of methods is based on exploiting the statistics of the fluctuation in photon emission coming from individual emitters [9]. Therefore, inspired by the fact that image stacks have been used for denoising on one hand and super-resolution on the other, we consider if they can also play a role in contrast enhancement. Contrast enhancement is quite a vaguely used term, unfortunately, and is difficult to define or quantify in an absolute manner. While the general hallmark of contrast enhancement is visual quality improvement, it entails a composite of multiple effects such as reduction of out-of-focus light, suppression of background from non-structural randomly distributed fluorescence debris and camera noise, denoising of structures and improving their sharpness, as well as visual enhancement in separability of different structures. For the same reason of diversity of causes, it is difficult to design one technique that addresses all the root causes. Nonetheless, methods that computationally contribute optical sectioning, noise-suppression, and potentially super-resolution or feature sharpening are of interest. Methods in the family described in [9] are likely to provide at two of these benefits at a time. Among the various super-resolution methods, MUSICAL [10] is an algorithm based on fluctuation in which there is a component in addition to super-resolution (as will be clarified later) that includes local context of intensities. It has been shown in [11] that MUSICAL provides super-resolution, optical sectioning and suppression of background due to debris.

Therefore, we consider if MUSICAL can be modified to achieve contrast enhancement instead of super-resolution. In particular, we identify that a recent work in generalizing indicator function of MUSICAL [11], the core function that operates in the local region, allows us in investigating its adaptation for contrast enhancement. As a result, we provide here a novel technique based on MUSICAL and stacking as an alternative to averaging whose purpose is to provide better local contrast by exploiting the fluctuation acquired in the dataset. We compared the results on a set of samples with different degrees of complexity: In-vitro F-Actin filaments, microtubules in U2OS cells, placenta tissue and heart tissue with the contemporary approaches of averaging (mean) and Richardson-Lucy deconvolution (RL), as well as state-of-the-art ACsN.

The structure of the paper is as follows: Section 2 presents the theory behind MUSICAL with focus on the indicator function, and then the new proposed functions. The results are shown in Section 3 which summarises the results and comments of using the function on different experimental samples, in addition to a comparison with state-of-the-art denoising methods. The last part includes a discussion of the results in Section 4 and finally our conclusions and future work are presented in Section 5.

## 2. Methods

MUSICAL is an image processing technique based on the decomposition of a sequence of images that exhibit the intensity fluctuation of the independent fluorophores in fluorescence microscopy into a sequence of orthogonal vectors, also called *eigenimages*, that explain the observed dataset. This can be considered similar to the Fourier Analysis where signals can be represented as a sum of weighted exponentials, with the difference that the eigenimages represent the most prominent spatial structures in the temporal stack. Every eigenimage  $\mathbf{u}_i$  has an *eigenvalue*  $\sigma_i$  associated that can be seen as a relevance factor in explaining the observed data with higher values representing the more relevant ones. This value also allows to order the set of eigenimages and we refer to this as the order of the eigenimage. In this definition, higher order means higher index but lower eigenvalue as they are sorted decreasingly. By doing this, we can expand on the relation with the Fourier domain as an interesting feature of eigenimages is that high order ones can be linked to high spatial frequencies [11,12].

By using the eigenvalue value the eigenimages are classified into two subsets that then form two mutually orthogonal subspaces: the signal subspace  $\mathcal{S}$  and the noise subspace  $\mathcal{N}$ . The assumption is that eigenimages with eigenvalues above certain *threshold* correspond to actual signal, while the ones with values below correspond to noise (therefore the name of the subspaces).

Regarding threshold selection, which is in principle chosen manually by the user, two rules have been suggested [11], both based on the second eigenvalue distribution across patches. From these rule, the simplest one is to use the lowest second eigenvalue (the so-called rule A) which corresponds to the first guideline published with MUSICAL and stated as the knee criterion. Unless the contrary is indicated, this is rule we followed during this work.

The workflow of MUSICAL is based on processing small overlapping sections or three dimensional patches (the third dimension being the time index) of the image of the size of the main lobe of the PSF of the system. This is estimated using four parameters given by the user: the emission wavelength of the fluorophore, the numerical aperture, magnification of the microscope and the pixel size of the microscope camera. Each 3D patch is processed to obtain a two dimensional image, that is then aggregated to the results of all the patches in the set. In the conventional MUSICAL used for super-resolution, the super-resolution is achieved during this patch-processing step by the use of an *indicator function*  $f$  whose value depends on whether an emitter is present or not at a particular test location. MUSICAL's original publication presented a single function, which later was generalized by Acuña et. al. [11] to:

$$f(\mathbf{r}_{test}) = \left( \sqrt{\frac{\sum_{i=1}^N a_i g_i^2}{\sum_{i=1}^N b_i g_i^2}} \right)^\alpha \quad (1)$$

The function  $f$  is evaluated at an arbitrary test point  $\mathbf{r}_{test}$  and its definition is made in such a way that it produces high values when tested at emitters locations, and low otherwise. When the aim of MUSICAL is super-resolution, the test points are chosen as belonging to a grid finer than the original image grid by a factor of a parameter called *sub-pixelation* in the original MUSICAL paper. However, since the aim of this paper is contrast enhancement only, the  $\mathbf{r}_{test}$  in our case correspond to the image grid itself i.e the center of each of the pixels. For each point, the estimated PSF  $\mathbf{g}$  (considered as a vector) is generated and then projected into the eigenimage

space of the patch being processed. Therefore, in this space, the PSF can be represented by its projections elements as  $\mathbf{g} = [g_1, \dots, g_N]$  where  $g_i$  represents the projection of the PSF in the eigenimage  $u_i$ . The variables  $a_i$  and  $b_i$  are part of the generalized framework [11] as they can be defined flexibly allowing the definition of different indicator functions, from where the original function is just a particular case of the general function. The last parameter,  $\alpha$ , while it plays a role in contrast, it can be considered as fix as no further studies have been made after MUSICAL publication and there its recommended value was 4.

We provide MUSICAL's coefficients definition in the context of the generalized indicator function next.

**MUSICAL:** in its original form,  $a_i$  and  $b_i$  are binary and complementary (i.e.  $a_i + b_i = 1$ ) and their value depends on whether the eigenimage  $u_i$  is classified as signal ( $a_i = 1$  and  $b_i = 0$ ) or noise ( $a_i = 0$  and  $b_i = 1$ ). Therefore, the coefficients can be defined as follows:

$$a_i = \begin{cases} 1 & \text{if } \mathbf{u}_i \in \mathcal{S} \\ 0 & \text{otherwise} \end{cases} \quad b_i = \begin{cases} 1 & \text{if } \mathbf{u}_i \in \mathcal{N} \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

In this definition, the projections of the PSF in the signal subspace are allocated in the numerator while the projections in the noise subspace are placed in the denominator. This allows to exploit the fact that every frame in the dataset is the result of a linear combination of the actual emitter's PSF and therefore belongs to the span of the collection of emitters' PSF. In MUSICAL, this space is also estimated using the eigenimages in  $\mathcal{S}$ . Hence, as  $\mathcal{N}$  is by definition orthogonal to  $\mathcal{S}$  it is also orthogonal to the PSFs. As a result, for actual emitters, the components in the denominator will be low as the PSF should be orthogonal to the noise subspace and therefore, the indicator function is high. It is importance to remark that in MUSICAL design, it is the denominator the component that enables super-resolution. The numerator purpose is to provide a local context and it is used exclusively for stitching between nearby patches.

Based on the generalized version of MUSICAL, we designed two indicator functions whose purpose is to enhance the contrast of fluorescence image. Their main difference with MUSICAL is that they do not rely on the noise subspace but only the signal subspace, which reflects their main purpose of contrast enhancement and no super-resolution. The new methods are defined next.

**Contrast enhancement 1 (CE1):** while in MUSICAL's indicator function the numerator contains only the projections in the signal subspace and the denominator only the projections in the noise subspace, in this new design shown in Eq. (3) only the signal part is used as the denominator contains all the projections regardless how they have been classified.

$$a_i = \begin{cases} 1 & \text{if } \mathbf{u}_i \in \mathcal{S} \\ 0 & \text{otherwise} \end{cases} \quad b_i = 1 \quad (3)$$

As a result, we have removed high spatial frequencies associated to eigenimages with low eigenvalues as they are filtered out from the numerator which is now the main component of the indicator function. On the other hand, it provides a normalization component that is not present in MUSICAL: eigenimages with high eigenvalues are weighted the same in the numerator regardless of their relevance and also without being affected by their projections on the noise subspace. This can help to enhance edges or dim structures with good fluctuation levels.

**Contrast enhancement 2 (CE2):** instead of weighting every eigenimage equally as in CE1, we assign a weight based on their corresponding eigenvalue  $\sigma_i$ . As a result, while all the eigenimages are present in the denominator equally, the subspace associated with the signal will be weighted, with higher order eigenimages (the ones with lowest eigenvalues) being enhanced compared to the more prominent ones. By doing this, we expect that projectiona associated with

lower eigenvalues but still containing signal to be over enhanced in comparison to CE1.

$$a_i = \begin{cases} 1/\sigma_i & \text{if } \mathbf{u}_i \in \mathcal{S} \\ 0 & \text{otherwise} \end{cases} \quad b_i = 1 \quad (4)$$

We have also expanded the two methods described previously (CE1 and CE2) by relaxing their conditions on the threshold by including a smooth transition between signal and noise subspaces. We followed the soft-threshold scheme suggested in [11] and defined a linear decay for coefficients in CE1 and CE2. We call to these methods CE1s, and CE2s respectively and their definitions are shown in Eq. (5) and Eq. (6).

**CE1s:**

$$a_i = \begin{cases} 1 & \text{if } \sigma_i \geq \sigma_{\max} \\ 0 & \text{if } \sigma_i \leq \sigma_{\min} \\ \frac{\sigma_i - \sigma_{\min}}{\sigma_{\max} - \sigma_{\min}} & \text{otherwise} \end{cases} \quad ; \quad b_i = 1 \quad (5)$$

**CE2s:**

$$a_i = \begin{cases} \frac{1}{\sigma_i} & \text{if } \sigma_i \geq \sigma_{\max} \\ 0 & \text{if } \sigma_i \leq \sigma_{\min} \\ \frac{1}{\sigma_i} \frac{\sigma_i - \sigma_{\min}}{\sigma_{\max} - \sigma_{\min}} & \text{otherwise} \end{cases} \quad ; \quad b_i = 1 \quad (6)$$

The values for  $\sigma_{\max}$  and  $\sigma_{\min}$  can be picked programmatically by selecting two patches based on the mean image of the stack and computing their second eigenvalues. These patches correspond to the ones centered in the pixels with highest and lowest values, and their computed eigenvalue is assigned as  $\sigma_{\max}$  and  $\sigma_{\min}$  respectively.

As result, this methods are hybrids in terms of weighting. In the case of CE1s, in comparison to CE1, it adds a weight to the eigenimages proportional to their relevance. In the case of CE2, its the opposite as the extra term competes with the eigenvalue value. Therefore it alleviates the enhancement of lower order eigenimages.

**Further insights into the presented contrast enhancement functions:** It has been shown earlier in [13] that the the eigenimages and spatial frequencies are indeed related. The authors of [13] showed how the decomposition of the temporal auto-correlation matrix encodes not only temporal variance but also spatial distributions. The same article shows that there are multiple Fourier components of the sample present in each eigenimage, but the maximum frequency in the eigenimages increases with the order. In this sense, eigendecomposition is pseudospectral, and not precisely a spectral decomposition. Also, the presented contrast enhancement functions involve a pseudospectral low pass eigen-filtering approach, which is roughly analogous to a spectral filtering approach.

One may argue that high pass filter on the averaged image is analogous to the approach presented here. However, it is notable that the averaging operator first removes the information encoded in the temporal fluctuations arising from fluorescence photokinetics, causing a loss of information. Further, performing a high pass filter inadvertently suppresses or removes low frequency information without any consideration of statistical significance or presence of features with low frequency support, causing further loss of information. Similar arguments may be considered for bilateral filters as well. Further, the question of designing such spectral filters is also demanding. In comparison, the contrast enhancements proposed in this paper do exclude eigenimages of higher order, but the exclusion is based on the statistical evidence of their diminished significance. Therefore, similar to dimensionality reduction performed using principal component analysis, our approach chooses the statistically most significant information, which is replete with the frequency support in the eigenimages that do not get excluded.

Indeed MUSICAL is a non-quantitative and non-linear approach in its original form. However, due to the design choice  $b_i = 1$ , the contrast enhancement indicator functions presented here are



non-linear only in the sense of alpha, vaguely similar to the gamma correction. Nonetheless, MUSICAL's indicator functions, the original as well as the ones proposed in this paper indicate a confidence in the presence of an emitter better than the diffraction limited image. In this sense, the proposed approaches provide opportunities of better suppressing the out-of-focus light and providing a better contrast when the densities of intensities of the fluorophores vary across the sample. Examples to this effect are included in [Supplement 1](#) (Note S3 and S4), where simulated datasets with well-known ground truths are used. It is noted from the example in note S3 that CE2 and CE2s suffer from the non-linear weighing terms (of the form  $1/\sigma_i$  in  $a_i$ ) which gives increasing weightage to the terms of lesser statistical significance. This property indeed indicates that the presented approach can be considered as eigen-filtering where the choice of the coefficients  $a_i$  can be considered as the eigen-filter design. At the same time, the importance of  $g_i$ , i.e. the projection of the PSF at a candidate location upon the eigenimages is essentially analogous to identifying the coefficients of principal component analysis, and thereby indicating the significance of that location contributing to the measurements.

### 3. Results

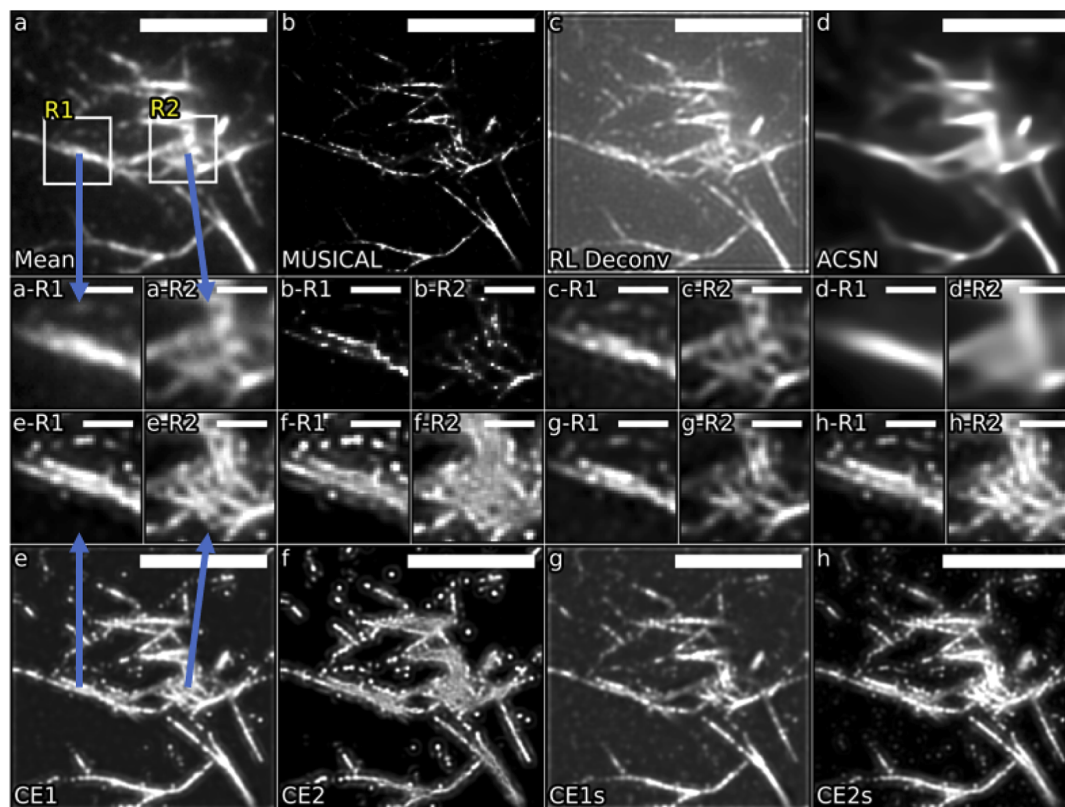
The methods were tested on experimental data available from previous works [9,10,12]. Four samples were used corresponding to in-vitro F-actin filaments [10], microtubules in U2OS cells [12], placental tissue and myocardial tissue [9]. We have added an extensive description of the acquisition methods used for tissue in Note S1 of our [Supplement 1](#). Details for the remaining samples can be found in their respective publications.

As a means of comparison, we included the results of applying Richardson-Lucy [5,6] (referred to as RL from now on) using the restoration package of scikit-image [14] and ACsN using the ImageJ/Matlab implementation. We have also included MUSICAL result with subpixelation of 1 for the different samples in order to analyze the difference between the proposed indicator functions (CE1, CE2, CE1s and CE2s) and the original one. The results of MUSICAL and our new methods were obtained using Python, by modifying just the window processing section of the MUSICAL implementation. MUSICAL results are included as the starting point as any implementation available could be used to obtain these results by setting the subpixelation to 1. The RL and ACsN results were obtained by applying both methods to the mean image which was obtained by taking the average for each pixel across time which as mentioned in the introduction, it already reduces noise. The RL was applied with the iteration parameter set to 30 and using the main lobe of the theoretical PSF for the corresponding optical system. These parameters were also the input for ACsN. For CE1, CE2 and original MUSICAL we set the threshold following the first rule of selection as described in Section 2. A special case is the threshold for placenta's MUSICAL result for which we generated an additional result using a higher value as it provided a better result. We have also included a close view of two regions of interest in most cases. These are labeled as R1 and R2 and are only marked in the mean image to provide more clarity in the results.

**In-vitro F-actin filaments:** the sample shows actin tagged with Phalloidin-Atto 565 imaged in a TIRF microscope, with numerical aperture of 1.49 and a pixel size of 65 nm (more details can be found in [10]). As the sample has already been shown suitable for MUSICAL and therefore shown to have good intensity fluctuations, it provides a good example for testing our CE methods both in terms of contrast and also in terms of the original MUSICAL's indicator function. It is also a sample with low background as TIRF illuminates a small volume of the sample and therefore it a good assessment of how CE methods deal with images with high difference between foreground and background. From a total of 10000 frames we extracted the first 500 frames and processed them to obtain the results shown in Fig. 1. The diameter of the strands is approximately 6 nm, therefore it is not possible to measure them through optical microscopy as the point-spread



function is 2 orders of magnitude larger. However, it is still valuable as single strands can be looked at separately and also globally as a network or distribution.



**Fig. 1. Improvement of contrast in in-vitro F-actin filaments labeled with Phalloidin-Atto 560 for a stack of 500 frames.** The CE methods provide better visualization of single strands and good contrast against a dark background than the mean, RL and ACsN. Additionally, dim structures are enhanced by our CE methods which facilitates visualization of the entire region. The results of the entire field of view are shown in the first row of the figure ([a-d]) and last row ([e-h]), with the name of the method written in the bottom left corner. The rows in between correspond to close views for each methods of the regions R1 and R2 labeled in **a** put close together for an easier comparison. They are labeled using the same letter as their corresponding method and the region as suffix. **a.** Mean image with two regions of interest R1 and R2 labeled with white squares. The same regions are used for the different methods. **b.** MUSICAL results for threshold -0.67. **c.** Richardson-Lucy deconvolution on the mean image. **d.** ACsN on the mean image. **e.** Contrast enhancement 1 for threshold -0.67. **f.** Contrast enhancement 2 for threshold -0.67. **g.** Contrast enhancement 1 soft. **h.** Contrast enhancement 2 soft. Scale bar is 5  $\mu\text{m}$  for whole field of view ([a-h]), 1  $\mu\text{m}$  for the close-ups.

Our CE methods provide good definition of the strands and a good delimitation between foreground and background. It is also interesting to note how dim structures that do not appear as part of strands but as single points are also enhanced. An example of this is shown in region 1 (marked in Fig. 1(a) as R1). In both MUSICAL and our methods, two strands are clearly visible, in contrast with the mean image. It is interesting to note that MUSICAL not only emphasizes the two strands but makes them thinner. This is expected as MUSICAL provides super-resolution and strands are supposed to be thin. We note that since the mean image does not show this double strand, it is expected that such structures do not appear in results of RL and ACsN. We also observed the expected local normalization properties of CE methods as dim and small structures are also

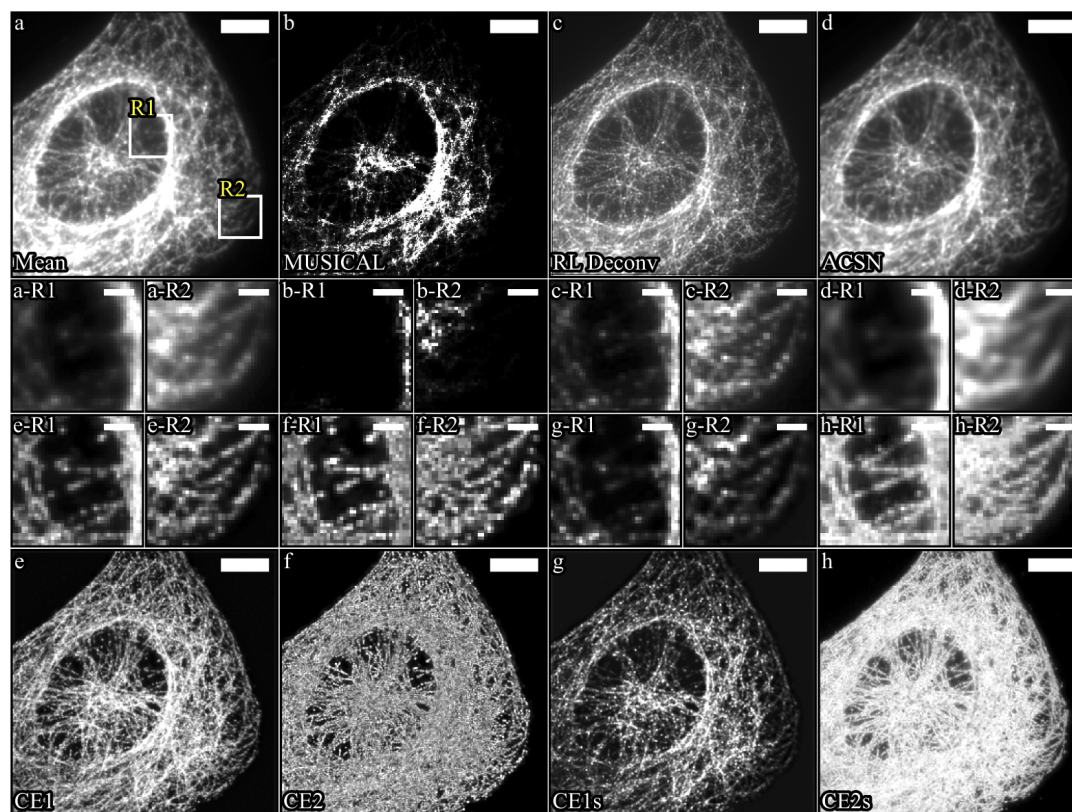
enhanced when compared to the mean. This can be appreciated in the structures placed on above of the two strands shown in R1. The second region (Fig. 1 R2) shows some strand cluttered and crossing each other. While this network can be partially seen in the mean image (Fig. 1(a) R2), its intricate structure is easily revealed after improving its contrast with CE1 and CE1s. In the case of CE2, we observed that the network becomes denser with no clear definition of single strands and with a structure that resembles a mesh. This is an example of over-enhancement where normalization of dim structure do not provide better insight into the sample. For this case, we consider this result as an artifact as the image provides an almost homogeneous region instead of lines. Such artifact is produced by the over-representation of high-frequency content derived from higher order eigenimages in CE2 due to the use of  $1/\sigma_i$  in the definition of  $a_i$ .

This is not the case for CE2s whose result is similar to the one obtained by CE1. Therefore, CE methods in general, manage to enhance the local contrast and details of bright single strands in the foreground while also enhancing the visualization of dimmer structures that are hard to observe in the mean image. In addition, the distinction between foreground and background is clear for every case.

**Microtubules:** this sample corresponds to U2OS cells in which microtubules were tagged with Alexa 647 and acquired in widefield mode at numerical aperture of 1.49 and pixel size of 108 nm. It provides a more challenging case than the actin sample presented before as the cell is a three-dimensional structure and the illumination can activate fluorescence in a larger volume than for TIRF. This contributes to poorer contrast and provides an example where contrast-enhancement can provide a different insight into the sample as several as the brightness can change across the sample due to both density and axial plane of emitters. The whole acquisition comprises 10000 frames (details can be obtained from [12]) from which we extracted and processed 500 frames to obtain the results shown in Fig. 2. From the mean it is possible to observe how the brightness of the sample is concentrated along a circular structure where the density of strands is larger than in other regions. From region 1 (labeled as R1 in Fig. 2(a)) it is possible to observe how single strand are hard to observe without saturating brighter regions of the image. The same happens in MUSICAL (Fig. 2(b)), RL (Fig. 2(c)) and ACsN (Fig. 2(d)).

On the other hand, our CE methods allow their visualization without saturating other regions, which provides a better assessment of the whole image as no fine tuning of particular regions is needed. The best results are obtained with CE1 and CE2s as the strands are enhanced when compared to the background. The results are not that clear in region 2 where the density of strands is higher. In this situation even though the methods contribute to improve contrast between strands and background, they look very similar to the mean and also to the RL results. Regarding the results of MUSICAL, there is subtle suppression of the structures in the middle and also in the periphery of the whole structure. This is alleviated by CE1 and CE1s without enhancing the already bright structures as much as the other methods. These results provide an example of intensity normalization where CE methods generate better local distribution of intensities than the mean. Background is clear and the dim structures are enhanced without affecting surrounding bright structures.

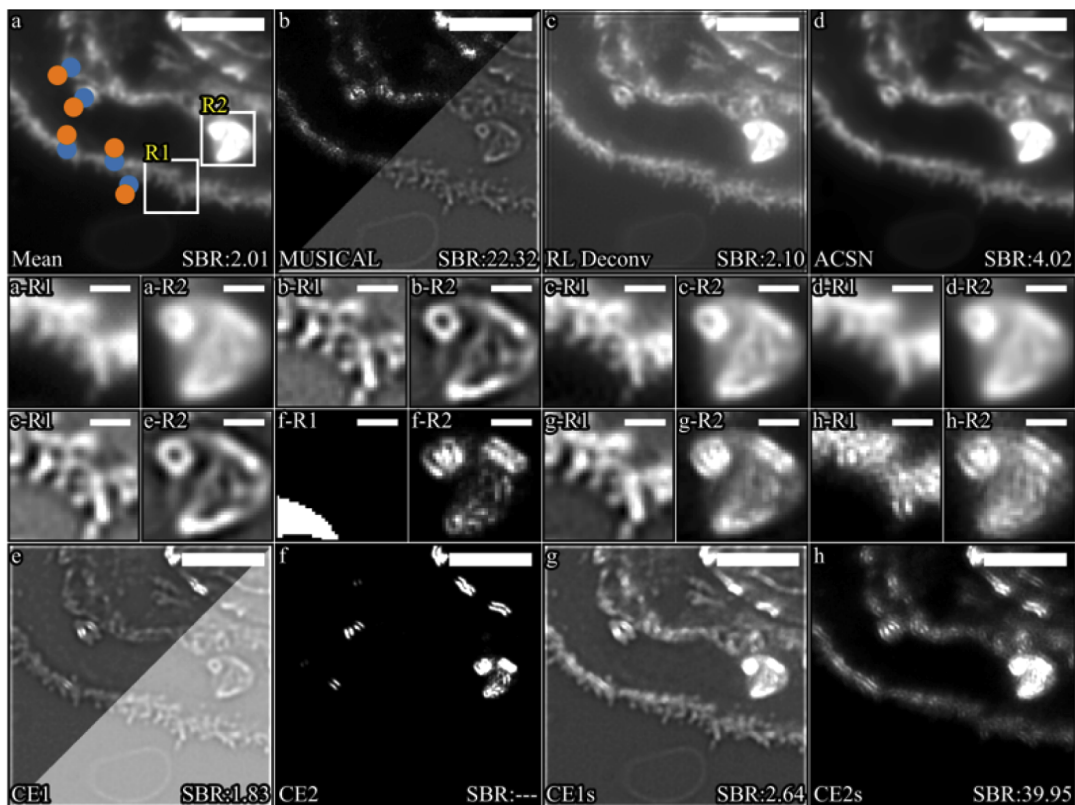
**Placenta tissue:** this sample corresponds to chorionic tissue from placenta and constitutes a different type of example when compared to the ones with strand-like structures presented before. Tissue samples are generally quite dense and therefore, differentiation between signal and background can be a problem. They may also have autofluorescence from a variety of proteins in the extra-cellular matrix. Therefore, contrast in tissue samples is a bigger challenge than in cells. For that reason, this sample provides a good study case where contrast between foreground and background is poor and enhancement may provide better insight. The results are shown on a single channel corresponding to actin labeled with Phalloidin-Atto647N and acquired in widefield illumination, with numerical aperture of 1.42 and pixel size of 80 nm. The sample thickness is 1  $\mu\text{m}$  and a stack of 500 frames was processed. As the sample is thick, contrast



**Fig. 2.** Improvement of contrast in U2OS microtubules filaments labeled with Alexa 657 for a stack of 500 frames. This sample shows the local contrast enhancement properties of CE methods as it is easier to observe the whole structure without saturation of bright regions. The results of the entire field of view are shown in the first row of the figure (**[a-d]**) and last row (**[e-h]**), with the name of the method written in the bottom left corner. The rows in between correspond to close views for each method of the regions R1 and R2 labeled in **a** put close together for an easier comparison. They are labeled using the same letter as their corresponding method and the region as suffix. **a.** Mean image with two regions of interest R1 and R2 labeled with white squares. The same regions are used for the different methods. **b.** MUSICAL results for threshold -1.5. **c.** Richardson-Lucy deconvolution on the mean image. **d.** ACsN on the mean image. **e.** Contrast enhancement 1 for threshold -1.5. **f.** Contrast enhancement 2 for threshold -1.5. **g.** Contrast enhancement 1 soft. **h.** Contrast enhancement 2 soft. Scale bar is 5  $\mu\text{m}$  for whole field of view (**[a-h]**), 1  $\mu\text{m}$  for the close-ups.



is compromised as the out-of-focus light and background compete with the signal of interest. The structure of interest corresponds to the microvilli brush border (MVBB), which are cellular membrane protrusions in the form of small hairs located at the periphery of the chorionic tissue, on the fetal side of the placenta. These are more apparent after processing the stack as can be seen on the results labeled with the R1 suffix at the middle rows of Fig. 3.



**Fig. 3. Improvement of contrast in microvilli brush border in placenta tissue with actin labeled with Phalloidin-Atto647N for a stack of 500 frames.** CE1 and CE1s make a good job at improving the edges of the structures, even though background is not suppressed. CE2 and CE2s fail as strong artifacts appear in the images in the shape of double layers. The results of the entire field of view are shown in the first row of the figure ([a-d]) and last row ([e-h]), with the name of the method written in the bottom left corner. The rows in between correspond to close-ups for each methods of the regions R1 and R2 labeled in a put close together for an easier comparison. They are labeled using the same letter as their corresponding method and the region as suffix. Additionally, each image has an estimation of its signal-to-background ratio (SBR) displayed in the bottom-right corner. **a.** Mean image with two regions of interest R1 and R2 labeled with white squares. The same regions are used for the different methods. The circles indicate the points used for computing the SBR for each image. Blue indicates signal and orange background. **b.** MUSICAL results for threshold -0.75 and -0.2. **c.** Richardson-Lucy deconvolution on the mean image. **d.** ACsN on the mean image. **e.** Contrast enhancement 1 for threshold -0.75 and -0.2. **f.** Contrast enhancement 2 for threshold -0.75. **g.** Contrast enhancement 1 soft. **h.** Contrast enhancement 2 soft. Scale bar is 5  $\mu\text{m}$  for whole field of view ([a-h]), 1  $\mu\text{m}$  for the close-ups.

Our methods produce a range of results, with the best one being CE1 and CE1s. These show a clear definition of the edges of the MVBB against the background, reducing blurriness in a similar way to the RL method across the entire structure. An example is shown in region 1 (R1). However, we note that background is not suppressed to a zero value. However, in the interface

between MVBB and the background it is possible to observe darker areas than the background which effectively helps to differentiate foreground from background. On the contrary, CE2 and CE2s completely suppress the MVBB, leaving only the brightest regions visible and several artifacts in form of artificial double layers. From the methods used for comparison, RL shows a great improvement over the mean image as the MVBB becomes visible and the background is suppressed. In the case of ACsN, this method produces a smoother results but not noticeable improvement.

Region 2 (marked as R2 in Fig. 3) shows a situation where CE2, CE1s and CE2s fail in reconstructing the right structure and is also an example where fine tuning may be necessary. On the other hand, CE1 provides the best result with some artifacts appearing for low threshold which can be improved upon increasing this value.

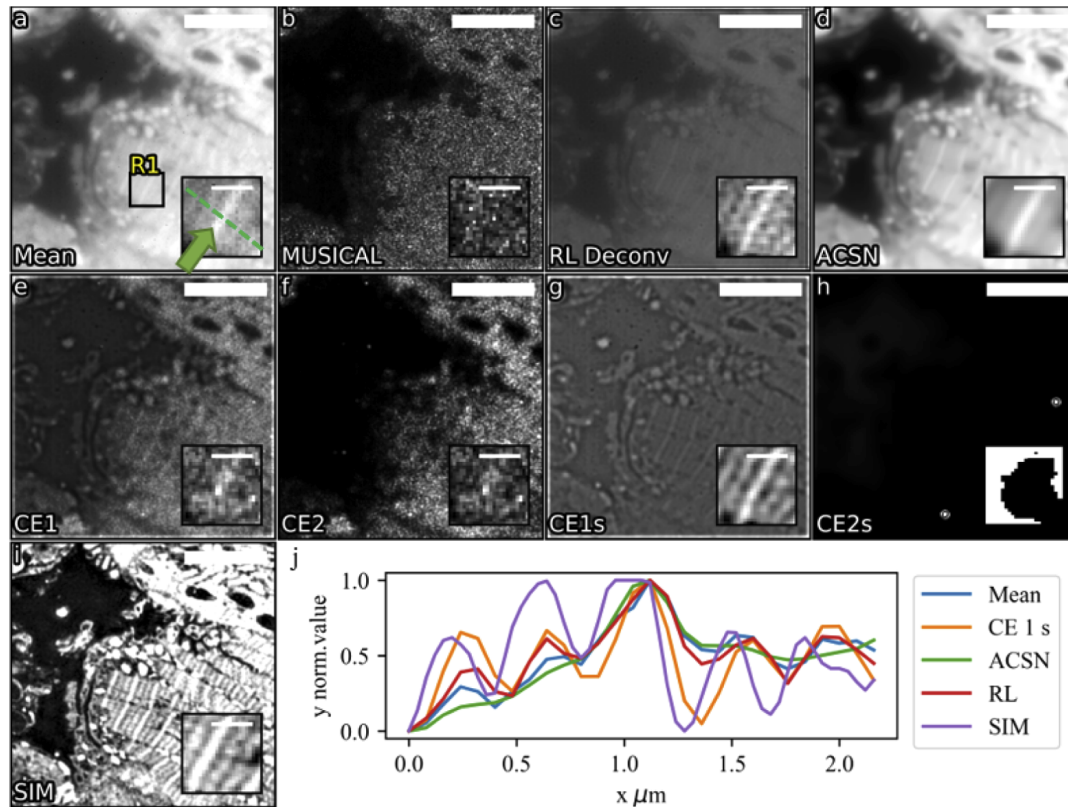
Finally, this sample also shows a case where threshold required tuning as the best results were obtained using a higher value than the default one (rule A). In this case, a higher value allows better results both for MUSICAL and CE1 as can be observed in Fig. 3(b) and 3(e) respectively. This also provides an example of the different sensitivity that CE1 and MUSICAL may have, as the former provided a good visualization of the MVBB even with rule A. Therefore, we observed that CE2 and CE2s failed in obtaining a valid result. On the other hand, CE1 and CE1s allow to effectively obtain better contrast on placenta tissue when compared to the mean, even if background is not removed.

In addition to the visual inspection, we have estimated the signal-to-background ratio for every image. In order to do so, we picked 5 pairs of signal and background locations, and took a region of 5 by 5 pixels around each. Then, we obtained the mean for each region and obtained the SBR for each pair of regions as a division of the mean intensity in the foreground region over the background region. The points are used for the estimation are shown in Fig. 3(a) in blue and orange color which represent foreground and background respectively. For the mean image, this results in a SBR of 2.01. In the case of CE2, the value was not possible to compute as the background was precisely zero in the background and therefore the division is undefined. Something similar occurs for CE2s which produces a rather large SBR due to small background values. However, this should not be considered as an indicator of good results as the visual inspection reveals easily artifacts on the image. A similar situation occurs with MUSICAL. On the other hand, CE1s seems to do a pretty good job at increasing the SBR slightly without introducing artifacts. This result correlates with the observation described before.

**Pig heart tissue:** this sample corresponds to a 100 nm thick sample, where the lipid membrane was tagged with the CellMask Orange fluorescent dye, acquired in widefield illumination, with numerical aperture of 1.42 and pixel size of 80 nm. While this sample is thinner than the placenta sample shown before, it is also denser in structures. As the lipid membrane is present across different organelles of the cell, the density of emitters can also be high which affects the contrast directly as PSF of large number of emitters overlap. In addition to the image stack, we have also used the same region taken using Structured Illumination Microscopy (SIM) [15]. SIM allows to obtain an image with improved spatial resolution compared to conventional fluorescence microscopy by a factor of two by combining several images through image processing taken under special illumination patterns (more details can be found at Note S2 of our [Supplement 1](#)). While SIM is a super-resolution technique, by producing in practice a smaller PSF, edges and also contrast are also better than conventional microscopy. Therefore, as fluctuation are not exploited, it provides an independent reference for both super-resolution and high-contrast.

Results are shown in Fig. 4 for all the methods, and the first we can observe is how the mean image (Fig. 4(a)) displays an almost uniform structure which require fine-tuning of the histogram for a nice visualization. We have marked a region of interest on it (R1) where the striations are visible as a ladder-like structure formed by the sarcomeres. The SIM image (Fig. 4(i)) makes a better job at showing such structure due to the improved resolution and contrast as mentioned

before. However, we note that that this is not a problem of resolution as the strides separation is not below the resolution limit but contrast.



**Fig. 4. Improvement of contrast in heart tissue with lipid membrane labeled with CellMask Orange for a stack of 400 frames.** The close up view of R1 (marked as black square only in **a**) included for each method as an inset shows the sarcomere, and a line profile (across the green dashed line shown in **a**) was drawn to show contrast improvement in R1. **a.** Mean image with region of interest R1 labeled with a black square and a green line use for profiling. **b.** MUSICAL results for threshold -0.38. **c.** Richardson-Lucy deconvolution on the mean image. **d.** ACsN on the mean image. **e.** Contrast enhancement 1 for threshold -0.38. **f.** Contrast enhancement 2 for threshold -0.38. **g.** Contrast enhancement 1 soft. **h.** Contrast enhancement 2 soft. **i.** SIM. **j.** Profile line across the sarcomere for the images obtained for the mean, CE1s, ACsN and RL. Scale bar is 5  $\mu\text{m}$  for whole field of view (**a-i**), 1  $\mu\text{m}$  for the close views.

From the tested methods, we obtained the best results with CE1s (Fig. 3(f)) which produced clear definition of details through contrast enhancement. As in the placenta case, background was not suppressed. A close result was obtained by deconvolution, where striation is partially visible. On the other hand, MUSICAL which performs super-resolution, reconstructs a mesh with all the details lost. This may indicate that fluorescence characteristics required by the method are not met for this dataset.

The remaining methods (CE1, CE2 and CE2s) do not show an improvement on the mean and several artifacts are visible (CE1 and CE2) or even missing information (CE2s). In the case of ACsN the image is smoother but in the process, details as striation are completely lost. Finally, as a means of comparison, we include a profile line (normalized) across the striation in Fig. 4(j). The results are shown only for the methods that worked for this sample corresponding to the

mean, CE1s, RL and ACsN. This provides an example of how contrast is improved which is reflected on the clear apparition of the ladder structure of sarcomeres.

#### 4. Discussion

The original MUSICAL indicator function is defined in a way that the denominator and numerator components have different purposes. While the denominator is the part that enables super-resolution as its value approaches zero when tested at emitters location, the numerator provides local context which enables stitching of the results of different patches. Therefore, as a super-resolution technique, it is a method that exploits predominantly the noise subspace of the imaging signal. On the other hand, our CE methods exploit only the signal part with the purpose of enabling only contrast enhancement. After testing the new methods on a variety of experimental samples, we have proven this concept for two of our methods. CE1 and CE1s present the most promising results as improvement in the quality of the images was consistent across the four samples. On the other hand, CE2 and CE2s appear to over-enhance features. This is noticeable for both tissue samples where several artifacts were obtained across the entire resulting image (Fig. 3(f,h), and 4(f,h)). However, CE1 and CE1s worked for every situation with almost no artifacts and great separation of foreground and background. In the case of CE1, we notice that even if the threshold is chosen roughly using rule A, which can be implemented programmatic, results were good. This is completely different to what was observed for the same sample in MUSICAL image where the result varied dramatically for different thresholds. An example of this is shown in Fig. 3(b) and Fig. 3(e) where results for two different thresholds are displayed. In addition, for CE1 the threshold appears to affect the local saturation but not structures recovered. This can be translated into a simple rule of thumb where the threshold can be increased until the desired saturation is obtained.

Regarding contrast, we noticed that CE1 and CE1s equalize the brightness of different structures for every sample. This is noticeable specially where structures that appear dim in the mean, are clearer and crisper after processing the stack. This can be observed in Fig. 4 where structures and details are enhanced regardless of their average brightness, with results very similar to the ones obtained by SIM. This can be explained by how MUSICAL's family of indicator function are sensitive not only to the mean but also to the second moment [12]. By decomposing the image into eigenimages and re-weighting the basis (in CE1 all the eigenimages weigh the same while in CE2 they are scaled according to their eigenvalues). As a result, CE methods can enhance dimmer structures provided they show intensity variations.

It was also observed that depending on the sample, background is not entirely suppressed by the CE methods. This is particularly visible in the tissue sample, which were precisely challenging samples as the background can be indeed a source of light. This is expected as these samples are dense or have several sources from off-focus planes. Therefore, as these methods work locally, they also enhance the low signal found in the background. However, edges were not compromised and therefore, the effect on the background is not entirely negative at least for CE1 and CE1s. In order to provide the user a darker background if needed, CE methods and the mean could be combined using image multiplication which would weight down the background and increase the foreground. However, the user should be careful as the mean could still have dim structures in the background.

As the result for CE1 and CE1s, it might be interesting to try on a larger number of samples or even revisit previous results where methods as MUSICAL may have failed. We include a small study about the sensitivity of CE1s to the accurate knowledge of the PSF in Note S5 of [Supplement 1](#). It shows an added advantage of less sensitivity of CE1s in particular and the contrast enhancement indicator functions in general over MUSICAL.

In order to avoid subjectivity, we computed some metrics on the image with the purpose of quantifying contrast enhancement. In particular, for each image we calculated the BRISQUE



index, standard deviation of the pixel intensities, and the entropy of the pixels. However, the results were inconclusive as no index was coherent with the observations or had a definitive trend that could be insightful. The numerical results are included in Note S2 of [Supplement 1](#) for the sake of completeness.

## 5. Conclusion

We have identified that it is possible to process a stack in order to obtain higher definition and contrast in fluorescence images. This work presents a novel approach for enhancing the local contrast of fluorescence microscopy images as alternatives to the mean. These methods, referred to as CE methods, are based on the indicator function of MUSICAL, originally designed for super-resolution. Such function was modified by dropping the super-resolution component and leaving only the local contrast part. We have tested the different designed methods in four samples with different degree of complexity and compare these results with Richardson-Lucy deconvolution and ACsN. Particularly, the results of our methods can be compared to the ones obtained with deconvolution but with an improved contrast of dimmer structures as they appear enhanced in the final image. From the four CE methods designed in this work, we checked that for tissue samples, CE1 and CE1s provide very detailed results. We concluded this after comparing the results of the same sample under SIM as it provided a good reference in terms of contrast and also resolution.

**Funding.** Norges Forskningsråd (288082, 804233); H2020 Excellent Science (749666); Universitetet i Tromsø (Publication fund, Strategic fund).

**Acknowledgements.** The authors thank Asa B. Birgisdottir for providing pig heart samples, Mona Nystad for providing the placental samples, and Randi Olsen for cutting all the tissue cryo-sections used in this study.

SA and MR share first authorship. KA conceived the idea. MR generated the results. LV and VD performed the tissue sample preparation, image acquisition and assisted with the image analysis under the supervision of BSA. SA supervised MR, made the figures and wrote the manuscript. All authors discussed the results and contributed to the final manuscript.

**Disclosures.** Authors declare no conflicts of interest.

**Supplemental document.** See [Supplement 1](#) for supporting content.

## References

1. E. Abbe, "Beiträge zur theorie des mikroskops und der mikroskopischen wahrnehmung," *Arch. für Mikroskopische Anat.* **9**(1), 413–468 (1873).
2. S. Koho, E. Fazeli, J. E. Eriksson, and P. E. Hänninen, "Image quality ranking method for microscopy," *Sci. Rep.* **6**(1), 28962 (2016).
3. D. Lim, T. N. Ford, K. K. Chu, and J. Metz, "Optically sectioned in vivo imaging with speckle illumination HiLo microscopy," *J. Biomed. Opt.* **16**(1), 016014 (2011).
4. R. Heintzmann and T. Huser, "Super-resolution structured illumination microscopy," *Chem. Rev.* **117**(23), 13890–13908 (2017).
5. W. H. Richardson, "Bayesian-based iterative method of image restoration," *J. Opt. Soc. Am.* **62**(1), 55–59 (1972).
6. B. L. Lucy, "An iterative technique for the rectification of observed distributions," *Astron. J.* **79**, 745 (1974).
7. B. Mandracchia, X. Hua, C. Guo, J. Son, T. Urner, and S. Jia, "Fast and accurate scmos noise correction for fluorescence microscopy," *Nat. Commun.* **11**(1), 94 (2020).
8. A. Krull, T. Buchholz, and F. Jug, "Noise2void - learning denoising from single noisy images," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, (2019), pp. 2124–2132.
9. I. S. Opstad, S. Acuña, L. E. V. Hernandez, J. Cauzzo, N. Akalko Basnet, B. S. Ahluwalia, and K. Agarwal, "Fluorescence fluctuations-based super-resolution microscopy techniques: an experimental comparative study," arXiv:2008.09195 (2020).
10. K. Agarwal and R. Macháň, "Multiple signal classification algorithm for super-resolution fluorescence microscopy," *Nat. Commun.* **7**(1), 13752 (2016).
11. S. Acuña, I. S. Opstad, F. Godtliebsen, B. S. Ahluwalia, and K. Agarwal, "Soft thresholding schemes for multiple signal classification algorithm," *Opt. Express* **28**(23), 34434–34449 (2020).
12. K. Agarwal, R. Macháň, and D. K. Prasad, "Non-heuristic automatic techniques for overcoming low signal-to-noise-ratio bias of localization microscopy and multiple signal classification algorithm," *Sci. Rep.* **8**(1), 4988 (2018).
13. K. Agarwal and D. K. Prasad, "Eigen-analysis reveals components supporting super-resolution imaging of blinking fluorophores," *Sci. Rep.* **7**(1), 4445 (2017).



14. S. Van der Walt, J. L. Schönberger, J. Nunez-Iglesias, F. Boulogne, J. D. Warner, N. Yager, E. Gouillart, and T. Yu, "scikit-image: image processing in python," *PeerJ* **2**, e453 (2014).
15. M. G. Gustafsson, "Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy," *J. Microsc.* **198**(2), 82–87 (2000).

## Deriving high contrast fluorescence microscopy images through low contrast noisy image stacks: supplement

**SEBASTIAN ACUÑA,<sup>1,3,\*</sup>  MAYANK ROY,<sup>2,3</sup> LUIS E. VILLEGAS-HERNÁNDEZ,<sup>1</sup> VISHESH K. DUBEY,<sup>1</sup>  BALPREET SINGH AHLUWALIA,<sup>1</sup>  AND KRISHNA AGARWAL<sup>1</sup>**

<sup>1</sup>*Department of Physics and Technology, UiT The Arctic University of Norway, 9010 Tromsø, Norway*

<sup>2</sup>*Indian Institute of Technology (Indian School of Mines), Dhanbad 826004, India*

<sup>3</sup>*Shared co-authors*

\**sebastian.acuna@uit.no*

---

This supplement published with The Optical Society on 11 August 2021 by The Authors under the terms of the [Creative Commons Attribution 4.0 License](https://creativecommons.org/licenses/by/4.0/) in the format provided by the authors and unedited. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

Supplement DOI: <https://doi.org/10.6084/m9.figshare.14544363>

Parent Article DOI: <https://doi.org/10.1364/BOE.422747>

## Supplementary notes: Deriving high contrast fluorescence microscopy images through low contrast noisy image stacks

\*[sebastian.acuna@uit.no](mailto:sebastian.acuna@uit.no)

### Note S1. Tissue sample preparations

Full term human placenta was collected immediately after delivery. A written consent from the participants was obtained according to the protocol approved by the Regional Committee for Medical and Health Research Ethics of North Norway (REK Nord reference no. 2010/2058-4). All the samples were anonymized. Chorionic tissue blocks of 1 mm<sup>3</sup> were dissected, washed and preserved following the Tokuyasu method for cryo-ultramicrotomy, as described elsewhere [1]. Myocardial samples (pig-heart tissue) were extracted from anesthetized pigs through a biopsy needle. After dissection into 1 mm<sup>3</sup>, the samples were further washed and preserved identically as the placental samples, including storage in liquid nitrogen. The collection of the pig heart samples followed the ethical protocols approved both by the Animal Welfare Board at UiT - The Arctic University of Norway (NO-9037 Tromsø, Norway), and the Norwegian Food Safety Authority Mattilsynet (NO-9008 Tromsø, Norway). The cryo-preserved placental tissues were sectioned into thin slices of 1 µm using a cryo-ultramicrotome (EMUC6, Leica Microsystems) and collected with a wire loop containing a 1:1 cryo-protectant mixture of 2.3M sucrose and 2% methylcellulose. Thereafter, the sections were placed onto poly-L-lysine coated microscope coverslip (#1.5) and labelled against F-actin, lipid membranes and nuclei using Phalloidin-Atto647N (1:100 solution in phosphate-buffered saline (PBS)), CellMask Orange (CMO) (1:2000 solution in PBS), and Sytox Green (10 µM in PBS), respectively. The pig heart samples were sectioned into 100 nm sections and placed onto a photonic chip [2] before further labeling against lipid membranes and nuclei using CMO (1:2000 solution in PBS), and Sytox Green (10 µM in PBS), respectively. All the labelling and mounting steps were identical as for the placental sections. The tissue sections were imaged using a DeltaVision OMX V4 Blaze microscope imaging system (OMX) (GE Healthcare) was used. The OMX microscope allowed both epi-fluorescence and structured illumination microscopy (used as a reference). The OMX was equipped with a 60X/NA1.42 oil-immersion objective lens (Olympus). For contrast enhancement, image stacks were obtained in epi-fluorescence (EPI) mode by matching the imaging channels of the microscope to the excitation wavelength of each fluorophore. For reference, structured illumination mode (SIM) images were also acquired. Laser illumination (488 nm, 568 nm, 642 nm) was used for SIM images, while the proprietary InsightSSI™ light sources were used for the EPI acquisition. The 642 nm channel was used for imaging F-actin (labelled with Phalloidin-Atto647N), and the 568 nm channel was used for imaging lipid membranes (stained with CMO). Similarly, the 488 nm channel was used for imaging nuclear content (stained with Sytox Green). Refractive index matching was performed before image acquisition. This was performed by changing the immersion oil and observing the orthogonal views of z-stack images until the point-spread function (PSF) of single emitters showed symmetrical shapes at two sides of the focal plane. Typical refractive indices used in the study ranged from 1.514 to 1.516. The image acquisition parameters were optimized according to the desired reconstruction technique. For the reference SIM images, the maximum intensity count was set to approximately 10000 for CMO, requiring acquisition times within 10 ms and 35 ms per frame. On the other hand, the intensity count in EPI mode was set below 3000, which allowed for shorter acquisition times (between 4 ms and 10 ms per frame). For the reference SIM images, the specimens were imaged at several focal planes at consecutive sampling z-steps of 125 nm. A total of 15 frames per imaging channel were acquired for each z-plane, corresponding to 5 phases and 3 orientations necessary for SIM

reconstruction. As for EPI mode, the focal plane was fixed while several images were recorded.

### Note S2. Contrast quantification

To measure contrast improvement quantitatively, we tried several strategies. While in literature comparisons are obtained by measuring PSNR or SSIM, both methods require ground-truth which is unavailable for experimental fluorescently labeled samples. Therefore, we have tried blind methods as metrics that only rely on a single image. We have collected some of the metrics in the following tables for the four samples presented in our work.

BRISQUE [3] is a blind image quality assessment algorithm available in MATLAB that analyzes the distribution of intensities across the images with the images of the same distribution and distortions (images that the algorithm has been trained with). These images have been manually evaluated in terms of quality, with 0 being the best and 100 the poorest. Therefore, an improved image should display a smaller BRISQUE index. However, we note that the algorithm was applied with the default MATLAB's model for BRISQUE which may not be suitable for microscopy images. The second metric corresponds to the standard deviation of all the intensities as a measure of dispersedness. We would expect a good image to have a larger value as it would mean that more range is used. On the contrary, an image with concentrated intensities would display a low value. The third metric is entropy which also measure the distribution and is a measure of the information contained in the image.

In the case of the In-vitro actin filaments (Table S1), we empirically observed that the best results were obtained by CE1 and C1S. However, from the BRISQUE analysis only the MUSICAL results showed improvement, while something similar occurs for standard deviation and entropy whose values are actually lower. On the other hand, ACsN achieves a better ACsN index, probably because of how the resulting image is smoother and therefore closer to natural objects. Similar results are observed for RL deconvolution. In terms of entropy, only MUSICAL achieves a large entropy which can be due to the high intensity of the reconstructed signal and very low background. This is expected as the indicator function of MUSICAL can produce very values in cases where the denominator becomes small.

Table S1. Quantification of results for In-vitro F-actin sample. BRISQUE: range of values is [0,100], lower is better. Standard deviation: higher values are better. Entropy: higher values is better.

Method	Threshold	BRISQUE	Standard deviation	Entropy
CE1	-0.67	57.033263	0.298227	0.464771
CE2	-0.67	45.233701	0.082632	0.059277
CE1s	—	45.933282	0.155533	0.167627
CE2s	—	52.747246	0.26646	0.391211
MUSICAL	-0.67	50.616861	0.035348	1.384741
ACsN	—	46.198068	0.228289	0.307909
Mean	—	53.020122	0.325196	0.529922
RL	—	44.579228	0.068589	0.043316

On the sample of microtubules (Table S2), CE1 and CE2, MUSICAL and RL gets better values. However, they differ in entropy as CE1 shows the largest among the four. In terms of standard deviation, CE1 is again among the best ranked and therefore, an indicator that its

intensity distribution could be in fact improved. This matches our empirical observations.

Table S2. Quantification of results for microtubules in U2OS cells. BRISQUE: range of values is [0,100], lower is better. Standard deviation: higher values are better. Entropy: higher values is better.

Method	Threshold	BRISQUE	Standard deviation	Entropy
CE1	-1.5	44.117168	0.44294	0.838632
CE2	-1.5	45.474566	0.2414	0.335858
CE1s	—	53.802671	0.215861	0.282117
CE2s	—	52.483188	0.471	0.917142
MUSICAL	-1.5	44.381886	0.003906	0.000266
ACsN	—	50.843798	0.307687	0.487365
Mean	—	51.734493	0.332461	0.547875
RL	—	44.429404	0.039227	0.016619

In the case of placenta (Table S3) there is almost no variation in the BRISQUE coefficient. This can be explained in most of the cases where the reconstruction shows the clear presence of villi. However, it is harder to explain for samples where the structure is not even present. In terms of standard deviation, neither of the methods show a better value than the mean. The same occurs with the entropy.

Table S3. Quantification of results for placenta tissue sample. BRISQUE: range of values is [0,100], lower is better. Standard deviation: higher values are better. Entropy: higher values is better.

Method	Threshold	BRISQUE	Standard deviation	Entropy
CE1	-0.75	44.618227	0.101571	0.083598
CE2	-0.75	44.504435	0.056479	0.03113
CE1s	—	44.854262	0.134484	0.132504
CE2s	—	44.574906	0.10776	0.092183
MUSICAL	-0.2	44.363584	0.01	0.001473
MUSICAL	-0.75	44.400236	0.036033	0.014338
ACsN	—	44.734983	0.165759	0.185666
Mean	—	45.816136	0.232132	0.316026
RL	—	44.722878	0.142464	0.145492

The pig-heart tissue sample is maybe the worst case as the SIM image which was expected to give the best result, gives the worst, which renders BRISQUE unusable as a comparison metric for this case. However, such image also shows the best (largest) deviation and entropy which can be correlated to the better quality of the image.

From the metrics used, neither seems to be useful for comparing the contrast of the results as

Table S4. Quantification of results for pig-heart tissue sample. BRISQUE: range of values is [0,100], lower is better. Standard deviation: higher values are better. Entropy: higher values is better.

Method	Threshold	BRISQUE	Standard deviation	Entropy
SIM	—	56.0157	0.492524	0.978478
CE1	-0.38	44.689645	0.101571	0.083598
CE2	-0.38	44.44668	0.024488	0.007287
CE1s	—	44.85316	0.101208	0.083105
CE2s	—	44.386956	0.01	0.001473
MUSICAL	-0.38	45.378781	0.079279	0.055298
ACsN	—	46.231884	0.476509	0.932744
Mean	—	46.037961	0.414606	0.761125
RL	—	44.363584	0.01	0.001473

they in general did not correlated with what was observed empirically. BRISQUE could still be useful in the future provided that good fluorescence images are available for a better model.

### Note S3. Contrast enhancement on simulated example

In order to achieve and quantify our contrast enhancement methods, we created a simple structure using squares. The groundtruth is shown in Fig. S1 where single emitters are placed in the perimeter of each structure, at different densities. We defined the densities in terms of the first square (labeled in blue) using a multiplier from 1 to 9 times for each case. As a result, the square at the upper-left is harder to appreciate in comparison to the one at the bottom-right. We used emission wavelength of 660 nm, numerical aperture (NA) of 1.2 and pixel size of 108 nm. The squares itself have the same size, with a side of 500 nm. The space between them is 500 nm as well. This is a distance above the resolution limit as the purpose was to analyze only contrast. Each emitters is active in average 1 every 4 frames, with a density of 100 emitters per  $\mu\text{m}$ . The noise is added later using a Poisson distribution in order to control the signal to background ratio. We used two values: 2 and 4. All the results are shown in Fig. S2.

The analysis was made by analyzing the profile line across the vertical line shown in blue in Fig. S2a. In these examples it is possible to observe how CE increase the contrast basically in two ways. One way is by looking to the centers of each square where it is possible to observe a deeper valley when compared to the mean. More interesting though, is the effect on the dimmer structure in the top left of the image, where CE1 allows an increase in the intensities, making its visualization easier.

In terms of the effect of the noise, CE1 is was clearly affected by it as the results get poor for a low SBR.

On the other hand, CE2 shows an interesting result as the 9 squares look similar in terms of brightness suggesting a normalization effect. However, definition is lost and a gradient effect is seen for both cases of noise, which is considered an artifact.

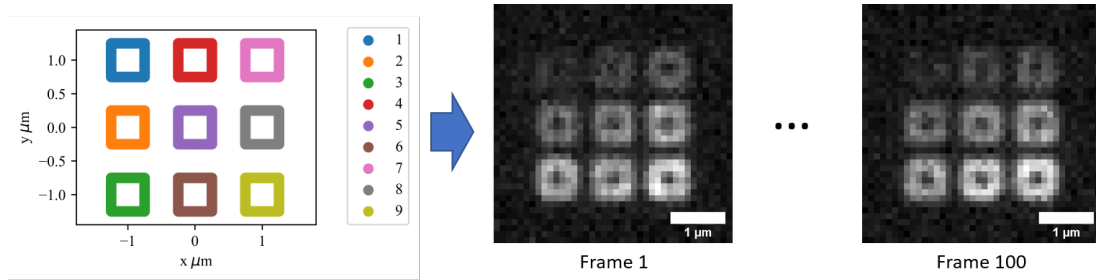


Fig. S1. Simulation of 9 flat square frames, with every square having different brightness proportional to its index (from 1 to 9).

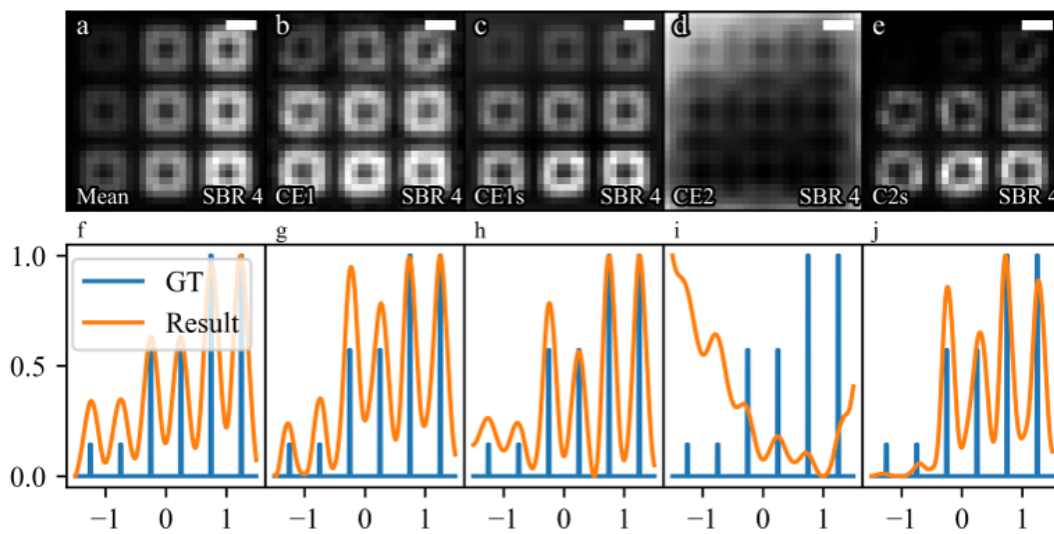


Fig. S2. Results on simulations show the effect of applying CE methods on the square sample. Scale bar is 500 nm. **a-e.** Results on sample with SBR 4. **f-j.** Profile across an horizontal line for the images shown in **a-e.**

#### Note S4. Out-of-focus removal - an illustration

We show an example of 3 spherical structures as a demonstration of the out-of-focus removal of our method. The three-dimensional structure is shown in Fig. S3a. Next to it, in Fig. S3b, we display the scatter plot of emitters located at 200 nm from the focal plane at most. The surface of the spheres is filled with emitters placed randomly. The large sphere has a radius of 1  $\mu\text{m}$  and a density of 1000 emitters per  $\mu\text{m}^2$ . The small green sphere (left) has a density of 2000 emitters  $\mu\text{m}^2$  and the orange (right) has a density of 1000. Emission wavelength is fixed at 660, with NA 1.42 and pixel size of 108 nm. Each emitter is active 1 every 4 frames on average, and Poisson noised is added for each pixel independently.

In the mean image, the haze around the two small spheres is produced by out-of-focus signal coming from emitters on the large sphere. As the PSF is wider at this positions, the number of overlap signal increases at the focal plane and as a result, we obtain an overall signal that seems constant and that does not fade away after averaging. The CE methods can filter out signal with low variance resulting from a larger number of overlapping small signals and produce an increased difference between signal and background. Further, the use of  $g_i$ , i.e. the projection of the PSF in the focal plane on the  $i$ th eigenimage allows for selectively enhancing the signal which

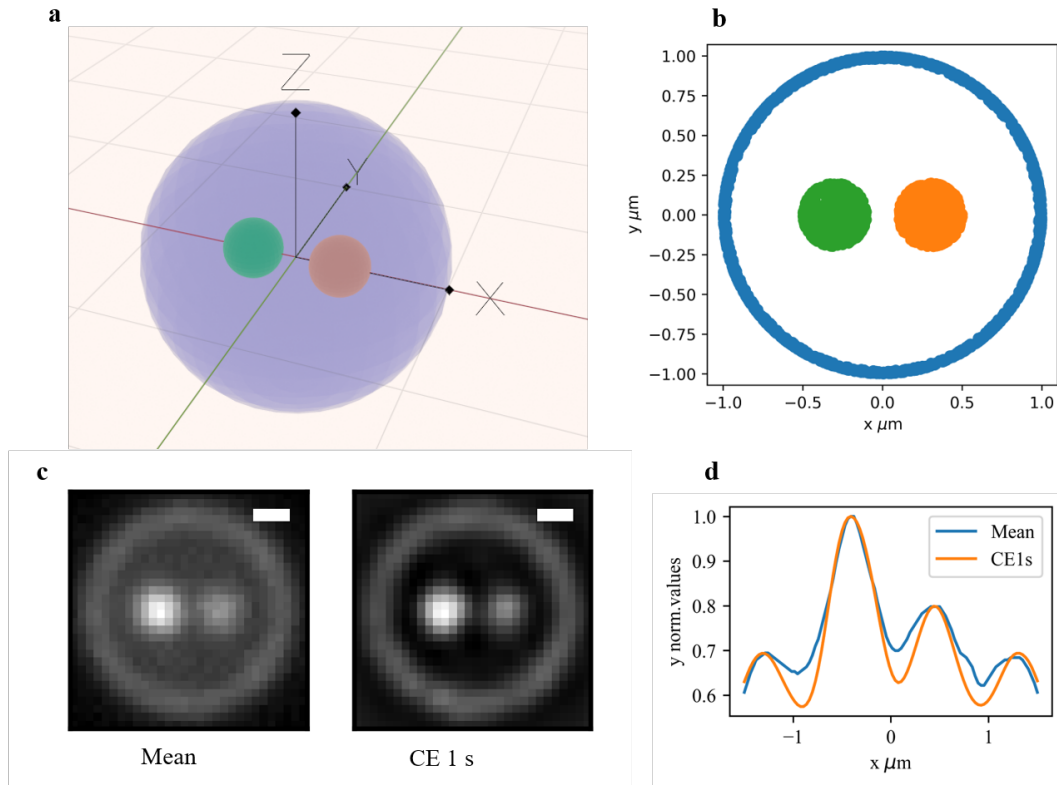


Fig. S3. Background suppression example shown on a simulation of spherical structures with labeling in the surface. The large sphere has radius  $1 \mu\text{m}$  with a density of emitter of  $1000 \text{ emitters per } \mu\text{m}^2$ , while the small ones inside have radius  $200 \text{ nm}$  with densities  $1000$  and  $2000 \text{ emitters per } \mu\text{m}^2$  respectively. The distance between centers of the two small spheres is  $600 \text{ nm}$ . **a.** 3D rendering of the structure of interest. The 3D scale has  $1 \mu\text{m}$  in each direction. **b.** Plot of actual emitters are single points. Only the ones at a maximum distance of of  $200 \text{ nm}$  from the focal plane (the one passing for all the 3 centers are display here). **c.** Result of applying the mean and CE1s. **d.**

comes from the focal region. The effect is evident in the profile across the horizontal mid-section of the images as seen in Fig. S3d. In these curves the valleys around the signal coming from the small spheres is pronounced for CE1s, indicating the enhancement in contrast.

#### Note S5. Effect of incorrect estimate of the PSF - investigation of sensitivity to PSF

Considering the dataset of pig heart tissue, presented in Fig. 4 in the main article, we investigate the sensitivity of CE1s to the accurate knowledge of the PSF. We choose CE1s since it clearly presents better performance than the other indicator functions across different challenging scenarios. While the actual dataset is acquired using an objective lens of numerical aperture (NA)  $1.42$ , we consider hypothetical situations where the NA of the objective lens is incorrectly estimated and used for constructing the CE1s image. We consider two candidates for incorrect NA, namely  $1.2$  and  $1.6$ . We present the CE1s images for all the three NAs, namely  $1.2$ ,  $1.42$ , and  $1.6$  in Fig. S4a-c respectively. The intensity at a cross-section of interest is plotted in Fig. S4d. It is seen that CE1s is not sensitive to the accurate knowledge or estimate of the PSF. On the other hand, the original MUSICAL [4] was comparatively more sensitive, presenting pronounced sharpening or blurring effect with the usage of higher or lower NA than the actual



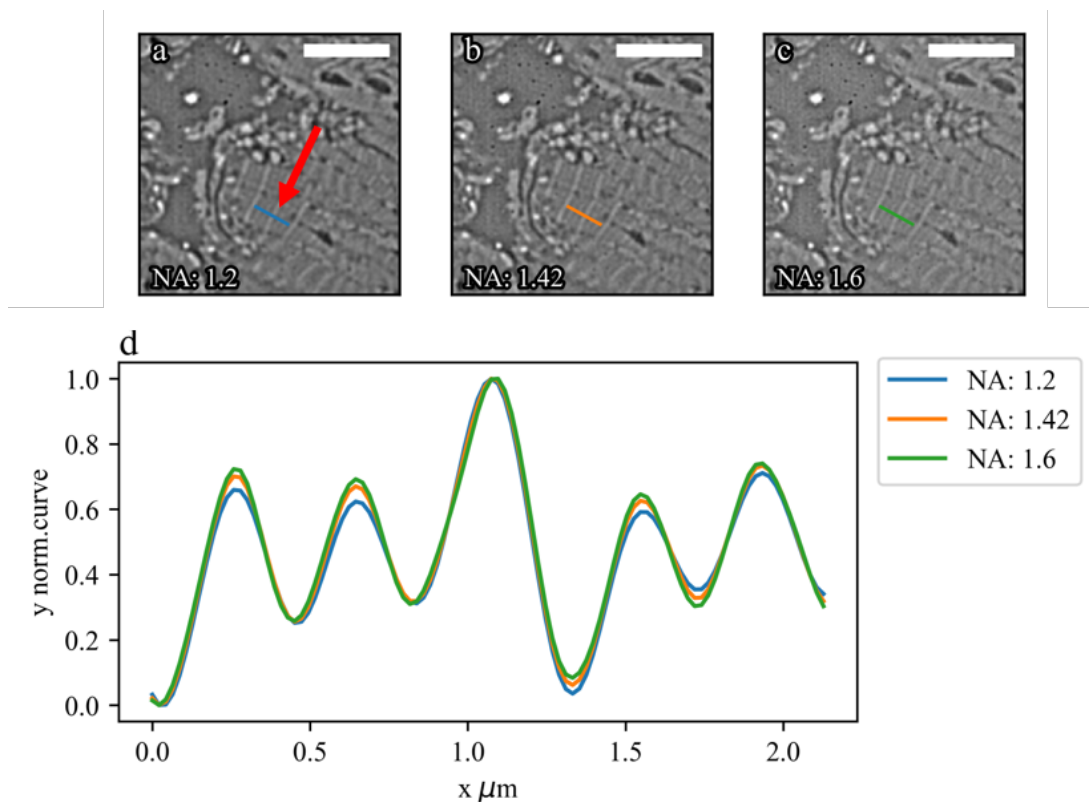


Fig. S4. Reconstructions using CE1s for heart tissue sample for different NA values that shows almost invariant results for incorrect value of NA. **a.** NA 1.2 (incorrect). **b.** NA 1.42 (correct). **c.** NA 1.6 (incorrect). **d.** Normalized profile for each case.

value. The reduced sensitivity to the incorrect NA in the indicator functions presented in this paper is attributed to the choice  $b_i = 1$ , which reduces the non-linearity of the indicator functions.

## References

1. L. E. Villegas-Hernández, M. Nystad, F. Ströhl, P. Basnet, G. Acharya, and B. S. Ahluwalia, “Visualizing ultrastructural details of placental tissue with super-resolution structured illumination microscopy,” *Placenta* **97**, 42–45 (2020).
2. J.-C. Tinguely, Ø. I. Helle, and B. S. Ahluwalia, “Silicon nitride waveguide platform for fluorescence microscopy of living cells,” *Opt. Express* **25**, 27678–27690 (2017).
3. A. Mittal, A. K. Moorthy, and A. C. Bovik, “No-reference image quality assessment in the spatial domain,” *IEEE Transactions on Image Process.* **21**, 4695–4708 (2012).
4. K. Agarwal and R. Macháň, “Multiple signal classification algorithm for super-resolution fluorescence microscopy,” *Nat. Commun.* **7**, 13752 (2016).








## Paper IV

Fluorescence fluctuation-based super-resolution  
microscopy using multimodal waveguided illumination





# Fluorescence fluctuation-based super-resolution microscopy using multimodal waveguided illumination

IDA S. OPSTAD,<sup>1,5</sup>  DANIEL H. HANSEN,<sup>1</sup> SEBASTIAN ACUÑA,<sup>1</sup>   
FLORIAN STRÖHL,<sup>1</sup>  ANISH PRIYADARSHI,<sup>1,2</sup> JEAN-CLAUDE  
TINGUELY,<sup>1</sup> FIREHUN T. DULLO,<sup>1</sup> ROY A. DALMO,<sup>3</sup>  TORE  
SETERNES,<sup>3</sup> BALPREET S. AHLUWALIA,<sup>1,4</sup>  AND KRISHNA  
AGARWAL<sup>1,6</sup>

<sup>1</sup>Department of Physics and Technology, UiT – The Arctic University of Norway, Tromsø, Norway

<sup>2</sup>Optoelectronics Research Centre, University of Southampton, Southampton SO17 1BJ, United Kingdom

<sup>3</sup>Norwegian College of Fishery Science, UiT – The Arctic University of Norway, Tromsø, Norway

<sup>4</sup>Department of Clinical Science, Intervention & Technology, Karolinska Institute, 17177 Stockholm, Sweden

<sup>5</sup>ida.s.opstad@uit.no

<sup>6</sup>uthkrishth@gmail.com

**Abstract:** Photonic chip-based total internal reflection fluorescence microscopy (c-TIRFM) is an emerging technology enabling a large TIRF excitation area decoupled from the detection objective. Additionally, due to the inherent multimodal nature of wide waveguides, it is a convenient platform for introducing temporal fluctuations in the illumination pattern. The fluorescence fluctuation-based nanoscopy technique multiple signal classification algorithm (MUSICAL) does not assume stochastic independence of the emitter emission and can therefore exploit fluctuations arising from other sources, as such multimodal illumination patterns. In this work, we demonstrate and verify the utilization of fluctuations in the illumination for super-resolution imaging using MUSICAL on actin in salmon keratocytes. The resolution improvement was measured to be 2.2–3.6-fold compared to the corresponding conventional images.

© 2021 Optical Society of America under the terms of the [OSA Open Access Publishing Agreement](#)

## 1. Introduction

Total internal reflection fluorescence microscopy (TIRFM) is a method used to obtain optical sectioning and sharp contrast of the sample close to the substrate layer [1]. The high-contrast images resulting from the elimination of out-of-focus signal of TIRFM have been exploited in 2D implementations of essentially all fluorescence-based super-resolution microscopy techniques, e.g. structured illumination microscopy (SIM) [2], stimulated emission depletion (STED) [3] and single molecule localization microscopy (SMLM) [4].

The typical lens-based implementation for TIRFM leads to disadvantages such as inhomogeneous illumination field (Gaussian-like intensity profile) and lack of flexibility in the choice of objective lens, limiting the associated sample illumination area and field-of-view (FOV). Approaches to achieve a more homogeneous illumination in lens-based TIRFM have been developed [5,6], and greater experimental flexibility—like a larger FOV and excitation area—can be achieved using prism-based illumination [7]. Nevertheless, the typical lens-based drawbacks still seem to limit most TIRFM systems.

Mass-producible photonic waveguide chips have been introduced as an illumination source for TIRFM (c-TIRFM). Transcending the main limitations of the lens-based approach and allowing for the possibilities within integrated optical systems, c-TIRFM has been demonstrated for

super-resolution applications such as SMLM [8–10] and SIM [11]. By using high refractive index materials, waveguides achieve higher spatial frequencies for illumination than far-field optics, allowing to push the achievable resolution beyond conventional implementations [8,11].

Experimental setup adaptations have also extended the gentle and arbitrarily large c-TIRFM illumination area towards live-cell imaging. For example, c-TIRFM imaging of living cells was demonstrated by Tinguely et al. [12] on cell lines (Merkel cell carcinoma (MCC13) and human trophoblast (HTR-8)) and by Opstad et al. [13] on primary neurons (from rat hippocampus and *Xenopus* retina). The natural next step of development is super-resolution c-TIRFM for live-cell imaging applications. In terms of power and acquisition time, waveguide-based SIM can be seen as the most promising implementation. However, the necessity of multiple waveguide arms, inputs, splitters and phase controllers adds steps to the fabrication process and parts to the experimental setup [11]. A more simple avenue is the exploitation of multimodal waveguide illumination patterns for fluorescence fluctuation-based super-resolution microscopy (FF-SRM) techniques. Similar to SIM and opposed to SMLM techniques, FF-SRM techniques allow for high data acquisition speed together with almost free choice of imaging media and (non-blinking) fluorophores. Photonic chip-based FF-SRM has been explored by Diekmann et al. [8] for entropy-based super-resolution imaging (ESI) [14], Priyadarshi et al. [15] for super-resolution radial fluctuations (SRRF) [16] and by Jayakumar et al. for balanced super-resolution optical fluctuation imaging (bSOFI) in combination with Haar wavelet kernel (HAWK) analysis [17].

In particular in [17], the use of computational approaches for breaking the illumination pattern correlation, which enhances the illumination artefacts in these methods, was explored. In essence, the multimodal illumination presented an obstacle instead of an opportunity in a simple application of these techniques. The underlying cause is that the temporal auto-correlation operation of these techniques correlates the illumination patterns as compared to the non-illuminated regions, artificially improving the contrast of the illumination patterns, which does not often get filtered out after a low-pass operation such as averaging or summation over time.

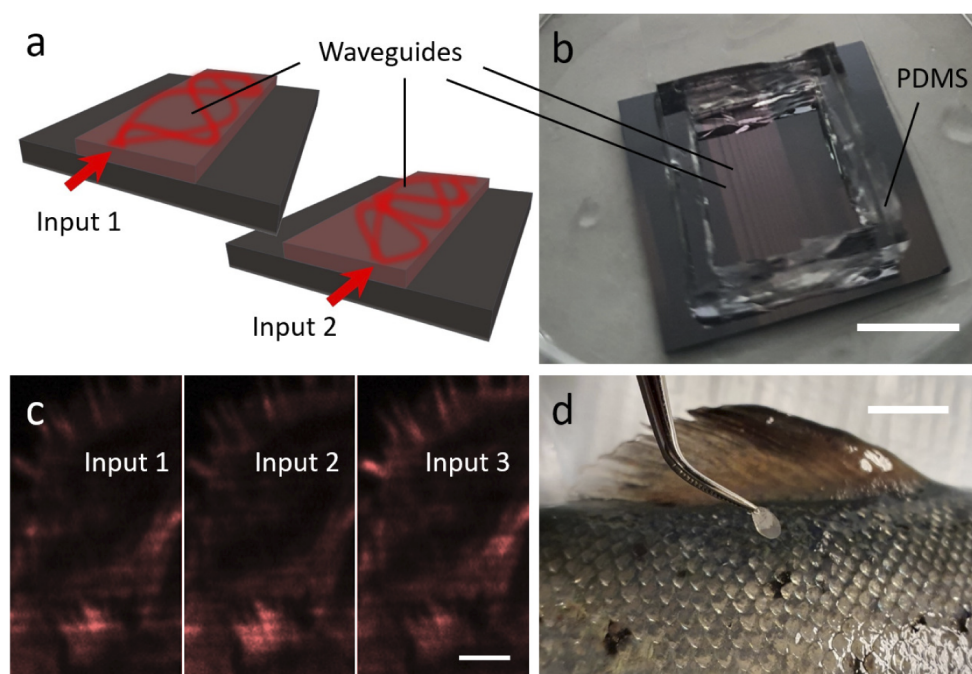
FF-SRM techniques typically rely on intrinsic low levels of signal fluctuations commonly exhibited by fluorophores under a temporally stable illumination field. Most of them work under the assumption that the temporal fluctuations in the fluorescence emissions from different fluorophores are independent of each other. This condition is violated when using illumination as a mechanism for introducing the signal fluctuations. Different from other FF-SRM methods, MUSICAL [18] utilizes eigenanalysis to extract super-resolution details from the fluctuation image sequence, allowing to exploit other sources of fluctuations, such as within the illumination pattern. Image generation from the intrinsic intensity situation of fluorophores has previously been found challenging and the cause of deleterious reconstruction artifacts for both the MUSICAL and SOFI [19] techniques [20]. The difficulties and need of long image sequences when reconstructing images from the random and often faint intrinsic fluctuations of fluorophores make it particularly interesting to explore if engineered illumination can be exploited for super-resolution techniques beyond SIM and STED.

In this work, we have explored a multimodal waveguide platform to boost the signal fluctuations for the FF-SRM technique MUSICAL. Opposed to previous efforts, the illumination variation is experimentally controlled and maximized by acquiring only one image per illumination point. We show proof of concept super-resolution reconstruction fidelity via imaging of the same sample region using different numerical aperture (NA) objective lenses, display the influence of the number of frames for resolution and reconstruction accuracy, and discuss the importance of the scan starting point of the waveguide illumination coupling. Moreover, the successful on-chip cultivation and imaging of a new type of sample—primary salmon keratocytes—is demonstrated.

## 2. Methods

### 2.1. Sample preparation and imaging

An outline of the waveguide mode-scanning and sample preparation is displayed in Fig. 1. As one point on the facet of the waveguide is illuminated, one set of multimoded waveguide modes is launched in the waveguide. Such a single coupling point corresponds to one of the inputs in Fig. 1(a). The evanescent field intensity on the top surface of the waveguide is proportional to the intensity pattern of this set of modes, which is the illumination pattern that excites the fluorescent molecules in the sample close to the waveguide surface. Each input point on the facet corresponds to a distinct illumination pattern. Therefore, the changing waveguide illumination patterns were achieved by the sequential excitation of different sets of guided modes, formed by laser coupling at different locations along the waveguide input facet (Fig. 1(a)).



**Fig. 1. Waveguide mode-scanning and sample preparation.** (a) The waveguide mode-scanning (or stepping) is achieved by changing the laser coupling point along the waveguide facet, sequentially exciting different sets of modes; (b) photonic chip with a PDMS chamber for sample and liquid confinement; (c) TIRFM images resulting from three different mode illumination patterns (650 nm excitation) and 0.3 NA collection objective; (d) salmon scale and cell harvesting. Approximate scale bars: b: 1 cm, c: 10  $\mu\text{m}$ , d: 1 cm.

To allow for on-chip cell-culture and liquid confinement, the photonic waveguide chips were prepared with a PDMS chamber as displayed in Fig. 1(b).

Cells can be prepared on the chips as on cover glasses for other types of microscopy. However, the chips used for this work are opaque and must be imaged from the top using upright microscopy. To address this limitation, Priyadarshi et al. [15] have developed photonic chips on transparent substrate for compatibility also with inverted microscopy setups. The further development of chip-bottom dishes (in a similar fashion as glass-bottom dishes for confocal microscopy) would be helpful for the chip imaging technique to become popular and widespread.

An example of signal fluctuations achieved through the described mode-scanning approach is shown in Fig. 1(c). Additionally, a video of the mode fluctuations for the same data set is available in [Visualization 1](#).



The cells used were primary keratocytes obtained from Atlantic salmon scales. These cells are of particular research interest in connection with the wound healing of farmed fish. The scale harvesting process is displayed in Fig. 1(d). The scales were placed on the photonic chips to let the keratocyte skin cells migrate from the scales onto the waveguides before chemical fixation and fluorescence labelling of filamentous actin. This approach gave a sheet of variable density of labelled keratocyte cells attached to the chip and waveguide surfaces. Scales that loosened during cell culture and sample preparation were removed from the chip culture dish, but with most of the scales still attached to the chip surface during imaging. As TIRFM cannot be performed through the scales, imaging was focused between the scales on migrated cells.

The imaging was performed using a 650 nm excitation laser together with a 670 nm longpass filter for signal collection. Complete sample preparation protocols, system description and an explanation of the waveguide mode formation are available in the detailed description of methods presented in [Supplement 1](#), chapter 1 and Fig. S1.

### 3. Results and discussion

#### 3.1. Verification of super-resolved image reconstruction

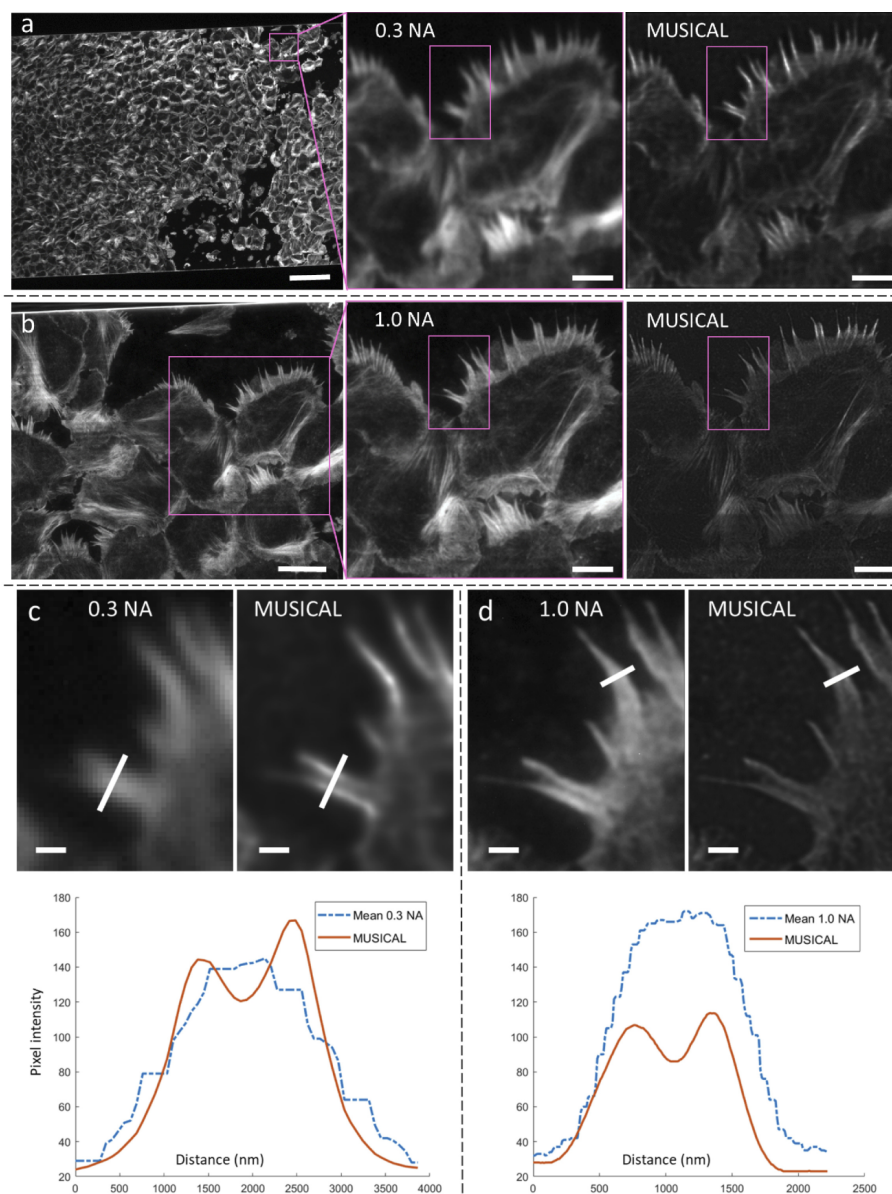
To verify that the MUSICAL super-resolution images accurately represent the underlying biological structures, the acquisition of 0.3 NA image stacks (obtained using waveguide mode-scanning) were accompanied by corresponding image stacks collected using a 1.0 NA objective lens, featuring a 3.33 times higher optical resolution according to the Abbe resolution criterion. Figure 2 shows the same sample area (filamentous actin in keratocytes) imaged using a 0.3 NA objective lens on the left (c-TIRFM mode-averaged), the MUSICAL reconstruction of the same image stack in the middle, and the mode-averaged c-TIRFM image using a 1.0 NA objective lens on the right. The results display good correspondence between the MUSICAL reconstruction and the high-NA ground-truth reference. Of particular interest are the closely spaced actin filaments unresolvable in the 0.3 NA image, but seen as clearly distinct features in the MUSICAL and 1.0 NA images. This is in contrast to the corresponding results provided by bSOFI and SRRF ([Supplement 1](#), Fig. S8), showing the two filaments as connected. The MUSICAL reconstruction of the 1.0 NA data (Fig. 2(d), right panel) shows only minor resolution improvement over the corresponding sum image (left panel). This is likely explained by the frequency support of the waveguide illumination patterns.

We bring to the readers' notice that despite a good qualitative match between the MUSICAL image and the 1.0 NA image, there are evident differences. For example, the dynamic range of the MUSICAL image is quite different compared to the 1.0 NA image, and some features appear sharper in the MUSICAL image. For example, the pink circle in Fig. 3(c) indicates a region where the actin strands are more visible in the 1.0 NA image than in the MUSICAL image, while the yellow circle in panel d indicates strands appearing with relatively higher sharpness in the 0.3 NA MUSICAL image than in the 1.0 NA reference image.

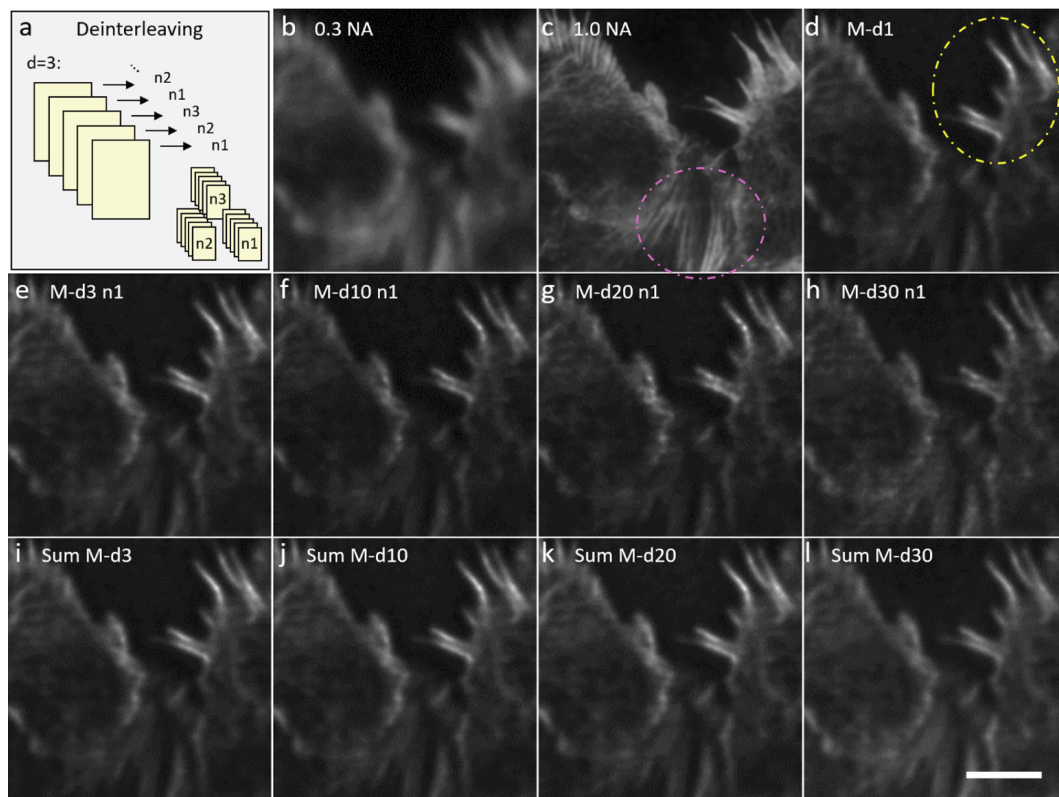
The region in the pink circle potentially corresponds to the situation where MUSICAL loses contrast due to computationally enhanced axial sectioning at the focal plane of the 0.3 NA collection objective, which may be offset from the focal plane of the collection objective in the 1.0 NA image. While the optical sectioning property of MUSICAL has been reported earlier [21], validation of this conjecture in the future will be quite important when using a high-NA image as a "ground truth reference" for MUSICAL and other super-resolution approaches. Such investigation may be further useful in the context of TIRFM systems in assessing super-sectioning ability as well as the potential loss of details. However, these investigations deserve a dedicated and elaborate treatment and are left for future work.

Now, we consider the situation of yellow circle. Regarding this situation, we note that the so-called "ground truth reference", although an image of much higher resolution than the 0.3 NA image, is still a blurred version of the underlying sample, with structural details gradually more





**Fig. 2. Large FOV c-TIRFM and computational super-resolution.** All panels are c-TIRFM images of the same 600  $\mu\text{m}$  wide waveguide covered with salmon keratocytes (labelled for F-actin). The extensions from the cells are pseudopods which aid cellular movements. (a) *Left*: Sample overview displaying the entire waveguide width captured using a 0.3 NA water dipping objective (mode-averaged intensities). Scale bar: 100  $\mu\text{m}$ . *Mid panel*: magnified view of the indicated region, scale bar: 10  $\mu\text{m}$ . *Right* MUSICAL reconstruction using the same data as averaged in the mid panel. Scale bar: 10  $\mu\text{m}$ . The indicated rectangles are displayed magnified in panel c. (b) Same as in a, but using 1.0 NA. The magnified view (mid-panel) serves as a ground truth reference for the 0.3 NA MUSICAL image in panel a. The right panel shows the MUSICAL reconstruction for the mode-averaged 1.0 NA data on the left. The indicated area is displayed magnified below in panel d. Scale bar left panel: 20  $\mu\text{m}$ ; mid and right panels: 10  $\mu\text{m}$ . (c) 0.3 NA intensity line profiles for mode-averaged (left) and MUSICAL (right); (d) the same for 1.0 NA. The scale bars are 2  $\mu\text{m}$ . The line profiles demonstrate a clear resolution improvement in the MUSICAL images as compared to the conventional (mode-averaged) images.



**Fig. 3. Effect of reconstruction stack size on image quality (0.3 NA).** (a) Subdivision of an image stack ( $d1$ ) into  $n_1, n_2, \dots, n_d$  via deinterleaving. Conventional images (mode-average) for (b) 0.3 NA and (c) 1.0 NA water dipping objectives. The pink circle indicates an area where the conventional image has significantly higher contrast than the MUSICAL image. (d) MUSICAL reconstruction of the complete 0.3 NA mode-scanned image stack ( $d1$ , 1499 frames). The yellow circle indicates a region where the MUSICAL image has higher contrast than the conventional image. (e-h) MUSICAL results using deinterleaved sub-stacks for deinterleaving parameter  $d = 3, 10, 20$  and  $30$  (for the first ( $n_1$ ) out of  $d$  total stacks). (i-l) Sum of MUSICAL reconstructions of deinterleaved image stacks, of deinterleaving parameter  $d = 3, 10, 20$  and  $30$ . The sums of MUSICAL reconstructions of smaller image stacks have significantly shorter reconstruction times without apparent loss of resolution or image quality. The scale bar is  $5 \mu\text{m}$ .

attenuated towards higher frequencies until completely invisible. Therefore, it is possible that the structures appearing sharper in the MUSICAL than in the 1 NA image are a more accurate representation of the underlying sample. In SIM imaging for comparison, the structural details in a frequency range corresponding to that of the illumination pattern become modulated with additional contrast (and image intensity) and clarity, which result into better reconstruction of features that remain unclear or invisible in the conventional image. This is not an issue unless one intends to quantify fluorescence intensities (or something derived from these) from the super-resolved images in addition to the qualitative rendering of morphological features of the sample like shape, co-localization, and orientation. In this sense, the MUSICAL images reconstructed from illumination-modulated raw data are analogous to the SIM approaches as well. We also include a more detailed note on resolution comparison between SIM and MUSICAL approaches in [Supplement 1](#), note 2.F.

For the common case of a fixed objective lens magnification and a limited camera size, the main advantage of using a 0.3 NA objective over a higher NA lens, is the possibility of (all at

once) collecting an extremely large FOV appropriate for the large waveguide excitation area. For the particular example displayed here, the 10X 0.3 NA objective captures a 36 times larger area than the 60X 1.0NA objective. As we demonstrate, the loss of resolution can be at least partly compensated for by applying a computational super-resolution algorithm post-acquisition. The MUSICAL reconstruction of the entire 0.3 NA collection area displayed in Fig. 2(a) is provided in [Supplement 1](#), Fig. S7.

For all results presented here and in the following sections, the MUSICAL threshold parameter was chosen automatically (and individually for each sub-stack) according to the soft thresholding scheme presented in [21]. This is a first demonstration of this novel thresholding scheme in the context of waveguide generated excitation patterns, resolving some major challenges in the reliability and practical usability of MUSICAL. The stack-specific soft thresholding scheme (MUS-S) was found greatly superior compared to other MUSICAL reconstruction schemes for the case of multimodal waveguides. A detailed comparison of the different thresholding schemes is presented in [Supplement 1](#), chapter 5 and Fig. S6.

### 3.2. Effect of number of multimoded illumination patterns

The number of frames, i.e. the stack size, required for reliable MUSICAL image reconstruction is an important parameter in several aspects. To avoid unnecessary lengthy reconstruction times, it is desirable to use just the sufficient number of frames for the best possible or required image quality. Long time-sequences also cause undesirable photobleaching and loss of image quality over time. From the perspective of live-cell imaging, the image acquisition time and potential phototoxicity are of considerable importance and they scale linearly with the number of frames. The fewer frames required for each MUSICAL time-point, the better the time resolution and more time-points can be acquired before photobleaching.

The stack size is equivalent to the number of multimoded illumination patterns, and consequently to the number of probing or sampling points on the input facet of the waveguide. The question arises of how to optimally choose a limited number of mode patterns for fluctuation-based super-resolution microscopy. A detailed mathematical treatment of this design problem and guidelines to how to solve it in practice is given in [Supplement 1](#), chapter 2. Below, we present the experimental verification of the effect of the input facet sampling on the MUSICAL reconstructions of actin in keratocytes. Additional studies regarding this data in connection with the mathematical aspects discussed in [Supplement 1](#), chapter 2 are presented in [Supplement 1](#), chapter 3 and Figs. S2-S4.

In order to perform an experimental verification of the effect of the input facet sampling on the MUSICAL results, we oversampled the input facet and acquired c-TIRFM images of the sample to generate an image stack with many more frames than considered necessary. For the 600  $\mu\text{m}$  wide tantalum pentoxide waveguide considered here, we acquired images by sampling the input facet every 100 nm. Then, we created smaller stacks out of the original oversampled stack through a deinterleaving approach, i.e. picking frames at a regular interval to create sub-stacks. For convenience, we refer to the over-sampled stack as the original stack, the stacks obtained by deinterleaving the original stack as the deinterleaved sub-stacks, and the interval at which the frames are picked for deinterleaved sub-stacks as the deinterleaving parameter  $d$ . Further, when performing deinterleaving, we have a choice to pick the first frame from the first  $d$  frames to create sub-stacks of the same size. The starting frame is represented using  $n$ .

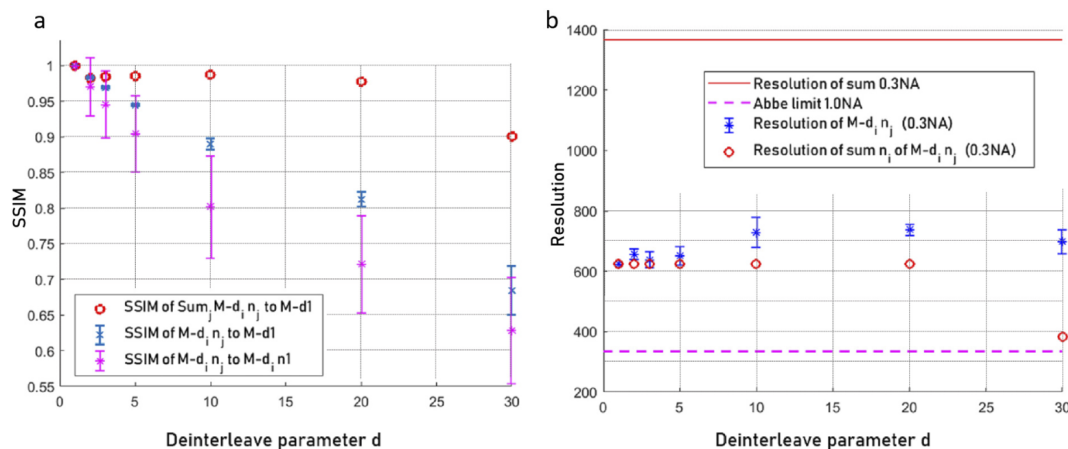
The division of the original image stack (deinterleaving parameter  $d = 1$ ) is illustrated for Fig. 3(a). Panel b show the sum image of the undivided stack  $d1$  for the 0.3 NA objective (the data used for the MUSICAL images) and panel c the equivalent for the 1.0 NA objective. The sum of mode pattern images correspond to conventional c-TIRFM images. Additionally, the 1.0 NA conventional image, serves as a ground truth reference for the 0.3 NA MUSICAL reconstructions (panels d-l) of Fig. 3. For visibility of minute differences, only a magnified region is displayed.

Figure 3(d) displays the MUSICAL reconstruction resulting from using the entire stack (referred to as M-d1 hereon). Panels e-h show reconstructions from the first sub-stacks ( $n = 1$ ) following deinterleaving into 3, 10, 20 and 30 sub-stacks. Panels i-l display the sum of all the MUSICAL reconstructions from the individual sub-stacks. The similarity and differences between these images are the topics of the following section.

### 3.2.1. Number of illumination patterns and structural similarity of MUSICAL reconstructions

The overall structural similarity between the images appears high and in accordance with the 1.0 NA ground truth reference. However, in the MUSICAL images, there is a slight appearance of horizontal stripes which are not visible in either of the conventional images. The stripes are more apparent for higher values of the deinterleaving parameter but less visible in the corresponding sum of MUSICAL images obtained using different values of  $n$ . We believe these stripes are a result of higher-order waveguide modes interfering in the direction parallel to the propagation direction, that are not visible in the conventional diffraction limited images but are recognised by the MUSICAL reconstruction algorithm due to its super-resolution ability. The illumination pattern's Fourier transform (displayed in Supplement 1, Fig. S5) shows a strong directionality in the vertical direction, which emphasized and better resolved structural features (sample and mode patterns) in the vertical direction.

The high similarity between the deinterleaved image reconstructions seen from the images in Fig. 3 were measured quantitatively using the structural similarity index (SSIM) from Wang et al. [22]. The results of this analysis are summarized in Fig. 4(a).



**Fig. 4. Effect of reconstruction stack size on image resolution and structural similarity.**

(a) Mean structural similarity index (SSIM) between the MUSICAL images obtained from sub-stacks and the M-d1 (full 1499 frame stack reconstruction) reference image (blue cross). The red circles mark the SSIM between the sum of MUSICAL images for each number of deinterleaving to the M-d1 reference. The magenta asterisks mark the mean SSIM for the individual (deinterleaved) MUSICAL images to the first sub-stack image ( $n_1$ ). The error bars indicate the SSIM standard deviation. (b) Resolution as measured by decorrelation analysis. The upper line provides a reference to the optical resolution of the conventional image (sum of modes), and the lower dashed line to the Abbe limit for the 1.0 NA objective (used as a ground truth reference in Figs. 2 and 3).

First consider the magenta bars with magenta asterisks in Fig. 4(a), which present SSIM computed between  $M-d_i n_1$  and the remaining MUSICAL images with the same value of  $d$ . These indicate the similarity between the individual MUSICAL reconstructions for the same deinterleaving parameter (i.e. sub-stacks of the same size but with different starting points),



using the MUSICAL image from the first deinterleaved image stack ( $n_1$ ) as the reference image. The SSIM between the individual MUSICAL reconstructions remain acceptably high until  $d=5$  (SSIM > 0.9), but drops to 0.80 for  $d=10$ , 0.72 for  $d=20$ , and 0.63 for  $d=30$  (mean values). This means that the sensitivity to the starting point is higher as we use fewer frames. We attribute this to two closely related factors. First, the fewer frames used, the smaller the probability of spanning all modes with sufficient intensity. Hence, the differences between the individual reconstructions result from the different sub-sets of illumination mode patterns used for each of the individual fluorescence images. Second, and associated with the first reason, the exact high frequency multimoded illumination pattern picked by MUSICAL in the background is different with each starting point and the resulting stack. Since generally the microscopy images have much more background area compared to the foreground, we suppose that the second reason is the main factor in the degrading SSIM values.

In order to better understand the waveguide mode behaviour and effect of the excitation laser step size, we performed correlation analysis of mode image stacks (Supplement 1, chapter 3). The results showed a significant correlation of the mode patterns up to a step size of about 2000 nm. As a step size of 400 nm was used for the data analysed here, deinterleaving parameter  $d=5$  correspond to the sweet spot, where redundant mode information is minimal, while optimally sampling the waveguide illumination frequencies.

Now, we consider the blue bars in Fig. 4(a). The blue crosses indicate the mean value of the SSIM measurements for the individual MUSICAL images to the M-d1 reference. This SSIM falls off fast compared to the sum MUSICAL images, but remains high above 0.9 until  $d=10$ . As compared to the magenta curves, these SSIM measurements are higher and far more consistent (i.e. low standard deviation) and definitely present a more optimistic prospect concerning the use of fewer frames for the MUSICAL images. This is with an exception of—if looking closely—the presence of high frequency horizontal stripes in the background. The better mean value of SSIM and the smaller standard deviation of the magenta compared to the blue curves is attributed to the following. The reference M-d1 image has insignificant horizontal stripe artifacts in the background, and the foreground patterns in both the M-d1 reference and the M- $d_i n_j$  match well in general. Therefore, the structural dissimilarity arises only in the locations where M- $d_i n_j$  have the strip patterns.

While the blue and magenta plots in Fig. 4(a) considered the effect of using fewer frames for super-resolution, we also consider the effect of using the same number of frames as the original stack but processed differently. This is in hope of reducing the presence of the undesirable horizontal features in the sum of M- $d_i n_j$  for a fixed value of  $d$ , as seen in Fig. 3. The SSIM of the sum of MUSICAL images with the reference M-d1 images are shown using red circles in Fig. 4(a), and they indicate a remarkably high similarity to the M-d1 reconstruction. The SSIM remains above 0.994 until  $d=10$ , 0.981 for  $d=20$ , and 0.955 for  $d=30$ . These are intriguing and interesting results from the aspects of reconstruction time and computational resources, attributed to the reconstruction of long image sequences being slow or even impossible to reconstruct beyond the computers' available memory. The possibility of splitting up the reconstructions into smaller time-windows both significantly speeds up the reconstruction process, but also allows for both the MUSICAL reconstruction of larger FOVs (as in Supplement 1, Fig. S7) and of longer time-sequences, possibly enabling better quality super-resolution images. The discussion on deinterleaved MUSICAL reconstruction and memory requirements is further expanded in Supplement 1, chapter 6.

### 3.2.2. Resolution measurements

To objectively quantify the global image resolution and the effect of reduced number of frames, we used the parameter-free image resolution estimation based on decorrelation analysis from Descloux et al. [23]. The resolution measurements (Fig. 4(b)) display a less clear dependency on

the number of frames than observed for the SSIM. As a reference, also the resolution measurement of the corresponding conventional c-TIRFM image is provided (upper red line at 1368 nm). The lower dashed line provides a reference to the Abbe diffraction limit for the 1.0 NA ground truth reference image (335 nm), using an emission wavelength of 670 nm (the lower cutoff for the emission filter passband). The coarsest resolution for individual stacks is observed at  $d = 20$  and is approximately 750 nm, which is 1.82 times better than the reference resolution for the c-TIRFM image. Furthermore, the sum of different sub-stacks with the same value of  $d$  consistently results in better resolution than the individual sub-stacks. Interestingly, the sum of M-d30 images is measured with significantly higher resolution than any of the other measurements: 384 nm as compared to 623 nm for the M-d1 image. This is 3.56 times higher resolution than the measurement for the corresponding conventional image, and only 50 nm worse than the theoretical maximum for the 1.0 NA objective. Comparing the images of Fig. 3 by eye, we can verify that indeed the resolution of the MUSICAL d30 sum is better than the conventional 0.3 NA image and of better quality than the single M-d30 image. Further, we can subjectively judge the resolution of the M-d30 image to be slightly below the 1.0 NA conventional image. Nevertheless, the large resolution increase given by decorrelation analysis appears to be a greatly exaggerated estimate. Even with parameter-free methods available, resolution estimates remain challenging and only estimates; an exact value is normally not available for experimentally obtained data and must be expected to vary even within the same microscopy image.

### 3.3. Live-cell imaging

The chip surface has repeatedly been found suitable for both live-cell imaging and even the cultivation of primary neurons over the course of several weeks [13]. Also, for this work, the salmon keratocytes were kept alive and migrating on the chips for a couple of weeks. Biocompatibility is therefore not an issue. We present results demonstrating live-cell chip-based MUSICAL images in the [Supplement 1](#), Fig. S9.

Concerning imaging time in the context of live-cell imaging, several aspects of chip-based MUSICAL can be further improved and optimized. The total imaging time is a factor of (a) the total number of frames needed for sufficient reconstruction quality, and (b) the acquisition time for each individual frame. (a) can be improved via more sophisticated engineering of the illumination, or potentially, just including more mode patterns into the same images (i.e. mode-stepping while the camera is collecting signal). For the results presented in this article, only the mode patterns from a single coupling point was included in each frame to maximize the pattern contrast for MUSICAL reconstruction. The mixing of mode patterns in each frame might be beneficial for both the reconstruction quality and the imaging time, i.e. fewer frames needed to sufficiently cover both the sample area and to span frequency space. This investigation is the scope of future work.

Factor (b), the acquisition time for each individual frame, is the sum of camera exposure time, read-out time, and the system settling time between each laser coupling point in the case of mode scanning. The latter one has in the current implementation been greatly improved over previous set-ups, where the input-facet scanning is done rapidly via a galvo-mirror solution (described in the [Supplement 1](#), chapter 1) instead of stepping using a piezo electric actuator as in previous implementation, effectively reducing the settling time from about 20 ms per coupling point to  $< 1$  ms. The camera exposure time (about 100 ms for the results presented in Figs. 2 and 3) can be reduced by finding the minimum sufficient signal intensity needed for sufficient reconstruction quality, and/or applying a high-power excitation laser. The last factor, the camera readout time (30 ms for the full FOV used for the present work), can be improved by reading out only a smaller region of interest, or by investing in a faster camera technology. For the results presented in Figs. 2 and 3, the camera exposure and readout time were 100 ms and 30 ms, respectively. If we consider a sufficient quality d10 (using 149 frames), the total acquisition time would be 19 s.

To achieve resolution doubling while imaging living cells, TIRF SIM is likely the fastest solution, although at the cost of system complexity and with a much smaller FOV than what we have presented. On the other hand, MUSICAL can achieve a higher resolution if the intensity fluctuation images sufficiently span Fourier space, either via intrinsic fluorophore fluctuations, via engineered illumination pattern or a combination of the two. The MUSICAL-on-chip technique could undoubtedly benefit from a more sophisticated engineered illumination as illustrated by [Supplement 1 Fig. S5](#), which shows a strong directionality of the Fourier transform of the acquired image stack. This is the scope of future work.

#### 4. Summary and conclusion

We have demonstrated the super-resolution technique MUSICAL on fluctuation image data obtained using c-TIRFM with multimodal waveguides. The core source of super-resolution is the fluctuations introduced by the variety of multimoded evanescent field intensity patterns achieved by illuminating different points on the input facet of the waveguide. This is a first explicit demonstration of MUSICAL using pseudorandom illumination patterns instead of intrinsic fluorescence fluctuations for super-resolution. This is also the first time that the systematic excitation of different multimoded illumination patterns was exploited for super-resolution in a c-TIRFM imaging system. The super-resolution reconstructions from 0.3 NA image data were verified by the correlated acquisition of the same sample region using a 1.0 NA collection objective, featuring a 3.33 times higher optical resolution. The main advantage of using a lower NA lens for image acquisition is the significantly larger FOV: 36 times for this particular example.

Further, we have two contributions from an application perspective. First, we have successfully demonstrated the cultivation of a completely new sample type on chip, namely primary keratocytes harvested from Atlantic salmon. These immune cells are of immense importance and research interest from the perspective of fish wound healing, especially in the context of fish farming. Second, we show that chip-microscopy is a very good application for the recently proposed soft thresholding scheme of MUSICAL, which also makes MUSICAL free of thresholding heuristics, thereby resolving a significant bottleneck of MUSICAL. Furthermore, as a development of MUSICAL, we show that creating many sub-stacks of the original stack, performing MUSICAL on each of them, and then computing the summed image may be an interesting way of adapting MUSICAL for faster computation and artifact suppression. For illumination-induced fluctuations, this approach might achieve an even better artifact suppression than the deep-learning based approach [24].

To understand the effects of using fewer frames per super-resolution image, which is especially relevant for fast, live-cell imaging, we performed image stack deinterleaving and compared the results using SSIM and decorrelation analysis for resolution measurements. The SSIM measurements showed a good structural similarity ( $SSIM \geq 0.9$ ) between both the individual MUSICAL reconstruction of the same deinterleaving parameter and compared to the full stack reconstruction ( $d = 1$ ) until  $d = 5$ , i.e., reconstruction using only 1/5 of the full stack image frames. For the sum of MUSICAL reconstructions, the SSIM remained above 0.9 until  $d = 30$ . Correlation analysis of the different waveguide mode images revealed a considerable correlation of the illumination patterns until a laser stepping distance of 2000 nm, corresponding to  $d = 5$  in the analysed keratocyte image data acquired on 600  $\mu\text{m}$  wide tantalum pentoxide waveguides. This indicates that for these waveguides, an input facet probing step finer than 2  $\mu\text{m}$  is likely to yield significant redundancies in the illumination patterns and the available frequency support for the super-resolution image reconstruction.

Faint horizontal stripes that are not visible in the conventional images were observed in the MUSICAL reconstructions. The stripes were more pronounced in the reconstructions with fewer sampling points (i.e. larger  $d$ ). We believe that these stripes result from higher order waveguide modes that are not visible at conventional resolution, and/or that they are emphasized by the

fluctuation-based super-resolution reconstruction procedure. Fourier spectral analysis presented in Supplement 1 Fig. S5 confirmed a significant vertical directionality of the mode image stacks, which can cause these artifacts and also emphasize structural features along the vertical Fourier axis as compared to the horizontal axis.

The mode patterns are not visible in the 0.3 or 1.0 NA conventional (sum) images because the modes average out over many frames of different coupling points. However, although the intensity in sum is fairly uniform, the waveguide surface is not hit by the modes in a uniform manner. In particular, since all the modes are excited along the optical axis there is likely a dominating set of modes in this very direction. We believe the fluctuation-based technique MUSICAL captures the most prominent of these mode fluctuations in addition to the excited sample at the waveguide surface, although the effect is faint and hardly visible. The high ability of the 1.0 NA objective in resolving the mode patterns additionally leads to only a very minor resolution improvement using MUSICAL compared to what is achieved for the 0.3 NA data, for which the mode fluctuations are significantly below the resolution limit (important for fluctuation-based super-resolution methods).

The decorrelation analysis gave a resolution estimate of 1368 nm for the conventional c-TIRFM image, 623 nm for the full stack (1499 frames) MUSICAL image, and down to 384 nm for the sum of deinterleaved reconstructions (for d30), a resolution improvement of 2.2 to 3.6 compared to the conventional image. We conjecture that the higher resolution measured for the d30 compared to the other similar MUSICAL images, arose from the higher visibility of waveguide mode patterns rather than actual improved resolvability of the sample.

Although significant improvements can be made with respect to the waveguide modes' frequency content, chip-based MUSICAL appears overall as a promising platform for super-resolution microscopy that will be an interesting avenue to explore both for high-content screening and live-cell imaging applications in the future.

**Funding.** Universitetet i Tromsø (UiT Publication funding, UiT strategic funding); Norges Forskningsråd (301401); European Research Council (336716); H2020 Marie Skłodowska-Curie Actions (749666, 836355).

**Acknowledgments.** BSA, KA, and FS acknowledge Horizon2020 MSCA-IF fundings (749666, 836355). ISO, DHH, SA, and JC acknowledge UiT strategic funding. FTD and BSA acknowledge Horizon 2020 European Research Council funds (336716). RAD acknowledges financial support from RCN (Grant no. 301401) and from UiT – The Arctic University of Norway.

**Disclosures.** BSA has applied for a patent on chip-based optical nanoscopy. BSA is a co-founder of the company Chip NanoImaging AS, which commercializes on-chip super-resolution microscopy systems.

DHH and FS designed the imaging system. DHH built and synchronized the imaging set-up. DHH and ISO acquired the image data. JT did the waveguide design and preparation for imaging. AP did the waveguide chip fabrication. FTD contributed towards waveguide designs and initial experimental testing for the project. ISO designed the imaging experiments, harvested scale samples, cultivated the cells, fixed and labeled the cells and prepared the chips with chamber for cell culture and microscopy. RD and TS provided salmon for cell harvesting and advised on cell culture and cell biology. SA performed MUSICAL image reconstructions for various thresholding schemes and the mode cross-correlation study. KA performed the mathematical derivation of chip-based MUSICAL. ISO analysed the data, prepared the figures and a manuscript draft. All authors commented on and contributed to the writing of the manuscript. KA and BSA supervised the project and obtained the funding.

**Data availability.** Data underlying the results presented in this paper are available in [25].

**Supplemental document.** See Supplement 1 for supporting content.






## References

1. P. J. Verveer, *Advanced Fluorescence Microscopy* (Humana, Springer protocols, 2015).
2. L. J. Young, F. Ströhl, and C. F. Kaminski, "A guide to structured illumination tirf microscopy at high speed with multiple colors," *J. Vis. Exp.* **111**, e53988 (2016).
3. T. J. Gould, J. R. Myers, and J. Bewersdorf, "Total internal reflection sted microscopy," *Opt. Express* **19**(14), 13351–13357 (2011).
4. H. Ma, R. Fu, J. Xu, and Y. Liu, "A simple and cost-effective setup for super-resolution localization microscopy," *Sci. Rep.* **7**, 1542 (2017).



5. C. J. Rowlands, F. Ströhl, P. P. V. Ramirez, K. M. Scherer, and C. F. Kaminski, “Flat-field super-resolution localization microscopy with a low-cost refractive beam-shaping element,” *Sci. Rep.* **8**(1), 5630–5638 (2018).
6. A. L. Mattheyses, K. Shaw, and D. Axelrod, “Effective elimination of laser interference fringing in fluorescence microscopy by spinning azimuthal incidence angle,” *Microsc. Res. Tech.* **69**(8), 642–647 (2006).
7. D. R. Gibbs, A. Kaur, A. Megalathan, K. Sapkota, and S. Dhakal, “Build your own microscope: step-by-step guide for building a prism-based tfr microscope,” *Methods Protoc.* **1**(4), 40 (2018).
8. R. Diekmann, Ø. I. Helle, C. I. Øie, P. McCourt, T. R. Huser, M. Schüttpelz, and B. S. Ahluwalia, “Chip-based wide field-of-view nanoscopy,” *Nat. Photonics* **11**(5), 322–328 (2017).
9. Ø. I. Helle, D. A. Coucheron, J.-C. Tinguely, C. I. Øie, and B. S. Ahluwalia, “Nanoscopy on-a-chip: super-resolution imaging on the millimeter scale,” *Opt. Express* **27**(5), 6700–6710 (2019).
10. J.-C. Tinguely, A. M. Steyer, C. I. Øie, Ø. I. Helle, F. T. Dullo, R. Olsen, P. McCourt, Y. Schwab, and B. S. Ahluwalia, “Photonic-chip assisted correlative light and electron microscopy,” (2019).
11. Ø. I. Helle, F. T. Dullo, M. Lahrberg, J.-C. Tinguely, O. G. Hellesø, and B. S. Ahluwalia, “Structured illumination microscopy using a photonic chip,” *Nat. Photonics* **14**(7), 431–438 (2020).
12. J.-C. Tinguely, Ø. I. Helle, and B. S. Ahluwalia, “Silicon nitride waveguide platform for fluorescence microscopy of living cells,” *Opt. Express* **25**(22), 27678–27690 (2017).
13. I. S. Opstad, F. Strohl, M. Fantham, C. Hockings, O. Vanderpoorten, F. van Tartwijk, J. Q. Lin, J.-C. Tinguely, F. T. Dullo, G. S. Kaminski-Schierle, B. S. Ahluwalia, and C. F. Kaminski, “A waveguide imaging platform for live-cell tfr imaging of neurons over large fields of view,” *J. Biophotonics* **13**(6), e201960222 (2020).
14. I. Yahiatene, S. Hennig, M. Müller, and T. Huser, “Entropy-based super-resolution imaging (esi): From disorder to fine detail,” *ACS Photonics* **2**(8), 1049–1056 (2015).
15. A. Priyadarshi, F. T. Dullo, D. L. Wolfson, A. Ahmad, N. Jayakumar, V. Dubey, J.-C. Tinguely, B. S. Ahluwalia, and G. S. Murugan, “A transparent waveguide chip for versatile tfr-based microscopy and nanoscopy,” (2020).
16. N. Gustafsson, S. Culley, G. Ashdown, D. M. Owen, P. M. Pereira, and R. Henriques, “Fast live-cell conventional fluorophore nanoscopy with imagej through super-resolution radial fluctuations,” *Nat. Commun.* **7**(1), 12471 (2016).
17. N. Jayakumar, Ø. I. Helle, K. Agarwal, and B. S. Ahluwalia, “On-chip tfr nanoscopy by applying haar wavelet kernel analysis on intensity fluctuations induced by chip illumination,” (2020).
18. K. Agarwal and R. Macháň, “Multiple signal classification algorithm for super-resolution fluorescence microscopy,” *Nat. Commun.* **7**(1), 13752 (2016).
19. T. Dertinger, R. Colyer, G. Iyer, S. Weiss, and J. Enderlein, “Fast, background-free, 3d super-resolution optical fluctuation imaging (sofi),” *Proc. Natl. Acad. Sci.* **106**(52), 22287–22292 (2009).
20. I. S. Opstad, S. Acuña, L. E. V. Hernandez, J. Cauzzo, N. Škalko-Basnet, B. S. Ahluwalia, and K. Agarwal, “Fluorescence fluctuations-based super-resolution microscopy techniques: an experimental comparative study,” arXiv preprint arXiv:2008.09195 (2020).
21. S. Acuña, I. S. Opstad, F. Godtliebsen, B. S. Ahluwalia, and K. Agarwal, “Soft thresholding schemes for multiple signal classification algorithm,” *Opt. Express* **28**(23), 34434–34449 (2020).
22. Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, “Image quality assessment: from error visibility to structural similarity,” *IEEE Trans. on Image Process.* **13**(4), 600–612 (2004).
23. A. C. Descloux, K. S. Grussmayer, and A. Radenovic, “Parameter-free image resolution estimation based on decorrelation analysis,” *Nat. Methods* **16**(9), 918–924 (2019).
24. S. Jadhav, S. Acuña, I. S. Opstad, B. S. Ahluwalia, K. Agarwal, and D. K. Prasad, “Artefact removal in ground truth deficient fluctuations-based nanoscopy images using deep learning,” *Biomed. Opt. Express* **12**(1), 191–210 (2021).
25. I. S. Opstad, “Replication data for: Fluorescence fluctuation-based super-resolution microscopy using multimodal waveguided illumination,” Version 1, DataverseNO, 2021, <https://doi.org/10.18710/JEN4SB>.

## Fluorescence fluctuation-based super-resolution microscopy using multimodal waveguided illumination: supplement

IDA S. OPSTAD,<sup>1,5</sup>  DANIEL H. HANSEN,<sup>1</sup> SEBASTIAN ACUÑA,<sup>1</sup>   
FLORIAN STRÖHL,<sup>1</sup>  ANISH PRIYADARSHI,<sup>1,2</sup> JEAN-CLAUDE  
TINGUELY,<sup>1</sup> FIREHUN T. DULLO,<sup>1</sup> ROY A. DALMO,<sup>3</sup>  TORE  
SETERNES,<sup>3</sup> BALPREET S. AHLUWALIA,<sup>1,4</sup>  AND KRISHNA  
AGARWAL<sup>1,6</sup>

<sup>1</sup>Department of Physics and Technology, UiT – The Arctic University of Norway, Tromsø, Norway

<sup>2</sup>Optoelectronics Research Centre, University of Southampton, Southampton SO17 1BJ, United Kingdom

<sup>3</sup>Norwegian College of Fishery Science, UiT – The Arctic University of Norway, Tromsø, Norway

<sup>4</sup>Department of Clinical Science, Intervention & Technology, Karolinska Institute, 17177 Stockholm, Sweden

<sup>5</sup>[ida.s.opstad@uit.no](mailto:ida.s.opstad@uit.no)

<sup>6</sup>[uthkrishth@gmail.com](mailto:uthkrishth@gmail.com)

---

This supplement published with The Optical Society on 9 July 2021 by The Authors under the terms of the [Creative Commons Attribution 4.0 License](https://creativecommons.org/licenses/by/4.0/) in the format provided by the authors and unedited. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

Supplement DOI: <https://doi.org/10.6084/m9.figshare.14782359>

Parent Article DOI: <https://doi.org/10.1364/OE.423809>

# Fluorescence fluctuation-based super-resolution microscopy using multimodal waveguided illumination: supplemental document

## CONTENTS

<b>1</b>	<b>Methods</b>	<b>2</b>
A	Imaging System . . . . .	2
B	Waveguide fabrication . . . . .	2
C	Sample preparation . . . . .	2
D	Image acquisition . . . . .	5
E	Data analysis . . . . .	5
<b>2</b>	<b>Mathematical derivation of chip-based MUSICAL and measurement optimality</b>	<b>5</b>
A	The design problem . . . . .	5
B	The imaging model and the role of illumination . . . . .	8
C	Criteria of optimal measurements for waveguide based illumination engineering . . . . .	9
D	Strategy for identifying the optimal illumination approach . . . . .	10
E	Resolution limit and relationship with the spatial frequencies supported by the waveguide . . . . .	11
F	Resolution in the context of structured illumination approaches, including blind and speckle illuminations . . . . .	11
<b>3</b>	<b>Cross-correlation of illumination patterns</b>	<b>11</b>
A	Analysis of the mean and variance across multiple substacks . . . . .	12
B	Interpretation of eigenimages in each pixel window . . . . .	14
<b>4</b>	<b>Fourier analysis of illumination patterns</b>	<b>14</b>
<b>5</b>	<b>MUSICAL threshold selection</b>	<b>15</b>
<b>6</b>	<b>Extreme FOVs super-resolution microscopy</b>	<b>15</b>
<b>7</b>	<b>Comparison of MUSICAL-on-chip with bSOFI and SRRF</b>	<b>17</b>
<b>8</b>	<b>Live-cell imaging</b>	<b>17</b>

## 1. METHODS

### A. Imaging System

The imaging system, waveguide coupling and formation of illumination modes are outlined in Figure S1. The imaging system is based on an upright Olympus microscope, modified to accept a four-camera system for either simultaneous multi-channel imaging, or fast multiplexed imaging. The illumination system is a custom configuration made of mostly commercially available components, but with some parts designed and fabricated in-house. A list of components is included in Table S1.

**Imaging Module** The collection module is based on an upright Olympus microscope. It has been modified by replacing the tube lens and adding a set of four cameras. The excitation light is filtered out using a quad band emission filter. After the tube lens, the image beam is split into four using dichroic mirrors. The images are then digitized using compact and low cost industrial cameras. Adjustable length mechanical adapters make sure each camera is focused correctly.

**Illumination and Modulation** During image acquisition, the photonic chip is securely mounted using a vacuum chuck. A linear stage is used to move the photonic chip relative to the coupling objective allowing selection of a desired waveguide. The waveguide is excited by directly focusing a laser onto the input facet with a long working distance objective. Coupling efficiency is optimized by precisely aligning the excitation beam with the input of a waveguide along two axes using a flexure stage.

A mirror on the galvanometer allows for changing the input angle of the laser beam, the focused beam can then be scanned across the field of view of the coupling objective along the input facet of the waveguide. Fast settling time of the galvanometer (800  $\mu$ s) allows for switching the illumination pattern in the waveguide without significantly compromising acquisition rate.

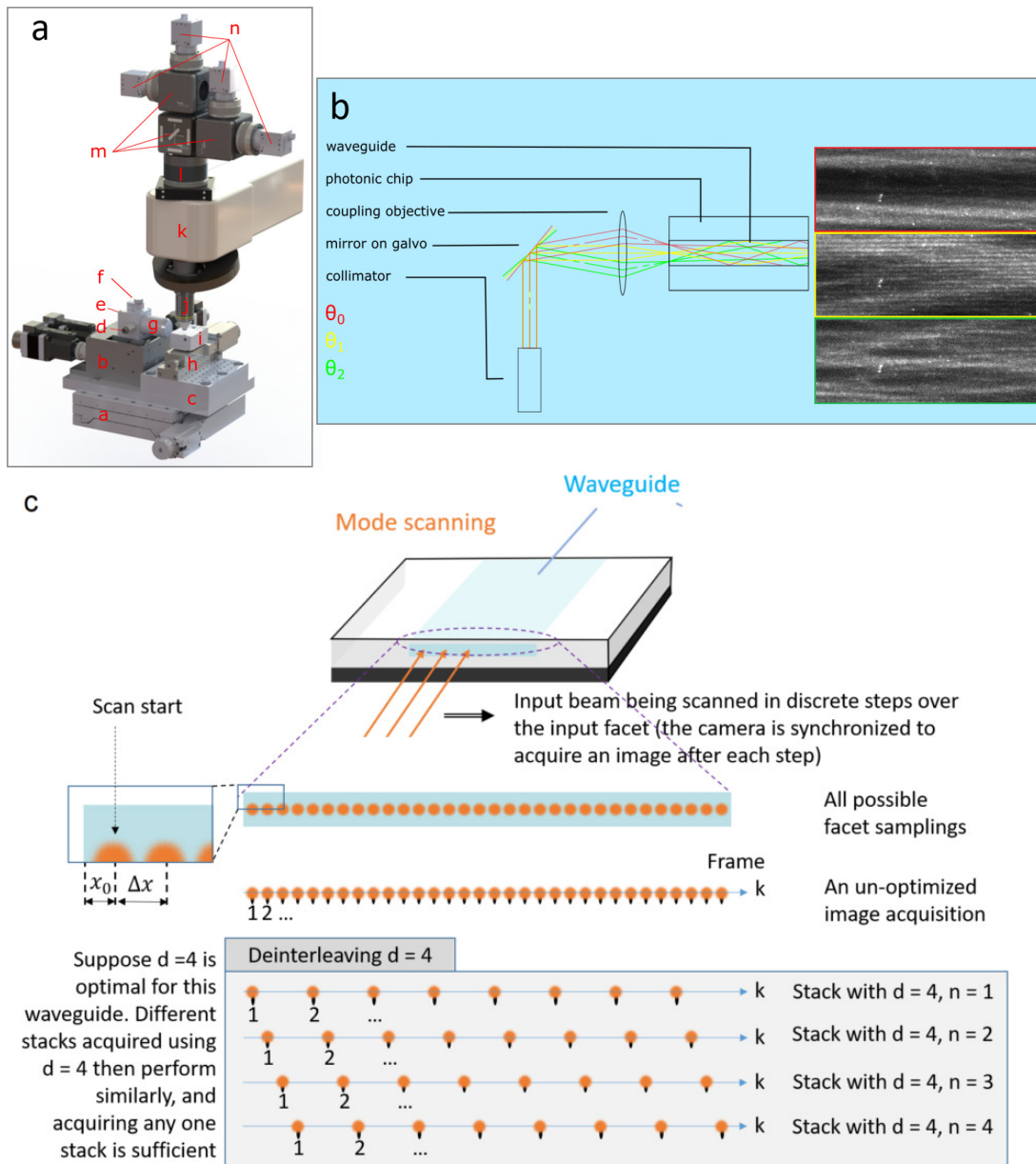
**System Control and User Interface** High level control and user interface for the system is provided through a custom LabVIEW application. Timing critical parts of the system, such as laser modulation and galvanometer position update are handled using trigger signals from the cameras, and a custom DAQ board monitoring these signals and updating the galvo position by means of a control voltage.

### B. Waveguide fabrication

A 250 nm thick Ta<sub>2</sub>O<sub>5</sub> film was deposited using a magnetron sputtering system on Si substrate with 2.5  $\mu$ m thick SiO<sub>2</sub> bottom cladding layer. The base pressure of the deposition chamber was evacuated up to  $1 \cdot 10^{-6}$  Torr, with sputtering gases being argon and oxygen (Ar:O<sub>2</sub> flow ratios of 20 SCCM: 5 SCCM). Substrate heating was applied to maintain a substrate temperature of 200°C throughout the deposition time [1][2]. The waveguide fabrication process was carried out using a conventional optical projection lithography method. First, 1  $\mu$ m thick positive resist (Shipley S1813) is coated on top of a 250 nm Ta<sub>2</sub>O<sub>5</sub> film and then prebaked at 90°C in the oven for a half-hour. Then, the wafer was mounted at the mask aligner (MA6), which illuminated the photoresist with the waveguide pattern on the mask. The masking process was optimized for the transparent wafer by varying the exposure time and development time (developed in Shipley MF319). Ion beam etching of the Ta<sub>2</sub>O<sub>5</sub> layer using argon gas was done to fabricate the waveguide. The etching was performed in the ion beam system (Ionfab 300+ from Oxford Instrument) using an argon gas flow rate of 6 SCCM. The process pressure ( $2.3 \cdot 10^{-4}$  Torr), beam voltage (500 V), beam current (100 mA), RF power (500 W), and substrate temperature (15°C) were kept constant. The etch rate goes down by reducing the crystal period. In the ion beam milling process, the substrate was placed at an angle of 45 degrees with respect to the incident ion beam to achieve low sidewall roughness. Finally, wafers are placed in a 3-zone semiconductor furnace at 600°C in an oxygen environment for 3 hours to reduce the stress and supplement the oxygen deficiency created in Ta<sub>2</sub>O<sub>5</sub> during the fabrication process.

### C. Sample preparation

The keratocytes were obtained from scales harvested from farmed Atlantic salmon. The fish were killed by a blow on the head which is allowed according to Norwegian Regulations for use of animals in experimentation ([https://lovdata.no/dokument/SF/forskrift/2015-06-18-761#KAPITTEL\\_10](https://lovdata.no/dokument/SF/forskrift/2015-06-18-761#KAPITTEL_10)). This method is also in compliance with corresponding EU legislation - Directive 2010/63/EU (<http://data.europa.eu/eli/dir/2010/63/oj>). The scales were placed on ethanol cleaned photonic chips



**Fig. S1. Chip imaging system with galvo mode scanning.** (a) Acquisition system: a: xy stage, b: coupling alignment stage, c: adaptor plate, d: fiber collimator, e: custom mount, f: galvo, g: coupling objective, h: linear stage, i: vacuum chuck, j: collection objective, k: upright microscope, l: tube lens, m: dichroic mirror mounts, n: cameras. (b) Overview of waveguide coupling, mode scanning and the formation of different TIRF illumination pattern from waveguide mode interference. (c) An illustration of image acquisition protocol for minimum distance of probing input facet ( $\Delta x$ ) and scan start location  $x_0$ , a deinterleaving parameter  $d$  for acquiring fewer frames, a stack offset parameter  $n$ . These notations are also used in supplementary note 2. Frame number is represented using  $k$ , which also indexes the probing points in the corresponding image stack.

Ref.	Description	Manufacturer	part number
a	Sample positioning stage	Standa	8MTF-75LS05
b	Coupling alignment stage	Thorlabs	MAX381/M
c	Mechanical adapter plates	Custom	-
d	Fiber collimator	Thorlabs	F280FC-A
e	Mount for collimator, galvo and coupling objective	Custom	-
f	Galvanometer with mirror	Cambridge tech.	6210H
x	Driver for Galvanometer	Cambridge tech.	671
x	DAC-board and timing controller	Custom	-
g	Coupling objective 50x0.5NA	Olympus	LMPLFLN50x
h	Waveguide selection stage	Standa	8MT30-50
i	Vacuum chuck for chip mounting	Custom	-
j	Collection objective	Nikon	MRH07120
k	Microscope body:		
	-Focusing mount	Olympus	BXFM
	-Illuminator module	Olympus	BX3M-RLA-S
	-Objective turret	Olympus	U-5RE-2
x	Camera port dovetail adapter	Thorlabs	LCPY2
x	SM30 to SM2 adapter	Thorlabs	SM2A12
x	SM2 to M30x0.5 adapter	Thorlabs	SM2A20
l	Tube lens	Thorlabs	ITL200
x	SM2 lens tube	Thorlabs	SM2M05
x	SM2 to SM1 adapter	Thorlabs	SM2A6
x	SM1 lens tubes	Thorlabs	3x SM1T2
m	Filter holders	Thorlabs	1x DFM1_M Thorlabs 2x DFM1L_M
n	Cameras	Daheng imaging	4x MER-502-79U3M
x	Focusing adapters	Thorlabs	4x SM1ZM
x	SM1 to C-mount adapters	Thorlabs	4x SM1A39

**Table S1.** Components of the imaging system. The letters in the 'Ref.' column refer to the letters in Figure S1a. Further, the letter 'x' denotes the components not directly shown in Figure S1a.



combined with a polydimethylsiloxane (PDMS) camber of height  $\approx 0.5$  cm and volume  $\approx 1$  mL (see Figure 1c). For the scales to attach to the waveguide surface, they were left to dry for about 2 min before adding antibiotic antimycotic buffer solution (Hank's Balanced Salt Solution (HBSS, Corning, #21-023-CM) with 100 IU/mL Penicillin (Alfa Aesar), 100  $\mu$ g/mL Streptomycin (VWR) (P/S) and 1  $\mu$ g/mL Amphotericin B (PanReac AppliChem, #A7009). The cells were kept at 4°C in an air tight box, replacing the old solution with fresh medium every 2-3 days. After day 4, the cells were kept in Leibovitz's L-15 medium (Gibco) with P/S. The cells and scales were fixed on chip (11 days after harvesting) in freshly thawed paraformaldehyde (4% in phosphate buffered saline (PBS)) for 1.5 hours. The cells were permeabilized using 0.1% trypsin for 4 min and washed using PBS before F-actin labelling using phalloidin-ATTO647N (ATTO-TEC) in a dilution from stock 3:100 in PBS for 1 hour while kept protected from light and at room temperature. The cells were washed in PBS and kept in the HBSS antibiotic antimycotic buffer solution in a sealed container in the fridge until imaging.

#### D. Image acquisition

Different waveguide illumination patterns were used to illuminate the sample by moving the excitation spot along the input facet of the waveguide as outlined in section S1.1. The illumination and image acquisition were synchronized such that the camera would start its exposure when the galvo mirror was stationary, and the laser would be turned on only while exposing. This approach was chosen to create a high contrast between the mode patterns captured in different image frames and to minimize photobleaching.

The acquisition time per frame is the sum of the camera exposure and readout time, plus the time of mode switching, of which the latter one is negligible compared to the former ones for our implementation with galvo scanning. In effect, the imaging rate is primarily determined by the camera technology and the fluorophore brightness.

In the experiments presented in the main article, the modes of a 600  $\mu$ m wide waveguide were excited using a 650 nm laser of power 150 mW. The distance between each coupling point along the input facet for each consecutive frame was 400 nm. The fluorescence images were collected through a 670 nm longpass emission filter. The choice of laser coupling step size and exposure times were made from a compromise on getting a high enough number of frames with a sufficient signal-to-noise ratio without deteriorating the image quality from photobleaching too much towards the end of the time sequence.

#### E. Data analysis

Image visualization, analysis and processing were conducted using Fiji [3] with ImageJ 1.53c. The MUSICAL image reconstructions were accomplished using an implementation of MUSICAL for Python, with the following parameters: 5 subpixels, lambda 660 nm, and alpha 2. The threshold was picked automatically using the work presented in [4]. More details are described in 5.

Resolution estimation based on decorrelation analysis were performed using the ImageJ plugin from [5] with default settings. SSIM calculations were performed using MATLAB (The MathWorks, Inc., version R2020a) calculating the SSIM from [6].

## 2. MATHEMATICAL DERIVATION OF CHIP-BASED MUSICAL AND MEASUREMENT OPTIMALITY

This section contains the mathematical analysis needed for optimizing the illumination design for chip-based MUSICAL. It is broken into the following sub-sections: description of the design problem needed for identifying the optimal measurement setup, description of the imaging model, the criteria designed for identifying the optimal illumination approach (and consequently the measurement approach), the protocol for designing the optimal measurement strategy for a waveguide prior to bioimaging, and lastly the resolution limit derived from the mathematical analysis. The complete list of notations are listed in Table S2.

#### A. The design problem

We first define the design problem related to the illumination approach. The input facet of a multimodal waveguide is a wide facet, and when using a focused laser beam, only a small width of the facet is sampled. Let the position of the center of the beam spot on the input facet with respect to the left edge of the waveguide facet be denoted as  $x$ . The beam is scanned over the input facet, and each illumination point (also called illumination probe) introduces a particular

Notation	Meaning
<b>Indices</b>	
$p$	Pixel number for pixels in camera. The pixels in 2D camera space may be numbered or indexed linearly using, for example, raster scan. There may be a separate look-up table for converting the 2D pixel coordinates into $p$ .
$m$	Pixel number for pixels in a super-resolved (SR) grid in the sample space. The pixels in the 2D plane illuminated by TIRF (z-axis neglected as small due to evanescent waves) may be numbered or indexed linearly using, for example, raster scan. There may be a separate look-up table for converting the 2D SR pixel coordinates into $m$ .
$k$	Frame number. This corresponds also to (a) the time coordinate as $t = k/F$ and the probing location $x$ on the input facet of the waveguide.
$q$	The index representing a propagating mode in the waveguide.
<b>Imaging setup related</b>	
$\Delta x$	The smallest scan step for the input facet of the waveguide used for designing the optimal measurement strategy.
$F$	The rate of image acquisition of the camera. In the case of the illumination engineering system used in this article, it is also the rate of switching of the scanning point on the input facet of the waveguide.
$t$	The time at which an image is taken.
$x_0$	The left edge of the waveguide (or a small practical offset point) which serves as the reference for the input facet scanning by the illumination spot.
$x$	The distance between $x_0$ and the illumination beam's spot (also referred to as illumination probe)
$d$	It is an integer, referred to as the 'deinterleaving parameter' that represents the distance between two consecutive positions of the illumination probes. Thus, if $x_k$ and $x_{k+1}$ represent two consecutive positions, then $x_{k+1} - x_k = d\Delta x$
$n$	It is an integer referred to as the 'offset parameter'. It indicates the offset from the $x_0$ in terms of $\Delta x$ for a given input facet scan. Therefore $x_k$ is completely specified as $x_k = x_0 + (n - 1)\Delta x + (k - 1)d\Delta x$ .

**Table S2.** Description of symbols and notations. Note: Table continued on the next page.



Notation	Meaning
<b>Physical quantities of interest, and related matrices if used</b>	
$I_{p,k}$ and <b>I</b>	Intensity at camera pixel $p$ in the $k$ th frame. Matrix $I$ contains $I_{p,k}$ such that one row corresponds to a pixel and one column corresponds to a frame.
$G_{p,m}$ and <b>G</b>	The point spread function mapping that maps the fluorescent emissions at an SR pixel $m$ in the sample space to a camera pixel $p$ .
$L_{m,k}$ and <b>L</b>	Illumination intensity at SR pixel $m$ in the sample region during the $k$ th frame.
$\mathcal{L}_{m,q}$	Evanescent field intensity of $q$ th propagating mode at the SR pixel $m$ in the sample region. These intensities are normalized in the sense that the maximum electric field amplitude is 1.
$a_{k,q}$	The coefficient of the $q$ th propagating mode that indicates the strength of mode in the illumination pattern in the $k$ th frame.
$f$	Spatial frequency (magnitude) in the Fourier domain
$f_{\max}$	The maximum spatial frequency of the propagating modes in the waveguide, which corresponds also to the maximum spatial frequency supported in the illumination patterns generated by the chip-based microscope
$f_{\text{obj}}$	The bandwidth of the optical transfer function of the collection system comprising of the conventional objective lens system.
<b>Statistical quantities defined for analysis</b>	
$V(p, n, d)$	The variance of intensity $I(p, k)$ across the frames (i.e. $\forall k$ ) at a given pixel $p$ when an imaging setup characterized by a particular combination of $n$ and $d$ is employed.
$A(p, n, d)$	The average of intensity $I(p, k)$ across the frames (i.e. $\forall k$ ) at a given pixel $p$ when an imaging setup characterized by a particular combination of $n$ and $d$ is employed.
$C(d, k)$	For a given frame $k$ and the deinterleaving parameter $d$ , the spatial correlation between the $k$ th frame $I(p, k)$ and its adjacent frame $I(p, k + 1)$ .

Table S2(contd.) : Description of symbols and notations

combination of the waveguide modes supported by the waveguide. Among these combinations, the excited modes propagate for long distances and produce evanescent fields at different regions of interest (ROIs) along the length of the waveguide. Therefore, by varying  $x$ , one can create illumination fluctuations in the ROIs which are to be exploited for super-resolution. Hence, the design of the optimal measurement approach corresponds to identification of a suitable set of illumination probing points  $\{x\}$ . The criteria of suitability will be derived later in this section.

From experimental practicality, we consider the constraint that the illumination facet can be probed in regular intervals via some form of scanning mechanism (employing e.g. piezo stage or galvo scanner) with the minimum scan distance between two consecutive probes being  $\Delta x$ . In other words, the input facet can be scanned in discrete multiples of  $\Delta x$ . Let the starting point of the scan be given by  $x_0$ , which is close to the left edge of the input facet. The starting point need not be particularly known as long as it is practically close to the facet edge, enabling the entire waveguide facet to be scanned. Then, the illumination probing distances can be specified as  $x_k = x_0 + (n-1)\Delta x + (k-1)d\Delta x$ , where  $d$  relates to the physical scan distance as  $d\Delta x$  and  $n$  relates to additional offset from the starting point creating the possibility of scanning starting at multiple locations. Therefore, the design parameters are  $n$  and  $d$ . We further incorporate in our design a choice of  $d$  such that the sensitivity of  $n$  is low, such that the offset from the left edge is not a factor of concern. In essence, our design parameter is  $d$  or  $d\Delta x$  for a given waveguide, illumination wavelength, and illumination probing setup.

In the experimental results presented in section 3.2 of the main manuscript, we start with an oversampled stack with  $\Delta x = 400$  nm and  $d = 1$  acquired at NA 0.3. In order to study the optimal stack size (i.e. number of illumination points) and sensitivity to the starting point, we create different stacks using different values of  $d$  and  $n \in [1, d]$ . However, if the stacks were created using optimal selection of  $d$ , the number of actual measurements would be significantly smaller. As noted in section 3.2, when we use  $d = 5$ , the structural similarity and resolution of MUSICAL images match well with the reference  $d = 1$ . In this case,  $d = 5$ , i.e. sampling distance  $d\Delta x = 2000$  nm between two consecutive frames is a good experimental design.

## B. The imaging model and the role of illumination

Consider a set of photostable emitters (fluorescent molecules) that have negligibly small standard deviation in the number of photons emitted over time. Let their average photoemission constant be  $e$  photons per unit illumination per frame. This assumption is made so that the analysis can focus on the fluctuations arising solely from variations in the illumination resulting from scanning the excitation laser along the input facet of the multimodal waveguide. Further, let us consider a hypothetical and sufficiently fine, super-resolved (SR) grid in the sample region. It is represented in the sample through the number of emitters  $E(m)$  in the  $m$ th pixel. The SR sample grid is significantly finer than the diffraction limit as well as at least two times smaller than the smallest feature of interest. Let the point spread function that maps the photons emitted at an SR pixel  $m$  in the sample to the  $p$ th camera pixel be  $G(p, m)$ . Let the illumination pattern in the sample region in the  $k$ th frame be  $L(m, k)$ . Then, the intensity image on the camera is given by:

$$I(p, k) = \sum_m eG(p, m)E(m)L(m, k) \quad (\text{S1})$$

In these notations, the following assumptions are made. The notation  $t$  denotes a particular frame number during which the camera sensor is exposed for a time  $\Delta t$ . The illumination  $L(m, k)$  is the sum of all illumination profiles delivered during the exposure of the camera. If the illumination point on the input facet is stationary during the image acquisition (such as used in this article), this consideration does not apply since the illumination is constant during the exposure time. However, the consideration applies when a continuous scan of the input facet is employed instead of discrete jumps characterized by  $\Delta x$ . We assume that the exposure time of the camera is large enough ( $>$  milliseconds), such that the photons' emitter-camera travel time is negligible. The observed intensity at a pixel  $p$  in a frame  $k$  indicates the total number of photons collected in the pixel region of the  $p$ th pixel during the exposure of the camera in that frame.

Relegating the measurements into matrix notations, eq. (S1) can be expressed as:

$$\mathbf{I} = e\mathbf{G}\mathbf{E}\mathbf{L} \quad (\text{S2})$$

where the  $(p, k)$ th element of  $\mathbf{I}$  corresponds to  $I(p, k)$ , the  $(p, m)$ th element of  $\mathbf{G}$  corresponds to  $G(p, m)$ , the  $(m, k)$ th element of  $\mathbf{L}$  corresponds to  $L(m, k)$ , and  $\mathbf{E}$  is a diagonal matrix in which the  $m$ th element contains  $E(m)$ .

MUSICAL involves eigenanalysis of  $\mathbf{J} = \mathbf{H}^T$ , which can be written as

$$\mathbf{J} = \mathbf{H}^T = \mathbf{G}\mathbf{E}\mathbf{L}\mathbf{L}^T\mathbf{E}^T\mathbf{G}^T. \quad (\text{S3})$$

Therefore, the illumination variation achieved in the waveguides contributes to MUSICAL through the mathematical support of  $\mathbf{L}\mathbf{L}^T$ . The  $(m_1, m_2)$ th element of the matrix  $\mathbf{L}\mathbf{L}^T$ , being denoted as  $l(m_1, m_2)$  for the ease of further reference, is specified as

$$l(m_1, m_2) = \sum_k L(m_1, k)L(m_2, k) \quad (\text{S4})$$

This means that each element consists of correlation in time with lag zero but for different points in the sample space. The exceptions are the diagonal elements, as they represent autocorrelation in time at a given point in the sample space. For these cases, the diagonal elements comprise the illumination's second order temporal raw moment (explained below).

### C. Criteria of optimal measurements for waveguide based illumination engineering

We first consider the diagonal elements, which correspond to the second order temporal raw moment of illumination at a given point. Mathematically, this term is the sum of the temporal variance and the square of the average signal strength of the illumination for each point in the sample space independently. In order to achieve a diagonal dominant matrix, it is important to have high variance as well as high signal strength across all the points in the sample space. In order to assess this property over all the sample points in a ROI, we can first compute the temporal variance and average signal strength at each individual point, and then form a histogram each for the distribution of variance and signal strength over all the points.

*Then the optimality in the sense of temporal variance is given by the input facet sampling distance such that the median value of the spatial distribution of the temporal variance is large but the spread of the distribution is small.*

*Similarly, the optimality in the sense of temporal average intensity is given by the input facet sampling distance such that the median value of the spatial distribution of this average intensity is large but the spread of the distribution is small.*

Furthermore, the rank of the matrix  $\mathbf{L}\mathbf{L}^T$  is quite insightful. It indicates the number of independent illumination patterns in  $\mathbf{L}$ . The number of electromagnetic modes inside the waveguide is finite, say  $Q$ , even though it might be difficult in practice to determine the value of  $Q$ . As a consequence of the finite value of  $Q$ , there is a bound to the rank of this matrix irrespective of how fine sampling is taken on the waveguide input facet. Therefore, while one may think that sampling the input facet with a very small step size between two consecutive frames provides the best measurement strategy if photobleaching of the fluorophores is not a problem, this is actually not the case. In fact, taking very many such measurements provide (a) different linear combinations of the same waveguide modes, and (b) generate independent noise patterns for every single measurement. Therefore, taking too many such measurements with small facet size does not increase the cardinality of illumination patterns but does increase the sampling of noise. This in turn leads to noisier eigenimages and therefore also to a poorer image reconstruction. On the other hand, using too few sampling points also leads to a rank below the maximum number of modes supported by the waveguide.

*Therefore, the optimal number of measurements is such that taking a higher number of measurements does not alter the rank of  $\mathbf{L}\mathbf{L}^T$ , while fewer measurements surely and systematically reduces the rank.*

An alternative way of assessing the optimal number of measurements is described next. Consider two consecutive frames that are acquired by two consecutive samplings of the input facet. If the sampling points are too close, they will result in similar or highly correlated illumination patterns. If the difference between the illumination patterns across a chosen ROI is comparable to the noise, then the two patterns do not necessarily provide exploitable information. An eigenvalue decomposition of a matrix  $\mathbf{L}$  comprising of just these two illumination patterns results in one significant eigenvalue associated with an eigenvector which is proportional to the average of the two illumination and the other eigenvalue is comparable to noise. In essence, the rank of the matrix is 1 for this noisy scenario even though two illumination patterns were used. On the other hand, as the distance between the consecutive input facet sampling points increases, the illumination patterns from such points become less correlated, providing more exploitable information. As the distance between the sampling points increases, the eigenvalue decomposition results in more and more prominent value of the second eigenvector, contributing to full rankedness of such matrix.

*In other words, the optimal sampling distance between the input facet is such that the correlation between two consecutive measurements decreases when the distance is further increased and increases consistently when the distance is decreased.*

This alternate approach for assessing the optimal sampling distance is more computationally efficient than assessing the ranks for all combinations of  $d$  and  $n$ , as required for the previously described optimization procedure.

#### D. Strategy for identifying the optimal illumination approach

In order to identify the optimal illumination approach prior to bioimaging, we performed some measurements as outlined next and processed the acquired data stacks using the criteria discussed in the previous subsection and outlined more explicitly in this section as a procedure or protocol.

**Acquiring images of illumination patterns under uniform sample conditions and over-sampling of the input facet** Since the illumination strategy needs to be designed a priori and irrespective of the sample, we consider uniform sample conditions for this process. This may be achieved in waveguide-based systems through one of the following two strategies: (a) Form a thin and uniform monolayer of fluorophores using standard protocols such as in [], or (b) autofluorescence spectrum of waveguide material may be exploited to image the illumination at the autofluorescence peak. The second strategy is more suited for the conventional homogeneous waveguide materials, but is useful only if the excitation frequency of the fluorophores of interest generates sufficient autofluorescence in the waveguide material. On the other hand, the fluorophores of interest can be directly used for the characterization by the former strategy if uniform layers can be created effectively on the waveguide surface.

Beside this, the normal microscopy setup for chip-based imaging such as described in Suppl. section 1A or other multimodal waveguide-based imaging setups described elsewhere [7] may be used. The only condition is that the same setup is used for applying MUSICAL on actual samples later. The waveguide input facet may be illuminated using optics such as described in the main paper with the only condition that as fine sampling of the input facet as possible (i.e.  $\Delta x$ ) is used and one illumination pattern is imaged for one sampling.

**Data analysis to determine the optimal  $d$**  We first define the quantities that need to be computed using the stack described above for determining  $d_{\text{opt}}$ .

Consider a pair of images  $I(p, k)$  and  $I(p, k + 1)$ . There can be  $K - d$  such image pairs for a given value of  $d$ , where  $K$  is the maximum number of frames when  $\Delta x$  is used for sampling the input facet. Let the spatial correlation between each pair be denoted as  $C(d, k)$ .

Also, let a stack with deinterleaving  $d$ , and offset  $n$  be defined as  $\mathbf{I}(n, d)$  that contains the image frames starting at frame number  $n$  and sampled every  $d$  frames thereafter for the experimental stack measured using the aforementioned experiment. Let  $V(p, n, d)$  represent the variance of intensity across the frames and  $A(p, n, d)$  represent the average intensity across the frames.

Based on the analyses presented in section C, we outline the criteria that imply a good measurement strategy for waveguide-based illumination engineering for MUSICAL.

- The univariate histogram of  $V(p, n, d)$  with respect to  $d$  is computed. In the optimal case, the median of such distribution is high but the spread is small.
- The univariate histogram of  $A(p, n, d)$  with respect to  $d$  is computed. In the optimal case, the median of such distribution is high but the spread is small.
- The univariate histogram of  $C(d, k)$  with respect to  $d$  is computed. In the optimal case, the median of such a distribution is small.

We note that the above analysis is easily generalizable to include images acquired over many ROIs, and the conclusions on  $d_{\text{opt}}$  can be derived through the above mentioned histograms by including evidences over multiple ROIs in the same histograms. The above criteria are described qualitatively and also used for illustration purposes in the following sub-sections. We recommend the above qualitative measures as more intuitive and amenable to human-based heuristics. Nonetheless, one is free to compute the medians and standard deviations of  $V(p, n, d)$ ,  $A(p, n, d)$ ,  $C(d, k)$  as functions of  $d$  for more quantitatively rigorous and automated conclusions. This aspect and also design of other single valued functions using a combination of the above mentioned histograms are out of the scope of the current work.

### E. Resolution limit and relationship with the spatial frequencies supported by the waveguide

The wide-field resolution of chip-based TIRFM and the potential of super-resolution through both single molecule localization and fluctuations based techniques have already been investigated [7–9]. At the same time, resolution and reconstruction quality of MUSICAL have also been studied previously [10, 11]. The scope of utilizing the chip-based illumination for super-resolution using MUSICAL is new and the focus of the current article.

In the above discussion, so far, we have discussed the optimality of the distance between consecutive sampling points of the waveguide input facet. Here, we discuss how the spatial frequency support of the waveguide illumination patterns in the eigenvalue decomposition contributes to resolution. Let the electromagnetic modes in the waveguides be represented as  $\mathcal{L}(m, q)$ ,  $q \in 1, Q$ , and each characterizing one spatial frequency. Let the coefficients of the linear combinations of these modes be denoted as  $a(k, q)$  such that  $L(m, k) = \sum_{q=1}^Q a(k, q)\mathcal{L}(q, m)$ . Therefore eq. (S4) can be rewritten as

$$l(m_1, m_2) = \sum_k \sum_{q_1=1}^Q \sum_{q_2=1}^Q a(q_1, k)a(q_2, k)\mathcal{L}(q_1, m_1)\mathcal{L}(q_2, m_2)$$

This indicates that the maximum spatial frequency support of the eigenimages derived using  $\mathbf{J}$  defined in eq. (S3) is given by the multiplications of all of the spatial frequency components supported by the waveguide with each other. In the frequency domain, this corresponds to the convolutions of all frequency components with each other. Therefore, if the maximum spatial frequency in the waveguide has magnitude  $k_{\max}$ , then the maximum frequency of  $\mathbf{L}\mathbf{L}^T$  is  $2k_{\max}$ . Further, multiplication with the  $G$  on either side in the definition of  $\mathbf{J}$  defined in eq. (S3) contributes additional  $2k_{\text{obj}}$  where  $k_{\text{obj}}$  is the maximum frequency supported by the optical transfer function of the collection objective. This results in the net resolution limit of MUSICAL for waveguide based illumination engineering to be determined by  $2(k_{\max} + k_{\text{obj}})$ .

### F. Resolution in the context of structured illumination approaches, including blind and speckle illuminations

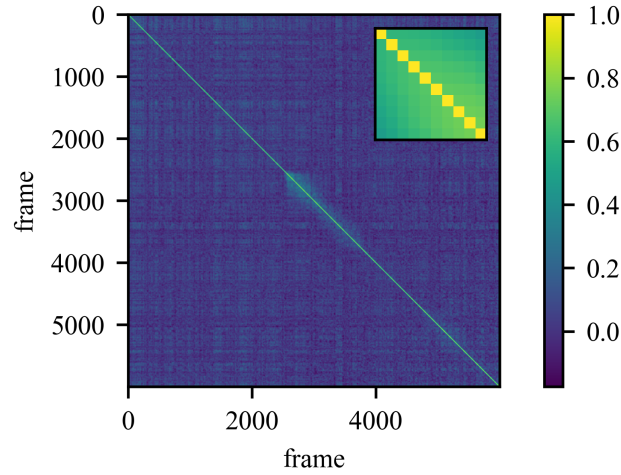
There is a certain degree of overlap between the structured illumination approaches with random or pseudo-random illuminations patterns [12–14] and the proposed chip-based MUSICAL. It is notable that the resolution of the proposed photonic-chip based nanoscope is limited in the sense of the hardware through the numerical apertures of the illumination and collection, similar to other structured illumination systems [7, 15].

However, we clarify an important difference with respect to the illumination numerical aperture between conventional optics-based systems versus chip-based imaging systems. We note that the objective lens used for coupling the light into the waveguide does not determine the spatial frequency support or the effective numerical aperture contributed by the waveguide-based illumination, as also discussed in detail in earlier works [15]. In the waveguide-based illumination system, the numerical aperture is indicated indirectly by the effective refractive index of the waveguide modes supported by the waveguide. Simulations provide an effective refractive index range from 1.9329 to 1.9317 for transverse electric polarization or 1.8469 to 1.8457 for transverse magnetic polarization at the geometry utilized for this work [7].

It is also interesting to consider the differences between MUSICAL and blind SIM given that both MUSICAL and blind SIM involve correlations. In blind SIM [13] and related algorithms, the sample's image is correlated with the illumination pattern, i.e. pattern-image correlation is performed. In MUSICAL, images taken under different illumination patterns are correlated, and therefore indirectly different illumination patterns are cross-correlated. We note that the pattern-image correlation, in the ideal case support the spectral span corresponding to  $2k_{\text{illum}} + k_{\text{obj}}$ . On the other hand, the cross-correlation of illumination patterns, such as indirectly performed in MUSICAL, supports spectral span corresponding to  $2(k_{\text{illum}} + k_{\text{obj}})$ . This is a direct consequence of the mathematical formulation presented in [16] although only fluctuations from photokinetics is considered there.

## 3. CROSS-CORRELATION OF ILLUMINATION PATTERNS

In this section we discuss the cross-correlation between frames located at different positions in time. The objective of this comparison is to analyse the difference between frames and the random



**Fig. S2.** Pearson's cross-correlation across 6000 frames for every pair of frames. The horizontal and vertical axes display the frame number.

The diagonal is full of ones as the correlation of a frame by itself is one. The inset shows a magnified-view of the top left corner of 10 by 10 frames.

patterns generated by light reflection in the chip. We captured the patterns by measuring images of the autofluorescence at discrete positions and measured the similarity between every pair of frames by obtaining the Pearson correlation coefficient between frames. The coefficient for frame number  $f_1$  and  $f_2$  where the operator  $\langle \cdot \rangle$  represents the mean of values in the vector is defined as follows:

$$C_p(f_1, f_2) = \frac{\langle (\mathbf{I}_{:,f_1} - \langle \mathbf{I}_{:,f_1} \rangle) (\mathbf{I}_{:,f_2} - \langle \mathbf{I}_{:,f_2} \rangle) \rangle}{\sqrt{\langle (\mathbf{I}_{:,f_1} - \langle \mathbf{I}_{:,f_1} \rangle)^2 \rangle} \sqrt{\langle (\mathbf{I}_{:,f_2} - \langle \mathbf{I}_{:,f_2} \rangle)^2 \rangle}} \quad (\text{S5})$$

To represent a column or row we use the colon sign ( $:$ ) to indicate that the whole column or row is considered. For instance  $\mathbf{I}_{:,1}$  corresponds to the first column, while  $\mathbf{I}_{1,:}$  is the first row.

When the correlation for every pair is obtained, they can be ordered in the correlation matrix, which is shown in Fig. S2 for the frames taken at NA 0.3 (6000 frames). As expected, the image is symmetrical since it contains every pair of frames twice, with the maximum value located across the main diagonal as it shows the correlation of each frame by itself. Due to the symmetry of the correlation matrix, we can analyse only half of the values in this similarity matrix.

Our interest lies in the off-diagonals as they represent different deinterleaving distances, starting from zero in the main diagonal. From this representation, it is also clear that for larger values of deinterleaving, fewer samples are available as the diagonals contain less pairs. Then, to make them comparable, we obtained the normalized histogram using a fixed number of bins (100) for each deinterleaving (or each off-diagonal) and collect them all in a single image as columns. The results are shown in Fig S3, where we added two curves. The mean curve is computed by getting the weighted mean of the histogram (continuous line), while the maximum is just the maximum value of each histogram.

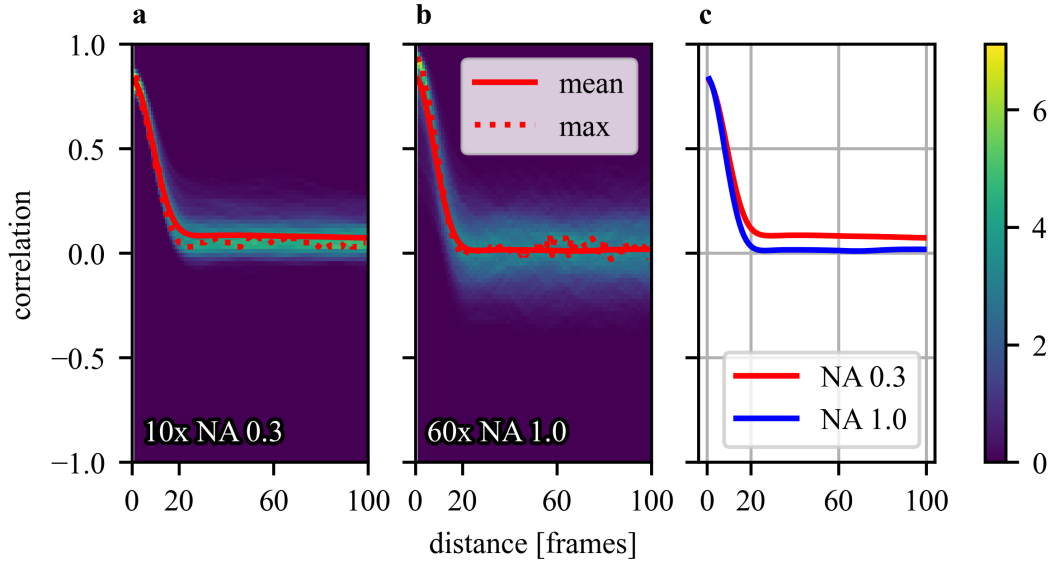
From the plot, we observe that correlation decreases as the distance increases, reaching a plateau at around 20 frames for both of the 0.3 NA and 1.0 Na case. It is also interesting to observe that for the higher resolution case (i.e. 1.0 NA), the cross-correlation goes closer to zero, as can be clearly seen from Fig S3c. Therefore, we conclude that using an deinterleaving of 20 frames would be enough for breaking the correlation between patterns.

#### A. Analysis of the mean and variance across multiple substacks

In this section, we compare the fluctuation properties captured in the mode pattern images acquired using the 0.3 NA and 1.0 NA objectives. We performed this analysis by getting  $V$  and  $A$  as defined in Section 2.F of this supplementary document for multiple substacks of the original image  $\mathbf{I}$  that depend on the deinterleaving  $d$  and also in the offset  $n$  as:

$$\mathbf{I}^{(n,d)} = \mathbf{I}_{:,n::d} \quad (\text{S6})$$





**Fig. S3.** Histogram of cross-correlation at different deinterleaving (measured as number of frames), with lines based on mean (weighted) and maximum values for each deinterleaving (across each column). The distance is measured in units of frames. Since each frame corresponds to a single position of the laser and the step size is 100 nm, this allows to easily convert frames to physical distance. **a.** NA 0.3 with 10x. **b.** NA 1.0 with 60x. **c** Mean curves only.

The double colon  $::$  is used to indicate a sequence of values starting at  $n$  with a step of  $d$  ( $n, n + d, n + 2d, \text{etc.}$ ). With this expression, we defined the following terms:

$$V(p, n, d) = \langle (\mathbf{I}_{p,:}^{(n,d)} - \langle \mathbf{I}_{p,:}^{(n,d)} \rangle) (\mathbf{I}_{p,:}^{(n,d)} - \langle \mathbf{I}_{p,:}^{(n,d)} \rangle) \rangle \quad (\text{S7})$$

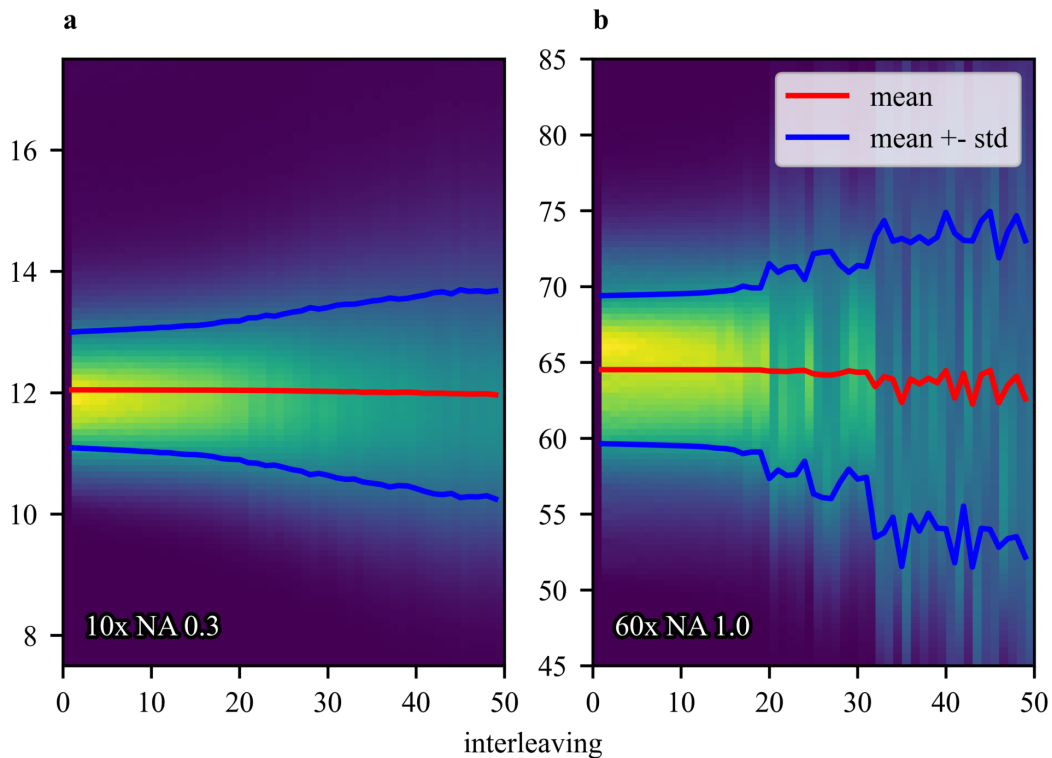
$$A(p, n, d) = \langle \mathbf{I}_{p,:}^{(n,d)} \rangle \quad (\text{S8})$$

From these expressions we get the ratio  $R(p, n, d) = V(p, n, d) / A(p, n, d)$  to observe variations in the fluctuations for every case. This expression allows to characterize every pixel at single stack in terms of its intensity values, which is normalized by the mean intensity in the stack. We note that instead of variance to mean ratio, standard deviation to mean ratio could have been used and would lead to similar interpretations. Since the interpretations are similar, we have chosen to present only variance to mean ratio.

As a result, for every choice of deinterleaving  $d$  we can obtain a distribution of  $R$  as there exist several substacks for each value. An important problem is that as the number  $d$  increases, so does it the number of substacks available (with each of them containing less frames as the total number of frames is limited).

To increase the number of samples, we extracted 10 small patches of 100 by 100 pixels from the original autofluorescence acquisition. The results can be visualized as living in a three-dimensional array whose axis correspond to  $d$ , substack index  $n$ , and pixel position  $p$ . Then, we compute  $R$  as an element-wise ratio and then compress the result to be represented as a function of  $d$ . Finally, we present the results in Figure S4 which shows the histogram across every value of  $d$ . In addition, we have added a line with the mean of the curve in red with an envelope showing the standard deviation for each  $d$ .

In this figure, it is possible to observe a change in the histograms and curves for deinterleaving values above 20 frames. For values below that, the standard deviation (blue curve) varies slowly suggesting that the variation of  $R$  (and therefore the variance across pixel intensities) does not change considerably in the range between 1 and 20. This represents the effect of picking different values of offset  $n$ . A large spread indicates large sensitivity to the offset, and vice versa. As  $d$  increases, indeed the number of candidate values for  $n$  increases, and indeed some amount of sensitivity to the value of  $n$  is expected. However, the fact that the spread becomes almost constant after a certain value of  $d$  indicates robustness to the value of  $n$  and therefore less critical need of controlling the offset during an experiment.



**Fig. S4.** Histogram of variance over mean with lines indicating the mean and standard deviation. **a.** NA 0.3 with 10x. **b.** NA 1.0 with 60x.

### B. Interpretation of eigenimages in each pixel window

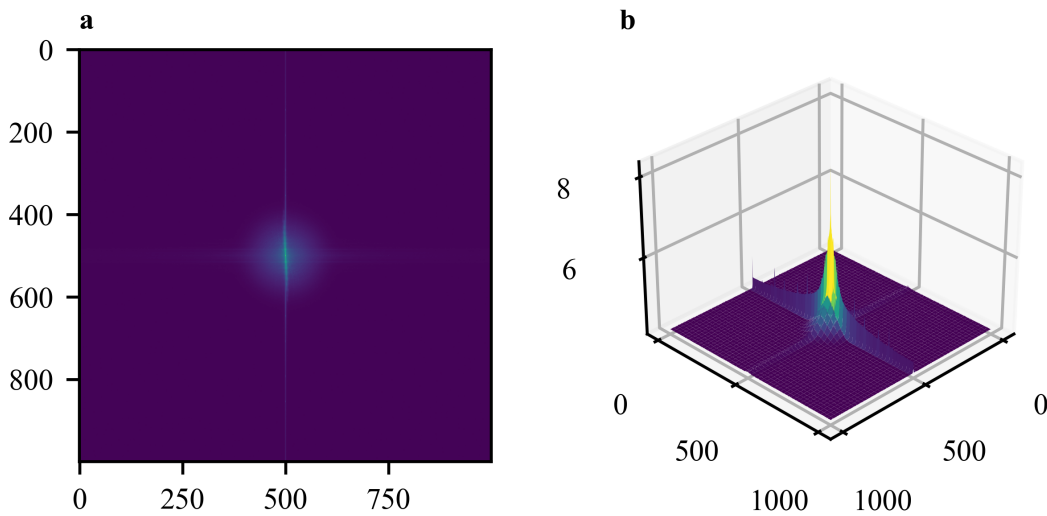
MUSICAL is based on the decomposition of small sub-sections into what is referred to as the eigenimages in related literature. In the original MUSICAL, where widefield illumination and a blinking behavior rooted in the fluorophores' photokinetics is expected, these eigenimages contain the most prominent spatial structures in the sample. Later, these are divided into two subsets called the signal subset and the noise subset depending on their statistical significance measured through their respective eigenvalues. On this new application of the MUSICAL algorithm, the fluorescence fluctuations is attributed largely to the illumination patterns generated in the chip. Therefore, in this new scheme, the eigenimages contain the different illumination modes (which encode the structural information) and the division into signal and noise subspaces is equivalent to selecting which modes represent statistically prominent illumination patterns within the pixel sub-window and which do not. This may be interpreted also as selecting fewer modes than pixels, namely the most relevant ones for super-resolution.

## 4. FOURIER ANALYSIS OF ILLUMINATION PATTERNS

The relevant waveguide mode patterns for this study, can in general be described as lines. While the patterns are random, due to the geometry of the waveguide and excitation laser, these tend to be horizontal or deviating with only a small angle compared to the waveguide optical axis. We applied a Fast Fourier transform to each frame in the stack in order to observe this effect. As the lines' position can be modulated by the phase in the Fourier domain, we are interested only in comparing the magnitudes. The results of the mean Fourier magnitude are shown in Figure S5.

The circular distribution shows that the illumination covers frequencies in both the horizontal and vertical axes. However, the strong vertical line indicates that the patterns are indeed emphasised in the horizontal direction. As both directions are not equalized, it is likely that resolution-limited structures along the vertical axis are better resolved than the ones along the horizontal axis.





**Fig. S5.** Fourier transform of illumination patterns in logarithmic scale. **a.** Mean magnitude of FFT, 2D view at 1.0 NA. **b.** Mean magnitude of FFT, 3D view at 1.0 NA

## 5. MUSICAL THRESHOLD SELECTION

The most difficult parameter to choose—and a potential source of subjectivity—for the MUSICAL reconstruction is the *threshold* parameter. This determines which fluctuations are regarded as signal and which are regarded as noise or background fluctuations.

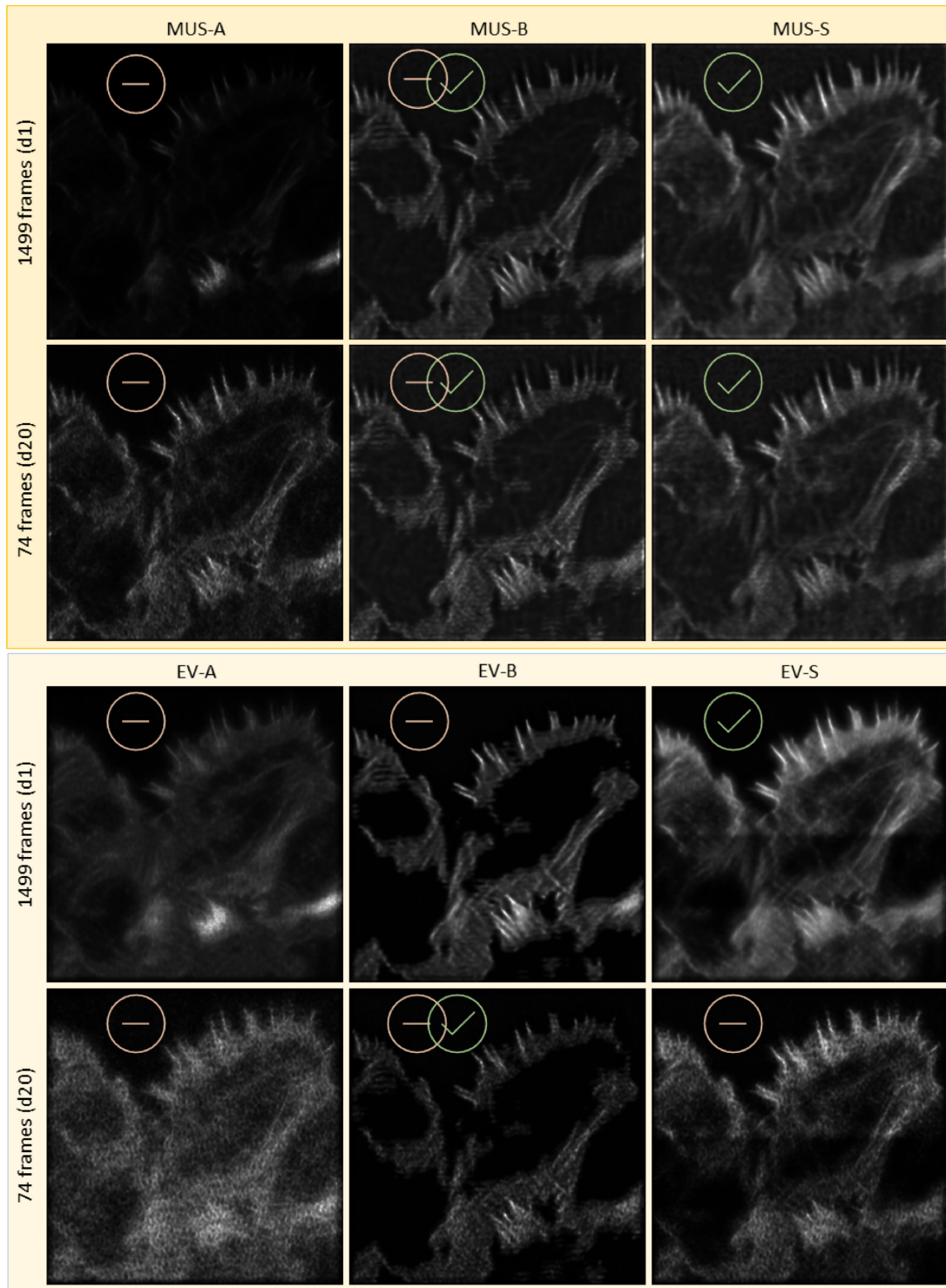
To overcome this challenge, various automatic thresholding schemes were developed by Acuña et al. in [4] and tested on both simulated and real microscopy data, but only for the case of homogeneous illumination and epifluorescence data. Here, we tested the different automatic thresholding schemes for the inhomogeneous multimodal waveguide (TIRF) illumination. We used the same deinterleaved image stacks as presented in the main article.

We found that for all cases of deinterleaving, the soft thresholding scheme MUS-S gave a satisfactory and the best performance compared to the other thresholding schemes. The reconstructions were in accordance with the 1.0 NA ground truth reference, low to no appearance of reconstruction artifacts and with an intensity scaling appearing in accordance with the actual sample brightness. The reconstruction procedures MUS-A, MUS-B and EV-B were seen to reject too many sample details for all cases of deinterleaving. The eigenvalue-scaled methods EV-A and EV-S appeared blurry and with a high presence of reconstruction artifacts for small stack sizes. Similar results as for the 0.3 NA data were also achieved for the 1.0 NA data, although the ground truth is not available for these images. The results for the 0.3 NA data are summarized in Figure S6.

## 6. EXTREME FOVS SUPER-RESOLUTION MICROSCOPY

For application scientists to efficiently harvest the advantages of the huge super-resolved areas provided by chip-based MUSICAL, the MUSICAL image computation must also be within reach of normal computers, e.g., via the open-source ImageJ plugin MusiJ [17]. Below we consider a practical scenario based on the reconstruction displayed in Figure S7.

Assuming that the image coming from a microscope is stored using 16 bits and with a grid of sensors of 2048 by 2048 elements, then each frame has a size of 8 MB. This means that a stack of 1500 frames requires approximately 12 GB in memory which is already larger than the common computer configuration of 8 GB. This can be solved partially by using ImageJ's virtual stacks where the images are not stored completely in the RAM but on the hard drive (disk resident). However, MusiJ requires the complete image stack to be loaded into the RAM in order to split the data across different processors in a way compatible with the MUSICAL computation. Moreover, we observed heap memory issues while processing larger datasets, obstructing the attainment of a best-quality MUSICAL image (processing all the available information) via the MusiJ plugin. We found the deinterleaved reconstruction strategy as a simple and elegant solution to this bottleneck of computational resources, namely a method for reconstructing all the available information



**Fig. S6. Selection of MUSICAL thresholding method.** The images show results from computing the MUSICAL images for different thresholding methods (by Acuña et al. [4]) on the 0.3 NA data as above. MUS-A and MUS-B are according to the original MUSICAL algorithm's hard thresholding scheme, while MUS-S and EV-S are modified MUSICALs employing automatic soft thresholding schemes. The EV-methods scale the contribution of every eigenimage (both signal and noise) by the inverse of their respective eigenvalues. On the other hand, the soft methods do not separate the noise and signal explicitly. They weigh every eigenimage both as signal and noise using two complementary functions that map eigenvalues to a real value in the range  $[0,1]$ . These functions are defined in terms of the highest and lowest second eigenvalue in the stack. For these particular c-TIRFM data, the soft thresholding method MUS-S proved the best for all deinterleaved stack sizes and data acquired via both the 0.3 NA and the 1.0 NA objectives.

which was also greatly sped-up compared to the previously applied "all-frames simultaneous" reconstruction strategy.

Figure S7 is an example of a large MUSICAL image resulting from such a deinterleaved processing with  $d=10$ , meaning a reduction of 10 times the required memory for processing. In terms of size, this sample contains originally 1500 frames of 2448 by 2048 pixels and requires approximately 14 GB in memory, impossible to process using MusiJ and standard laptop configurations. The deinterleaved stack of 150 frames (1.4 GB) on the other hand, is possible to process. Hence, the deinterleaved reconstruction approach with sum of sub-stack MUSICAL images outlined in the main manuscript, enables the obtainment of larger MUSICAL images of superior quality than what would have been obtainable without the deinterleaved approach.

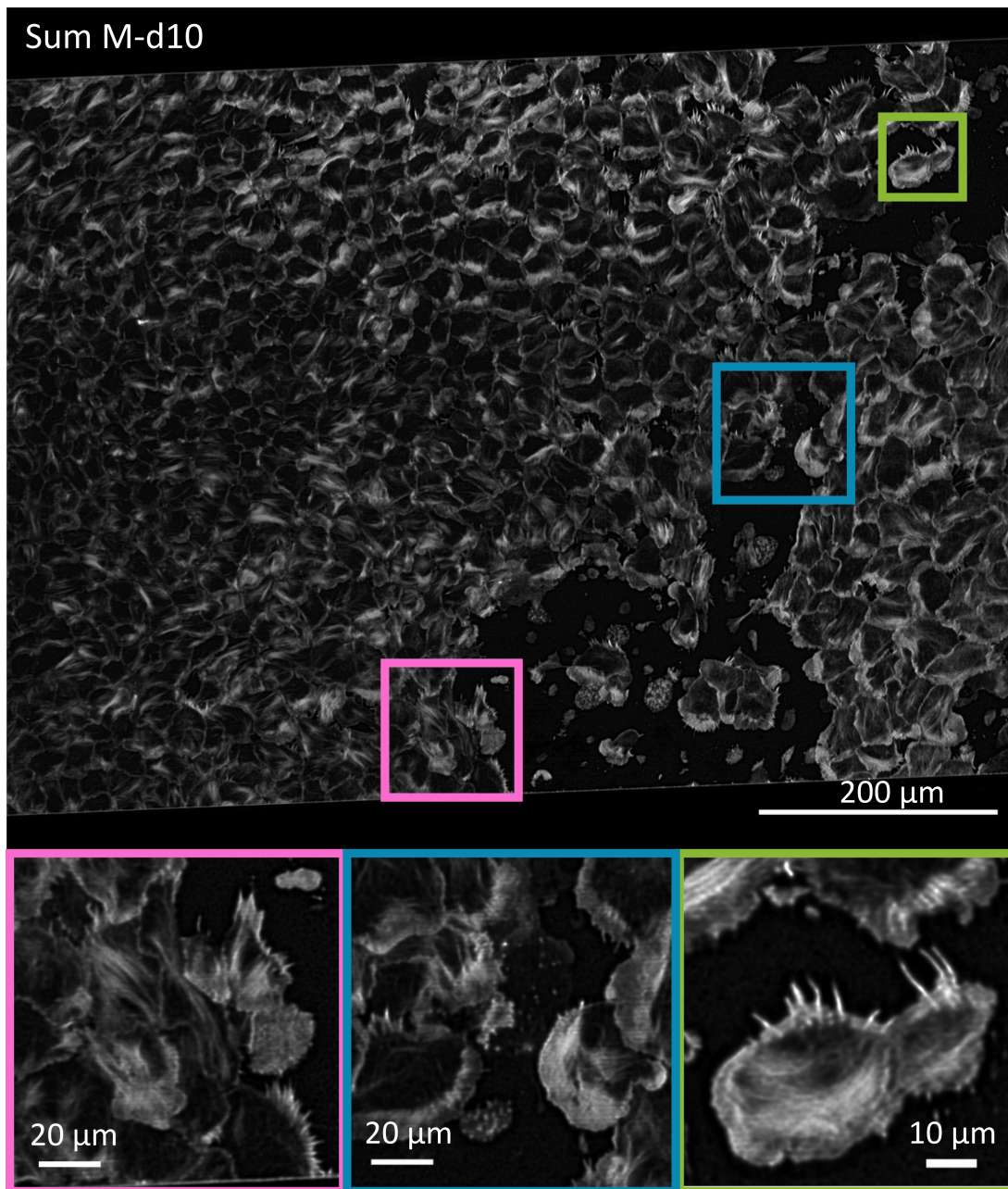
The MUSICAL computations presented in this work were achieved via a Python implementation of MUS-S [4]. The computer used was equipped with an Intel Xeon Gold 5118 processor with 12 cores and 128 GB DDR4 memory. For this configuration (and splitting the task into 8 parallel processes), the particular processing of the large FOV displayed in Figure S7 took 12 min 44 s for 150 frames, and 23 min 32 s for 1500 frames. Thus, the main advantage of the deinterleaved reconstruction approach is not saving reconstruction time (although this greatly depends on the particular hardware and software implementation), but to enable the computations to be performed on more memory-restricted systems.

## 7. COMPARISON OF MUSICAL-ON-CHIP WITH BSOFI AND SRRF

Although the multimodal chip illumination is expected to cause reconstruction artifact for other fluctuation imaging methods than MUSICAL, we have for completeness included a comparison of the MUSICAL results with the corresponding SRRF and bSOFI images in Figure S8.

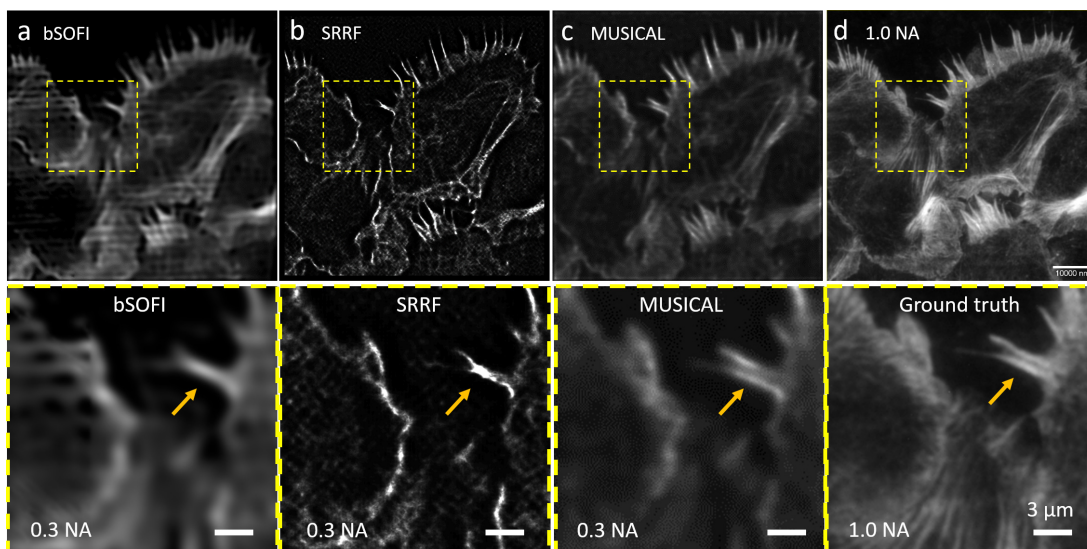
## 8. LIVE-CELL IMAGING

As the photonic chips are bio-compatible, the gentle TIRF illumination can also be exploited for live-cell imaging of a great variety of cell types. Figure S9 demonstrates imaging (conventional and MUSICAL) of actin in living cardiomyoblasts (labelled using SiR-Actin). The color coded panel (projection of three MUSICAL images at different time-points) highlights minute cytoskeletal changes over time. The optimization for high-speed super-resolution imaging of living cells is the scope of future work.



**Fig. S7. Super-resolution microscopy over extreme fields of view.** The upper panel show the same keratocyte sample and area as in Figure 2a, but instead of computing the mean of modes for a conventional TIRF image, the resolution is enhanced by computing the MUSICAL image. To speed-up and ease the computational burden, the SUM of 10 MUSICAL images from the deinterleaved image stack ( $d=10$ ) was used instead of all frames simultaneously. The super-resolved area is  $507\,000\ \mu\text{m}^2 \approx 0.51\ \text{mm}^2$  (here limited vertically by the  $600\ \mu\text{m}$  wide waveguide, and horizontally by the camera FOV, about  $845\ \mu\text{m}$ ). The color coded ROIs indicated in the upper panel are displayed magnified below. The large super-resolved image area is ideal for capturing the great variety of morphological features exhibited by keratocyte skin cells such as their extruding pseudopods.

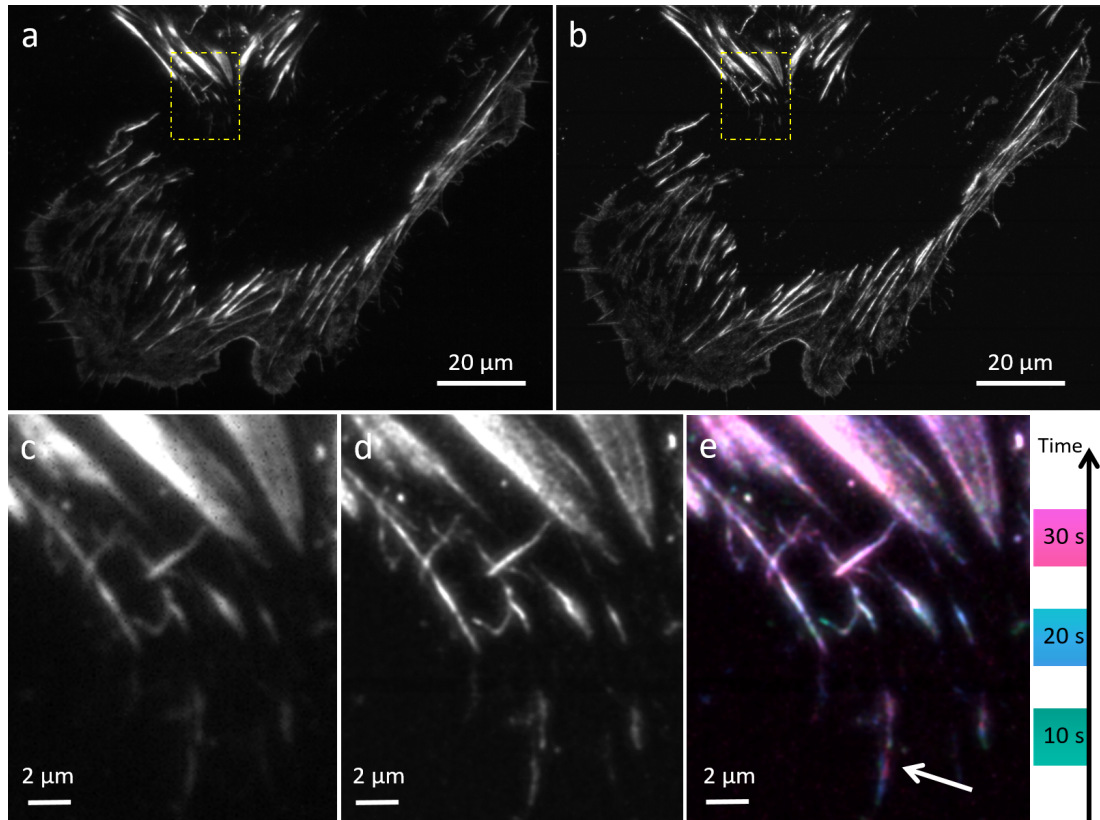




**Fig. S8. Comparison of bSOFI, SRRF and MUSICAL.** The ROIs indicated in the upper row are displayed magnified below. As expected from the theoretical considerations presented in this work, MUSICAL achieves a far better image reconstruction using the chip illumination than bSOFI and SRRF. This can be seen from the comparison with the ground truth image (panel d), especially clearly where indicated by the arrows. The bSOFI, MUSICAL and Ground truth (sum) images were generated using 1499 frames. The SRRF image was generated using 1400 frames (The software could not reconstruct 1499), Temporal radially auto-correlation of order 2, radius 0.5 and other default parameters. The bottom row scale bars are 3  $\mu\text{m}$ .

## REFERENCES

1. A. Z. Subramanian, G. S. Murugan, M. N. Zervas, and J. S. Wilkinson, "High index contrast: Ta<sub>2</sub>O<sub>5</sub> waveguide amplifier on oxidised silicon," *Opt. Commun.* **285**, 124–127 (2012).
2. C. Chaneliere, J. Autran, R. Devine, and B. Balland, "Tantalum pentoxide (ta<sub>2</sub>o<sub>5</sub>) thin films for advanced dielectric applications," *Mater. Sci. Eng. R: Reports* **22**, 269–322 (1998).
3. J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid *et al.*, "Fiji: an open-source platform for biological-image analysis," *Nat. methods* **9**, 676–682 (2012).
4. S. Acuña, I. S. Opstad, F. Godtlielsen, B. S. Ahluwalia, and K. Agarwal, "Soft thresholding schemes for multiple signal classification algorithm," *Opt. Express* **28**, 34434–34449 (2020).
5. A. C. Descloux, K. S. Grusmayer, and A. Radenovic, "Parameter-free image resolution estimation based on decorrelation analysis," *Nat. Methods* **16**, 918–924 (2019).
6. Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE transactions on image processing* **13**, 600–612 (2004).
7. R. Diekmann, Ø. I. Helle, C. I. Øie, P. McCourt, T. R. Huser, M. Schüttelz, and B. S. Ahluwalia, "Chip-based wide field-of-view nanoscopy," *Nat. Photonics* **11**, 322 (2017).
8. Ø. I. Helle, D. A. Coucheron, J.-C. Tinguely, C. I. Øie, and B. S. Ahluwalia, "Nanoscopy on-a-chip: super-resolution imaging on the millimeter scale," *Opt. express* **27**, 6700–6710 (2019).
9. N. Jayakumar, Øystein I Helle, K. Agarwal, and B. S. Ahluwalia, "On-chip tirf nanoscopy by applying haar wavelet kernel analysis on intensity fluctuations induced by chip illumination," (2020).
10. K. Agarwal and R. Macháň, "Multiple signal classification algorithm for super-resolution fluorescence microscopy," *Nat. communications* **7**, 1–9 (2016).
11. S. Jadhav, S. Acuña, I. S. Opstad, B. S. Ahluwalia, K. Agarwal, and D. K. Prasad, "Artefact removal in ground truth deficient fluctuations-based nanoscopy images using deep learning," *Biomed. optics express* **12**, 191–210 (2021).
12. E. Mudry, K. Belkebir, J. Girard, J. Savatier, E. Le Moal, C. Nicoletti, M. Allain, and A. Sentenac, "Structured illumination microscopy using unknown speckle patterns," *Nat. Photonics* **6**, 312–315 (2012).



**Fig. S9. Live-cell imaging of cardiomyoblasts.** c-TIRFM of actin in living cardiomyoblasts. a) The sum image of the first 150 frames (mode patterns) used for reconstructing the MUSICAL image in (b). The indicated ROIs in the yellow squares are shown magnified below in panels (c) sum and (d) MUSICAL image. (e) Color coded projection of MUSICAL images from time-point 1 (frames 1-150), time-point 2 (151-300), and time-point 3 (301-450). The arrow indicate dynamics of an extending actin filament. The image collection was done using a 1.2 NA water immersion objective. The image acquisition time was about 10 s per 150 frames used for 1 MUSICAL image.

13. L.-H. Yeh, S. Chowdhury, N. A. Repina, and L. Waller, "Speckle-structured illumination for 3d phase and fluorescence computational microscopy," *Biomed. Opt. Express* **10**, 3635–3653 (2019).
14. T. Mangeat, S. Labouesse, M. Allain, A. Negash, E. Martin, A. Guérolé, R. Poincloux, C. Estibal, A. Bouissou, S. Cantaloube *et al.*, "Super-resolved live-cell imaging using random illumination microscopy," *Cell Reports Methods* p. 100009 (2021).
15. Ø. I. Helle, F. T. Dullo, M. Lahrberg, J.-C. Tinguely, O. G. Hellesø, and B. S. Ahluwalia, "Structured illumination microscopy using a photonic chip," *Nat. Photonics* **14**, 431–438 (2020).
16. K. Agarwal and D. K. Prasad, "Eigen-analysis reveals components supporting super-resolution imaging of blinking fluorophores," *Sci. reports* **7**, 1–10 (2017).
17. S. Acuña, F. Ströhl, I. S. Opstad, B. S. Ahluwalia, and K. Agarwal, "Musij: an imagej plugin for video nanoscopy," *Biomed. Opt. Express* **11**, 2548–2559 (2020).





# Paper V

Scalable-resolution structured illumination microscopy





# Scalable-resolution structured illumination microscopy

ANKIT BUTOLA,<sup>1,2,\*</sup>  SEBASTIAN ACUNA,<sup>1,2</sup>  DANIEL HENRY HANSEN,<sup>1</sup> AND KRISHNA AGARWAL<sup>1</sup>

<sup>1</sup>*Department of Physics and Technology, UiT The Arctic University of Norway, Norway*

<sup>2</sup>*Co-first authors*

\*[ankitbutola321@gmail.com](mailto:ankitbutola321@gmail.com)

**Abstract:** Structured illumination microscopy suffers from the need of sophisticated instrumentation and precise calibration. This makes structured illumination microscopes costly and skill-dependent. We present a novel approach to realize super-resolution structured illumination microscopy using an alignment non-critical illumination system and a reconstruction algorithm that does not need illumination information. The optical system is designed to encode higher order frequency components of the specimen by projecting PSF-modulated binary patterns for illuminating the sample plane, which do not have clean Fourier peaks conventionally used in structured illumination microscopy. These patterns fold high frequency content of sample into the measurements in an obfuscated manner, which are de-obfuscated using multiple signal classification algorithm. This algorithm eliminates the need of clean peaks in illumination and the knowledge of illumination patterns, which makes instrumentation simple and flexible for use with a variety of microscope objective lenses. We present a variety of experimental results on beads and cell samples to demonstrate resolution enhancement by a factor of 2.6 to 3.4 times, which is better than the enhancement supported by the conventional linear structure illumination microscopy where the same objective lens is used for structured illumination as well as collection of light. We show that the same system can be used in SIM configuration with different collection objective lenses without any careful re-calibration or realignment, thereby supporting a range of resolutions with the same system.

© 2022 Optica Publishing Group under the terms of the [Optica Open Access Publishing Agreement](#)

## 1. Introduction

The past few decades have witnessed widespread application of fluorescence microscopy in biology by providing the unique ability of intrinsic selectivity, chemical specificity and signal to background ratio. In the recent times, technical advancement in fluorescence microscopy are primarily focused on simultaneous imaging of different cellular/sub-cellular components, three dimensional imaging of thick samples such as organoids, and improving spatial resolution of imaging [1–3]. In particular, several methods have been proposed in the past to break the classical resolution limit. These methods can be broadly categorized into the following groups: a) stimulated emission and depletion (STED) [4,5], b) single molecule localization (SML) i.e., rely on the fact localization precision [6–9], c) intensity fluctuations based optical nanoscopy (IFON) [10–14], and d) structured illumination microscopy (SIM) [15–18]. STED, SML and IFON all rely on photokinetics of fluorophores, i.e. the variation in the photon emission statistics either naturally (IFON), or using long-lived dark states (IFON, SLM), or trigger of different photophysical behaviour in presence of specific laser lights of high intensity (STED, SML). Consequently, strict control, special dyes and/or special imaging protocols are needed for these techniques. This restricts versatility of the methods. Furthermore, the temporal throughput of SML and IFON is quite low due to the need of adjusting the acquisition rate and the number of frames to the photokinetics. Temporal throughput of STED is limited due to point scanning based illumination and acquisition. These restrictions are absent in SIM, making it more versatile

and preferred in wide scenarios, despite lower factor of resolution enhancement than the other techniques.

Super resolved imaging in SIM is achieved by illuminating the sample by structured patterns and computationally combining the high frequency information conventionally in the Fourier space. Structured illumination extends the diffraction limit by an amount equal to the spatial frequency of the illumination pattern. Since the highest spatial frequency of the illumination pattern is diffraction limited, the best enhancement of lateral resolution in SIM is by a factor of about two [19]. Two-factor resolution enhancement is valid in the case of linear SIM i.e., when emission intensity is proportional to the excitation wavelength. Non-linear SIM can provide further resolution enhancement [18–20], but it uses extremely high intensity to saturate the excited state of molecules, which results in non-linear relation between the emission rate per fluorophore and the illumination intensity.

Another limitation of the conventional SIM-based system is the use of same objective lens for creating structured illumination and also for collecting the light from the sample. If one intends to support SIM with another objective lens of a different numerical aperture, recalibration and realignment of the system become imperative. This not only increases the demand on cost, but mandates complicated procedure during the process of imaging. The central reason for this limitation is that precise knowledge of the illumination patterns, including the pattern shift and rotation, is crucial for the reconstruction algorithm of SIM.

Our goal is to circumvent the limitation of fixed resolution of SIM system and the requirement of precise knowledge of illumination patterns to achieve super-resolution in SIM. We also intend to reduce the sensitivity of SIM reconstruction to the signal to background ratio. It has been reported that SIM reconstruction algorithms pose certain stringent requirements on the signal quality [21]. As explained in section 2, we achieve our goals by using (a) using non-saturated high frequency illumination pattern achieved within linear optical regime and (b) reconstructing the image using non-linear reconstruction algorithm that does not require illumination to be known a priori. We show in-principle support for resolution enhancement better than the linear structured illumination by decoupling the illumination and detection path and using high NA objective lens for illumination. Experimental results with multiple collection objective lens are shown for validation and wider applicability of the proposed system. We show good reconstructions at signal to background ratio as small as  $<3$ .

The motivation of using different collection objective lenses also comes from the potential applications. The dimensions of structures inside cells vary greatly. For example, cytoskeleton on its largest scale of visualization and function spans the entire cell (10s of  $\mu\text{m}$ ) and requires visualization of  $<200$  nm at its smallest scale. In addition, upgrading the system with switchable laser source and use the best super-resolution for actin and microtubules ( $<200$  nm), the next to best for mitochondria (0.5- 3  $\mu\text{m}$ ), and the subsequent one for the cell membrane of cytoskeleton scale. Further, operating in linear regime without the knowledge of illumination simplifies the optical system design and implied relatively inexpensive, easy-to-use instrumentation. We show that the same system can achieve scalable resolution by using the same illumination objective lens, but employing different collection objective lenses. As a consequence, our framework opens up a new possibility for structured illumination based super-resolution techniques.

## 2. Principle and deviation from the conventional SIM

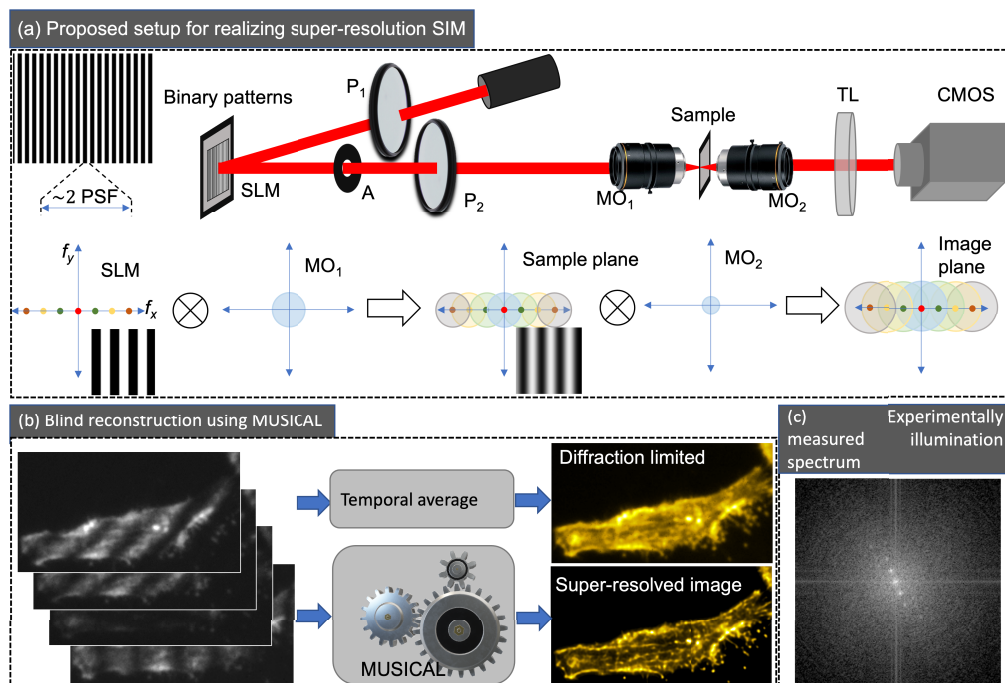
The essence of SIM lies in being able to introduce precisely known clean high frequency peaks in the illumination pattern in order to (a) down-modulate the high frequency components of the specimens into the collection optical transfer function (OTF) in a systematic manner, and (b) design a simple reconstruction algorithm that solves a linear system of equations derived from the prior knowledge of the illumination pattern. In order to circumvent the need of structured

illumination to achieve super-resolution, both the conditions above have to be relaxed by removing the need of precise peaks corresponding to sinusoidal patterns.

The present article proposes a unique solution to the above problem. We develop a simplified structured illumination system by decoupling illumination and detection arm and employ a reconstruction approach that does not assume sinusoidal illumination patterns. The fundamental empowering component of our approach is to deviate from the conventional SIM and NL-SIM in terms of not needing clean peaks corresponding to sinusoidal illumination patterns.

Our solution comprises of two complementary components shown in Fig. 1(a,b):

- Easy to align instrumentation with structured illumination pattern:** We present a simple and efficient optical system working in linear regime to induce structured illumination and thereby down-modulate the high frequency components of the specimens into the imaging objective lens's OTF. The presented framework may employ, in principle, random digital patterns or encoded patterns (such as Hadamard, etc.) with translations and rotations regardless of the final resolution desirable, making our method more robust and flexible. For simplicity, we use binary periodic patterns of sufficiently high frequency and project them onto the sample through an illumination objective lens (see Fig. 1(a) and Supplementary Fig. S1). Consequently, the higher frequency components of the binary patterns are blurred by the OTF of the illumination objective lens at the sample plane. The blurred patterns therefore do not have clean peaks corresponding to higher frequency components of the binary pattern needed for conventional SIM reconstruction (see Fig. 1(c)). Nonetheless, problems associated with high illumination intensity are avoided.



**Fig. 1. Overview of proposed imaging and reconstruction technique.** (a) Binary patterns are projected into the sample using spatial light modulator to encode high spatial frequency components. (b) Raw datasets with different illumination orientation processed using Multiple signal classification algorithm (MUSICAL). (c) experimentally measured higher order frequencies using binary pattern

- **Illumination blind reconstruction algorithm:** A non-linear reconstruction algorithm is used which does not require the prior knowledge of illumination patterns, nor needs to estimate the illuminations as an intermediary step. For this purpose, we use the principle of pseudospectral decomposition which isolates the spatial frequency components of the illuminated samples that can be well separated from noise for reconstruction. This alleviates the need of knowing or estimating individual illumination patterns. For simplicity of adaptation, we use multiple signal classification algorithm (MUSICAL [14]), which is based on pseudospectral decomposition and has been adapted as a reconstruction approach recently [22] (see Fig. 1(b)). We note that other blind reconstruction algorithms may be employed, and special illumination priors may be incorporated for example when using Hadamard or other computationally exploitable patterns.

### 3. Methods

#### 3.1. Experimental setup

The experimental setup and illumination strategy is shown in Fig. 1. A collimated laser light beam passes through a polarizer ( $P_1$ ). The polarized light is modulated using spatial light modulator (SLM), which has a binary pattern supporting high spatial frequencies. This pattern is projected on to the sample through the cross polarizer and illumination objective lens  $MO_1$ . Therefore, the pattern on the sample is the clean frequency peaks of the original digital pattern convolved with  $MO_1$ 's OTF. The sample is placed in a plane somewhere between the 'f' and '2f' planes of the illumination objective lens  $MO_1$ . The blurring due to  $MO_1$  and the imperfect projection of the pattern implies that only a few peaks that fall within the bandwidth of  $MO_1$  reach the sample. However, this does not pose a problem because we do not need to know or estimate the illumination pattern in our reconstruction.

The design of the illumination pattern is discussed in section 3.2. The sample, labeled with fluorescent dyes, emits photons in response to these illuminations. These photons are collected by imaging optics through imaging microscope objective  $MO_2$  and matching tube lens (TL) on an image sensor array. In order to fully automate the data acquisition and synchronize the SLM pattern projection and camera acquisition, we designed an image acquisition code in labVIEW. We used Cobolt laser of wavelength 660 nm, SLM Pluto-2.1 from Holoeye, 60X, 0.9NA (RMS60X-PFC, Olympus) objective lens for  $MO_1$ , different options for  $MO_2$ , tube lens matching the  $MO_2$ , and ORCA-Fusion C14440-20UP sCMOS camera for the image sensor array. In addition, polarization control is maintained by two polarizers in the illumination path and SLM.

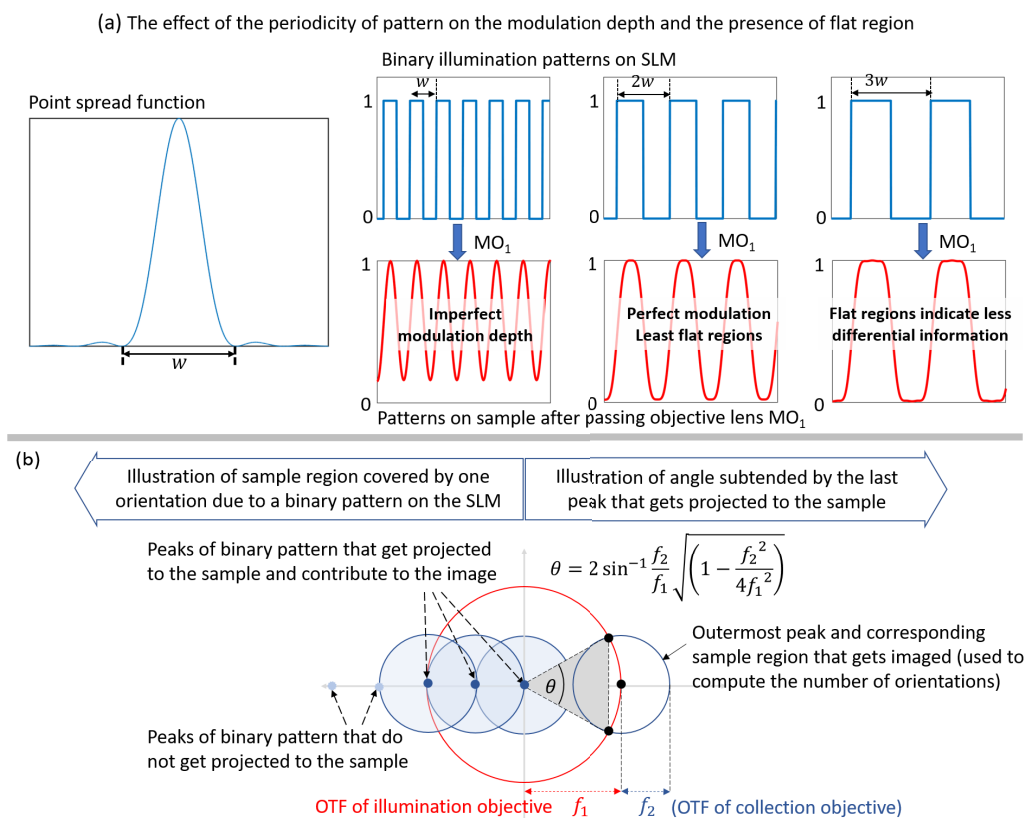
Our motivation to use SLM for illumination modulation was two-fold. First, SLM is an electronic device which implies illumination patterns can be changed without mechanical movements. Second, the pixel size of the SLM was 8  $\mu\text{m}$ , which together with the  $MO_1$  and  $MO_2$  gives sufficient sampling of the point spread function (PSF) of the  $MO_2$ . However, we note that other digital adaptive optical devices (such as digital micromirrors) or analog binary phase or amplitude masks could have been used instead of SLM without the loss of generality.

Further, we used  $MO_1$  and  $MO_2$  to be two distinct microscope objective lenses in transmission setup in order to investigate the effect of the illumination objective lens independent of the collection, and use different combinations for the objective lens for understanding the advantage accrued by the high spatial frequencies of illumination.

#### 3.2. Selection of illumination patterns

The selection of the illumination patterns, in our case, is treated as a geometric and signal processing problem.

**Periodicity** Given the numerical aperture (NA) of  $MO_2$ , twice the null-to-null width ( $w$ ) of the PSF is used as the period of the binary pattern. The duty cycle of the binary pattern is 50%. The choice of  $3\times$  or  $1\times$  the width of the PSF results in deteriorated results. The reason is explained through an illustration in Fig. 2(a). Consider an Airy point spread function with null-to-null width  $w$ . Convolution with a binary pattern of periodicity  $2w$  results in the best modulation depth together with minimum flat components of maximum and zero intensity. Having larger flat components implies less level of fluctuations in two consecutive pattern shifted images. The advantage of using larger period is also manifest in better robustness to noise, this is illustrated in Supplementary Fig. S2. Indeed, one could use smaller period (say only of size ( $w$ )), which results in half the number of images per orientation, and therefore better temporal resolution. However, the tradeoff is threefold: (a) poorer modulation depth, therefore (b) lesser fluctuations, and (c) poorer robustness to noise. More numerical experiments with different periodicities are presented in Supplementary Note S1 and Supplementary Fig. S4.



**Fig. 2. The selection of parameters of the binary pattern projected on  $MO_1$  before illuminating the sample.** (a) Selecting periodicity other than  $\sim 2w$  leads to either poor modulation depth or flat regions of maximum or no illumination, both of which reduce differential information between two consecutive translated illumination patterns. (b) The illumination and collection objective lenses determine what region in Fourier space of the sample gets imaged and how many orientations are needed for for imaging the sample without having shadow regions in the Fourier domain. (c) For a more general scenario, the scheme in (b) presents some limitations.

**Pattern shifts (i.e. translations)** In order to have uniform illumination for a given orientation, we introduce shifts in patterns by one pixel of SLM at a time for the entire period of the binary pattern. Uniformity of illumination can be achieved in principle using just two patterns, namely a binary pattern and its Boolean inverse, However, shifting one pixel at a time allows for introducing



the maximum diversity of spatial fluctuations at every point in sample. Further study on the pattern shift is presented in Supplementary Note S2 and Supplementary Fig. S4.

**Orientations** Given the desired order of resolution enhancement  $K$  over FWHM ( $\sim w/2$ ) and the effective size of OTF (proportional to  $1/(2w)$ ) of  $\text{MO}_2$ , we compute the number of orientations such that there is no shadow region within the desired bandwidth. This is illustrated in Fig. 2(b). It is simply specified using the smallest integer exceeding  $\pi/\theta$ , where  $\theta = \sin^{-1}(2/K)$ . However, this assumes that period  $2w$  was used. Attempting to incorporate a more general scenario, such as with a different selection of period of patterns and/or  $f_1$  including the effect of noise presents challenges. We therefore include a more sophisticated scheme of coverage of desired spectrum in Supplementary Note S3 and Supplementary Fig. S5. We note that the analyses discussed here pertain to the minimum number of orientations needed. We present a discussion regarding using more orientations than necessary in Supplementary Fig. S6 and Supplementary Note S4.

**The parameters derived from the above considerations** Given the 660 nm laser light and the  $\text{MO}_2$  i.e., 10X (0.25 NA, RMS10X Olympus), 20X (0.4 NA, RMS20X Olympus) and 40X (0.65 NA, RMS40X Olympus) the null-to-null width  $w \approx \lambda/\text{NA}$  is  $\sim 2640$ ,  $\sim 1650$  and  $\sim 1015$  nm respectively. Further, the pixel size of SLM ( $8 \mu\text{m}$ ) combined with 60X of  $\text{MO}_1$  illumination implies that the  $w$  in SLM pixels is approximately 20, 13 and 8 pixels. Thereby the periodicity of the patterns is 40, 26 and 16 pixels, the number of pattern shifts is 40, 26 and 16 per orientation. Further, we used 90 orientations, where two consecutive orientations differed by  $2^\circ$ . This covered the entire  $360^\circ$  since each orientation is symmetric along the origin. The oversampling in the orientation is intentional so that the best achieved resolution is limited only by noise in the images. Further, we optimized the number of pattern required to achieve super-resolution imaging for different objective lens in the Supplementary Table S1.

### 3.3. Blind reconstruction approach for super-resolution

We use MUSICAL for blind reconstruction of the sample. While MUSICAL was originally envisioned for widefield illumination, its capabilities allows its use also for structured illumination microscopy. This was proved already in [22] where MUSICAL was used with images where the sample was illuminated using a lattice projecting structured illumination on it. A nice feature of the algorithm is that every illumination is inherently used without a precise knowledge of it. The key feature of illumination patterns exploited in MUSICAL is that fluctuations at the same point in the sample region occur in different time frames from the variety of illumination patterns. This implies that no illumination estimation algorithm is needed as a pre-processing or tandem processing step. As a consequence, not only the entire computation is rather simple, also the sensitivity of the reconstruction to errors in illumination estimation is not a concern.

**MUSICAL's algorithm** The algorithm is a combination of Singular Value Decomposition (SVD) and the knowledge of the PSF of the system together with the mathematically linear model of imaging discussed in the supplementary document of the original article [14]. Let's consider a sequence of images taken for a particular sample stained with fluorescent molecules. The intensity emitted by these molecules is a random variable and they can be considered as independent sources or emitters. If this sample is imaged in time, we can obtain a sequence of images which we then call an image stack. If every image in the stack is flattened to be represented as a column vector, then the entire stack, can be expressed as single rectangular matrix  $I$  whose columns correspond to each single frame in the stack. This allows its decomposition using SVD as  $I = USV$  with  $U$  and  $V$  being square matrices and  $S$  a diagonal matrix. Due to the structure of  $I$ , we can conclude that the information of its columns is encoded in  $U$ . Further,  $S$  contains in its diagonal the singular values of the matrix associated to the respective columns of  $U$ . We refer to these as the eigenvalues and the eigenimages of the stack. Then, the columns of  $U$ , or eigenimages, form a basis of the image stack originally acquired. Among others properties, these vectors are orthogonal and ordered by their statistical importance in



terms of how much information they contain with respect to the particular sample. On the other hand, the stack can be modeled as a convolution of the original emitters distribution  $E(\mathbf{r})$  and the point-spread-function of the optical  $G(\mathbf{r})$ , where both terms are considered as scalar functions and  $\mathbf{r}$  is a two-dimensional vector containing the coordinates in the imaging plane. Then  $I(\mathbf{r}) = G(\mathbf{r}) * E(\mathbf{r})$  with  $*$  meaning convolution. Additionally, if the sample is formed by  $K$  different emitters, then its underlying structure can be expressed as a discrete sum of  $K$  impulses centered in the location of the emitters. This is  $E(\mathbf{r}) = \sum_{k=1}^K b_k \delta(\mathbf{r} - \mathbf{r}_k)$ . We can then use the properties of the impulse function and express the image as  $I(\mathbf{r}) = \sum_{k=1}^K b_k G(\mathbf{r} - \mathbf{r}_k)$ . Here,  $b_k$  is the intensity of each individual emitter. Furthermore, in the discrete case this can be more easily expressed as the matrix multiplication  $\mathbf{I} = \mathbf{G}\mathbf{E}$  where we have use bold letters to indicate column vectors. In this expression,  $\mathbf{I}$  corresponds to a single image whose values correspond to the intensity measured at different sensor or pixels.  $\mathbf{G}$  is the matrix of  $K$  columns containing the point-spread-function of the system but shifted at the location of the different emitters. Finally,  $\mathbf{E}$  correspond to the values of  $b_k$ . If  $b_k$  were to vary in time, then we can acquire a stack as defined previously. In this case, the stack can also be expressed as a matrix multiplication as  $\mathbf{I} = \mathbf{G}\mathbf{E}$ . From this, it is possible to conclude that the columns of  $\mathbf{G}$  span the entire space of images regardless of the values of  $\mathbf{E}$  which can be considered as random. If the same stack were to be decomposed using SVD as explained before, then the space spanned by  $\mathbf{G}$  would be the same than the one spanned by  $\mathbf{U}$ . If the number of pixels is  $P$ , then  $\mathbf{U}$  has  $P$  columns. The assumption in MUSICAL is that the rank, or the number of non-zero eigenvalues, of  $\mathbf{I}$  is less than  $P$ . If the rank is  $S$ , then it means that the image can be formed only from the span of the  $S$  first eigenimages. This is called the signal subspace. On the other hand, the remaining eigenimages are considered as forming the noise subspace. Due to the orthogonality of  $\mathbf{U}$ , both spaces are orthogonal. However, images are affected by noise and in turn, real images are full-ranked which makes separation non-trivial. In order to deal with this problem, MUSICAL separates the space using the eigenvalues and a threshold given by the user. The eigenimages with eigenvalues above such threshold are considered as the signal, and the one below, noise.

After such separation is made, we can exploit the fact that the signal space spanned by the eigenimages is the same than the space spanned by the columns of  $\mathbf{G}$ . As a consequence, the columns of  $\mathbf{G}$  are orthogonal to the noise space and therefore its projection on the noise space is zero. Based on the projections this can be used to define an indicator function which allows to test an arbitrary point by comparing the point-spread-function centered at that point, and the eigenimages obtained from the actual image. The indicator function is expressed below in Eq. (1).

$$f(\mathbf{r}_t) = \left( \frac{\sum_{i \in \text{Signal}} |\mathbf{u}_i \cdot \mathbf{g}_t|^2}{\sum_{i \in \text{Noise}} |\mathbf{u}_i \cdot \mathbf{g}_t|^2} \right)^\alpha \quad (1)$$

Here,  $\mathbf{u}_i$  corresponds to the columns of  $\mathbf{U}$  indexed by the index  $i$ , and  $\mathbf{g}_t$  is the image of a single test emitter indexed by  $t$  given by the shifted PSF of the system. The indicator function assigns a numerical value at points in the sample, where this numerical value indicates the presence of fluorophore. In practice, these points are located on a grid finer than the one given by original pixels, reducing the pixel size of the resulting image when compared to the original image. The factor  $\alpha$  is a contrast enhancement factor whose value is based on heuristics.

In terms of processing, MUSICAL is carried out using a moving window that process a small patch at the time, instead of the entire field of view. The patch corresponds to a section of the size of the main lobe of the point-spread-function in the lateral coordinates, but including all the temporal information. These patches overlap and their corresponding super-resolved images are then merged by averaging in order to obtain a MUSICAL image of the entire field of view. The total number of patches correspond to the total number of pixels in the lateral plane.

A final note on the indicator function is presented here. This can be modified to combine signal and noise using a weighting function [23]. This is called soft MUSICAL or MUSICAL-S and allows leave the threshold out of the indicator function. The weighting is a function of the eigenvalues and allows to include partially information of noise into the numerator of the indicator function, and signal into the denominator.

**Illumination patterns and MUSICAL** In the conventional form of MUSICAL, the fluctuations in fluorescence emissions from emitters occurs as the temporal information encoded in  $V$  and the spatial distribution of the emitters gets encoded in the signal space of  $U$ . However, in the application presented here, each frame corresponds to a different illumination pattern. Therefore, the variation of illumination intensities at pixels over time is encoded in  $V$  and the ensemble of spatial distributions of the illumination patterns is encoded in the signal space of  $U$  together with the sample distribution. The mathematical derivations related to the form in which the illumination patterns are encoded in  $U$  were presented in the supplementary information of [24] and [22]. The relevant portions of the derivations are reproduced in our Supplementary Note S5. Here, the derivations of [22] directly apply and thereby the maximum possible resolution supported by MUSICAL (in the case of best selection of algorithmic parameters) is given by  $2(k_{\max} + k_{\text{MO}_2})$ , where  $k_{\max}$  is the maximum frequency supported by the illumination patterns and  $k_{\text{MO}_2}$  is the maximum frequency supported by the collective objective lens. It is however worth noting in our case that  $k_{\max}$  depends upon the noise present in the image as the noise levels determine the farthest peak in the illumination spectrum that is separable from the noise floor.

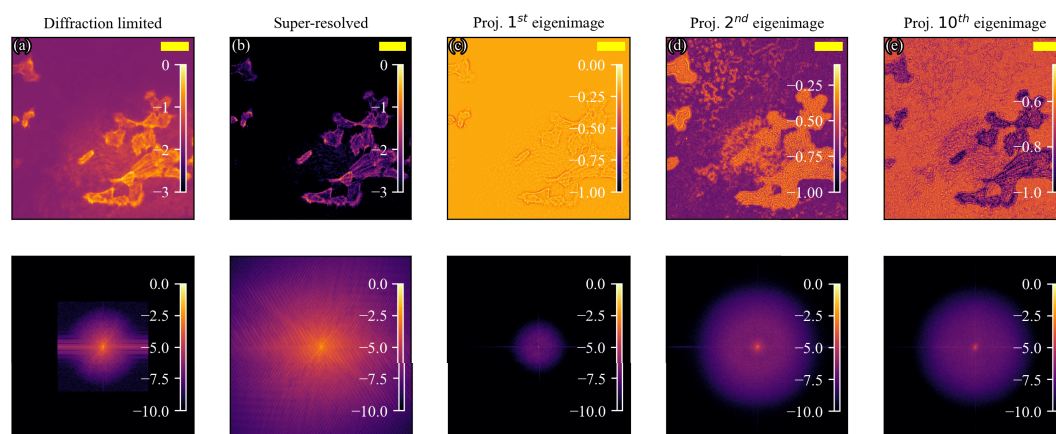
As described previously,  $U$  contains structural information in its columns or eigenimages. And since the number of matrices  $U$  is generally large, studying them is challenging. In the following section we explain how we have carried out a small analysis of the eigenimages of these matrices to show how they differ in terms of the information they carry.

The analysis is done by creating images using a single eigenimage from each patch. The selection criteria is based on the order that every eigenimage has in its own matrix  $U$ . So, an image of the eigenimage of order  $i$  is build using the following function:

$$f(\mathbf{r}_t)^{(i)} = |\mathbf{u}_i \cdot \mathbf{g}_t| \quad (2)$$

As can be seen Eq. (2), the expression corresponds to the projection of the PSF on a determined eigenimage. The stitching procedure is the same than for MUSICAL as describe previously with the patch-based processing. In Fig. 3 we show a visual example of the relation between  $U$  using the projection of eigenimages of different orders, the diffracted-limited image, and the MUSICAL result. The first row show the actual images that were obtained using this method, in addition to the diffracted-limited and MUSICAL image. To improve visualization, these were normalized following different procedures. For the diffracted-limited and super-resolved they were first normalized by its maximum value as  $I_N(\mathbf{r}) = I(\mathbf{r})/\max_{\mathbf{r}}(I(\mathbf{r}))$ . For the projections of eigenimages, the normalization involves the set of all of them. Let  $o$  be the order of an eigenimage. The, the normalization is carried as  $I_N^{(o)}(\mathbf{r}) = I^{(o)}(\mathbf{r})/\max_{\mathbf{r},o} I^{(o)}(\mathbf{r})$ . This allows a fairer comparison between different orders in terms of intensities value in the images. The images are presented in logarithmic scale of base 10. The bottom row shows the Fourier Transform of the corresponding image shown in top. These were normalized as  $I_N(\mathbf{r}) = |I(\mathbf{r})/\max_{\mathbf{r}}(I(\mathbf{r}))|^2$  and then plotted in logarithmic scale as well. Looking to the first to columns, the first we observe is how MUSICAL provides a richer frequency content thanks to its super-resolution properties. Second, we notice how the projection of the first eigenimage (Fig. 3(c),top) do not contain information for super-resolution as can be see in the smaller radius in the frequency domain (Fig. 3(c),bottom) when compared to the diffracted limit. This reduction is due to the stitching process as technically the first eigenimage corresponds actually to the diffracted-limited image when only a single patch is considered. More interesting is that the projection of higher orders (Fig. 3(e),top) provide in fact more frequency content than the diffraction-limited but not more

than MUSICAL as the latter combines all the information available in the eigenimages. Finally, it's interesting to note that even though the signal decreases and the background increases as we move to higher orders, the spectrum seems invariable in terms of coverage. We observed that the coverage in the frequency domain seems the same except for the first order. However, a study of this phenomena goes beyond the scope of this manuscript.



**Fig. 3.** The Fourier Transform allows to compare the spectrum of the diffracted image and the MUSICAL result in sample of Actin acquired with a 10x, 0.25 NA objective. Scale bar is 50  $\mu\text{m}$ . The upper row displays different images after a logarithmic transformation in base 10, while the bottom shows their corresponding Fourier Transform.

### 3.4. Sample preparation, data acquisition and data processing

We consider 2 types of samples in our experiments, namely bead samples with known bead diameters and actin in HeLa cells. The sample preparation, data acquisition, and data processing details for the different samples are presented here. A customized mount system is designed to hold and image the sample.

**Bead sample** The sample was prepared in the cover slip. Tetra beads of diameter 200 nm (Tetraspeck, Thermofisher T7279) were employed and excitation/emission wavelength of 660/690 nm were used for them. The stock solution of tetra beads was diluted in proportion 1:50 in distilled water. 100  $\mu\text{L}$  of the solution was added on the petri dish and allowed to dry completely. We used 40X, 0.65 NA objective lens for  $\text{MO}_2$ . The exposure time of acquisition is 200 ms per frame.

**HeLa cells sample** HeLa cells were grown in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% fetal bovine serum and 1% penicillin/streptomycin in a standard humidified incubator at 37°C with 5%  $\text{CO}_2$ . Cells were seeded into the glass bottom petri dish 1–2 days before imaging. For imaging, the cells were washed in PBS and fixed in 4% paraformaldehyde for 15 minutes. Cells were then washed with PBS and permeabilized with 0.1% Triton X-100 in PBS for 4 minutes, and washed three times in PBS. Cells were then incubated with Atto-647N phalloidin in PBS for 20 minutes in order to label F-actin in the cells.

These samples were imaged using different objective lenses for  $\text{MO}_2$  in order to assess the resolution scalability with the same illumination strategy but using different collection optics. These samples were also used to perform several supplementary studies. In the results reported in the main article, we used four candidate objective lenses for  $\text{MO}_2$ , namely 10X (0.25 NA), 20X (0.40 NA), 40X (0.65 NA) and 50X (0.55 NA). The exposure time in either case is 100 ms per frame.

### 3.5. Data acquisition and analysis

Total 1260 illumination patterns were used to generate super-resolution images in case of beads and Hela cells. However, number of patterns were further optimized to achieve super-resolution with minimum frames. All the raw datasets were processed using MUSICAL, without any preprocessing. Soft-MUSICAL version reported in [23], which alleviates the need of user-specified threshold. Python implementation of the algorithm was used. Further, we used  $\alpha = 4$  (refer to Eq. (1) above) and sub-pixelation of 10 i.e., one pixel in the raw image corresponds to  $10 \times 10$  pixels in the MUSICAL generated image. Raw and processed images presented in the article have been pseudocolored using standard colormaps available on imageJ platform. The selection of the colormap and the intensity range is done heuristically for good visualization, as conventional for microscopy data visualization.

### 3.6. Resolution estimation

In order to estimate the resolution enhancement, we used decorrelation analysis to estimate the resolution of the computationally generated super-resolved images [25]. It has been shown in the past to overcome the limitations of Fourier ring correlation [25]. While performing the decorrelation analysis, we used the Matlab implementation provided by the authors, and set the parameter for sampling the frequency domain as 50 and the number of low-pass filters as 30. In order to convert the cut-off frequency  $k_c$  given by the tool into resolution in metric system, we used the formula given by the authors,  $r = \frac{2 * p}{k_c}$ , where  $p$  is the pixel size.

## 4. Results

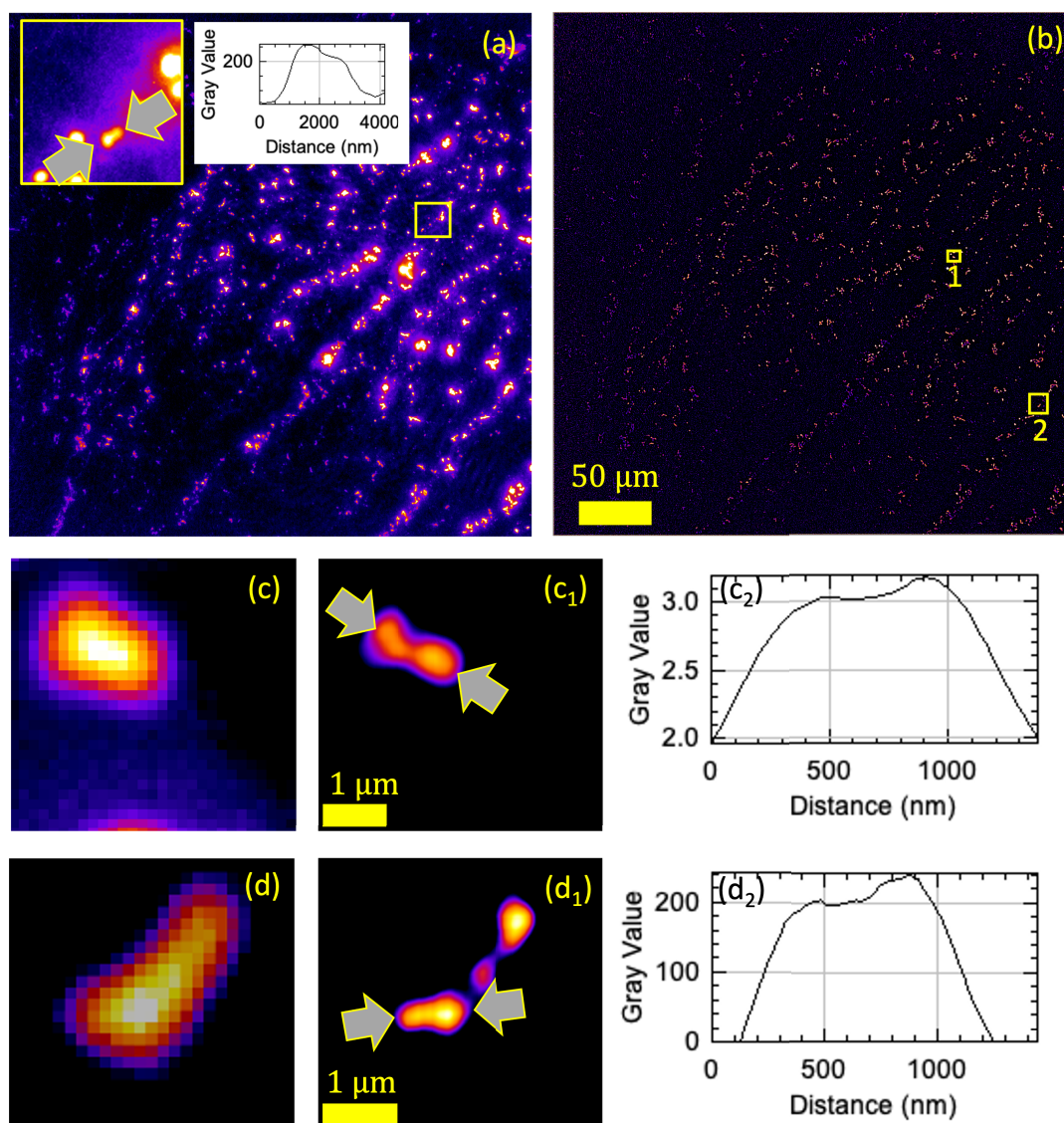
**Results on 200 nm beads** We first demonstrate the resolution improvement of the present approach by imaging 200 nm fluorescent beads. The collection objective lens is 40X (0.65 NA) and the illumination objective is 60X (0.9 NA), which corresponds to diffraction limit of  $\sim 445$  nm. Therefore the beads are smaller than the expected resolution enhancement of linear SIM. Further, some of these beads are spaced within the classical diffraction limit, also closer than the resolution achievable by linear SIM as well.

Average of the images taken using our illumination patterns is shown in Fig. 4 (left and middle panels) as a representative of diffraction limited image. The middle panel shows two region of interests (ROI), each with cluster of beads placed relatively densely. These images are reconstructed using MUSICAL to utilize non-linear patterns introduced in the acquired datasets. The results are shown in the right panel of Fig. 4 for the chosen ROIs.

In order to quantify the resolution enhancement achieved by our method, we measured the PSF from the image of few single beads across the sample in the diffraction limited (average of all images) and the super-resolved images. The average full width at half maximum (FWHM) of the single beads is 946 nm. On the other hand, the average FWHM of the same beads in the super-resolved image is  $\sim 347$  nm, roughly  $3 \times$  the diffraction limit.

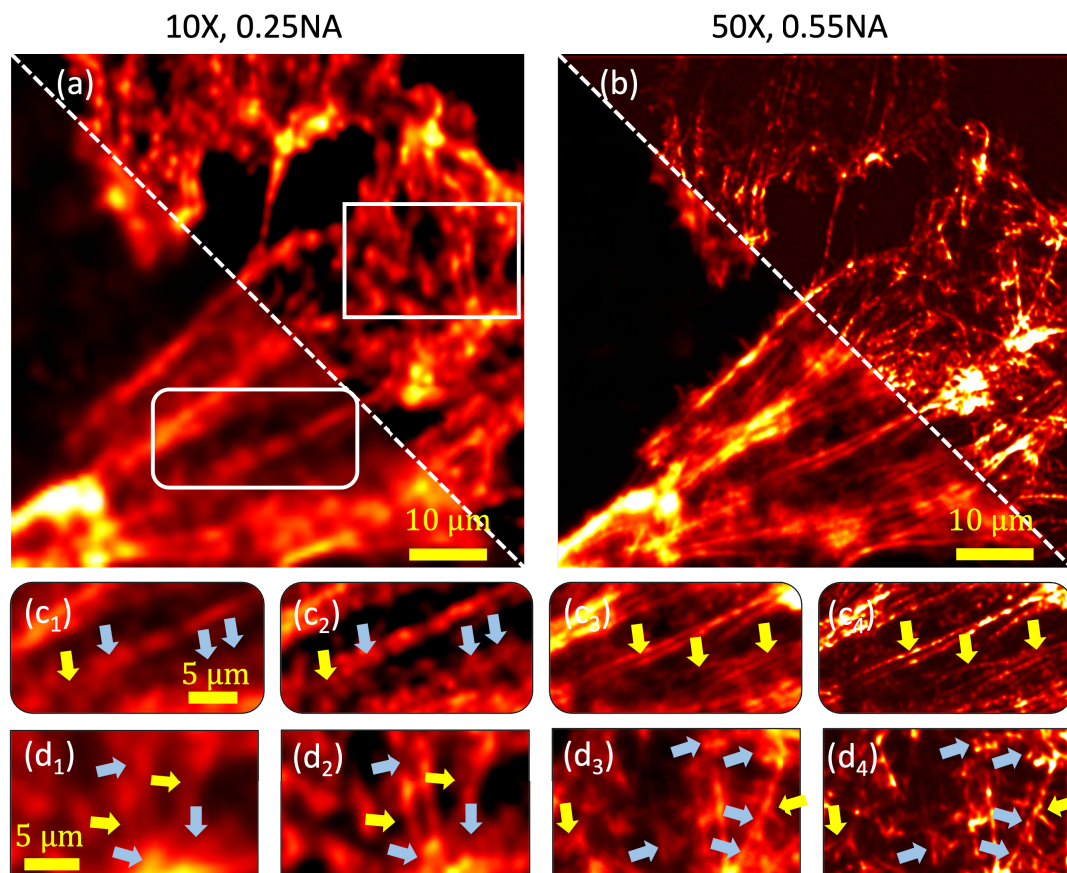
In addition to FWHM, we also estimated resolution using decorrelation analysis. The resolution observed in diffraction-limited image is 954 nm and in super-resolved 376 nm. Therefore, the resolution enhancement is found approximately 2.54. This agrees with the qualitative results as well, as we explain here. A crop from diffraction-limited image that appears qualitatively similar to two crops from super-resolved image are identified in Fig. 4. Each of these crops appear to have two beads placed close to each other. The line profiles across the two beads are also plotted for each crop. It is seen from the line profiles that the beads in the diffraction-limited crop are separated by  $\sim 1069$  nm, while for the super-resolved crops are 400 nm and 368 nm respectively. This indicates also a resolution enhancement by  $\sim 2.6$  to  $2.9$  times. The resolutions obtained by decorrelation analysis are plotted as a function of number of orientations in Supplementary Fig. S2.





**Fig. 4. Resolution improvement beyond the diffraction limit, shown using sample containing 200 nm beads.** 40X,0.65 NA objective lens is used for collection of light. Diffraction limit is  $\sim 445$  nm, resolution observed in diffraction limited image is 954 nm and in super-resolved 376 nm. Region a shown as an inset in top-left panel is diffraction limited and visually similar to the super-resolved versions of regions 1 and 2 shown in top-right panel and zoomed-in in the lower panels.

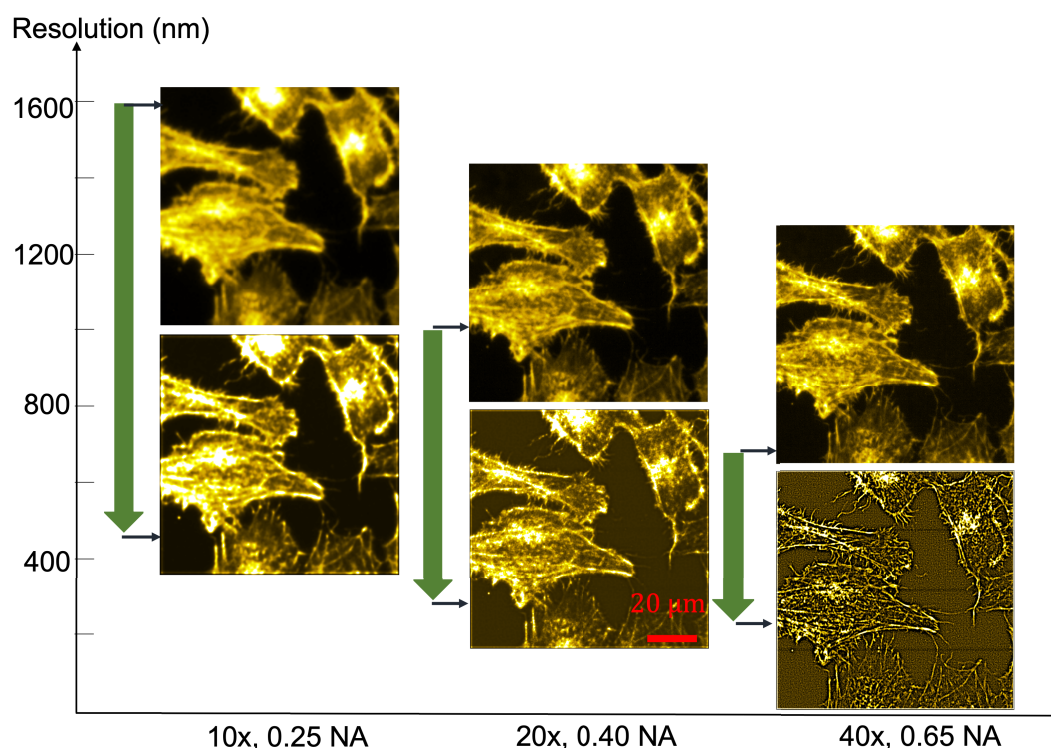
**Actin filaments in cells** Fig. 5 presents the results for actin filaments. It shows the comparison of diffraction limited image and our super-resolved reconstruction of actin filaments of Hela cell using two different objective lenses. Figure 5(a) depicts the superposition of average images and super-resolved images of the the sample using 10X, 0.25 NA. The same ROI is imaged by 50X, 0.55 NA lens and shown in Fig. 5(b). The zoom-ins of one region of interest, shown using white rounded rectangle in Fig. 5(a), are shown for the 10X diffraction limited, 10X super-resolved, 50X diffraction limited, and 50X super-resolved images in Fig.5 (c1-c4), respectively. Analogously, the zoom-ins of another ROI, shown using sharp cornered rectangle in Fig. 5(a), are shown in Fig. 5(d1-d4). The results show improvement of contrast and sharpness using our approach in



**Fig. 5. Super-resolution imaging of actin filaments in HeLa cells.** a) Image acquired with a 10X, 0.25NA lens and a super-resolved image. b) Similar field of view is imaged by 50X, 0.55 NA lens. Smaller ROI are magnified and shown in c1-c4 and d1-d4). Yellow arrows point at contrast enhancements leading to better visualization. Cyan arrows show resolution enhancement leading to better feature details.

sparse regions (shown using yellow arrows), and enhancement of resolution in the dense regions (shown using cyan arrows) of the diffraction limited image. The regions which are not resolved in 0.25 NA (Fig. 5(c1, d1)) are marked and compared with the super-resolved image (Fig. 5(c2, d2)) generated from the datasets acquired using 0.25NA. Figure 5(c2, d2) can also be compared with the Fig. 5 image of 0.55 NA (Fig. 4(c3, d3)). It is seen that the features in Fig.5(c2, d2) generally match well with Fig. 5(c3, d3), with some mismatch of finer details. The possible reason could be the sample preparation, mechanical instability in the system and polarization sensitivity of the SLM. Since SLM is a polarization sensitive device, contrast of the illuminated pattern could be directional dependent. Nonetheless, the issue could be overcome simply by replacing the sample mounting stage with the fine rotational stage and synchronize with the illumination pattern.

The decorrelation analysis for the data acquired using 10X indicated a resolution of 1856 nm in the diffraction limited image and 480 nm for the super-resolved image. This indicates a resolution enhancement of approximately 3.77 times. The decorrelation analysis for the data acquired using 50X indicated a resolution of 697 nm for the diffraction-limited image and 254 nm for the super-resolved. This indicates a resolution enhancement of approximately 2.74. The resolution measured in the diffraction limited and super-resolved images for both collection objective lenses are presented in Supplementary Fig. S2. Furthermore, we have shown the application of current work only using low NA detection since. Since, high NA lenses have limited working distance,



**Fig. 6. Pattern optimization for super-resolution imaging with different collection lenses.** Similar region of interests is imaged by three different NA (0.25, 0.40 and 0.65) without any re-calibration. Scale bar: 20  $\mu\text{m}$ .

more precise alignment of the system will be required to achieve super-resolution on higher NA lenses. Further accommodation for using liquid immersion medium is a consideration if a special sample holder is designed for the purpose. On the other hand, low NA lens capture images with larger depth of field and larger field-of-view, which is desired in many applications.

**Multiscale resolution** Finally, we optimized minimum number of pattern and angles required to achieve super-resolution images using different collection lenses. We used three different collection objective lens i.e., 10X (0.25NA), 20X (0.45 NA), 40X (0.65 NA) and acquired the datasets of actin filaments of the Hela cells. Period of the binary pattern is selected twice the null-to-null width of the PSF of collection lens. Details of the optimum number of angles is simulated and presented in Supplementary Fig. S5 and Supplementary Fig. S7 and Table S1. Simulated protocols is verified experimentally where 20° rotation angle is found suitable for 10X, 0.25 NA to fill OTF symmetrically. On the other hand, 60° is well suited for 20X, 0.45 NA and 40X, 0.65 NA lens. Therefore, total number of frames required to achieve super-resolution images in case of 10X, 20X and 40X are 360, 78 and 48 respectively. The diffraction limited and super-resolved images for each case is shown in Fig. 6. Experimentally observed resolution and the resolution enhancement for each collection geometry is shown in Table 1. Note that, increasing numerical aperture of the collection objective lens decreases the number of photons per pixel therefore the signal to background strength decreased in high NA objective lens. Experimentally calculated resolution in case of 0.25 NA, 0.40 NA and 0.65 NA is found to be 1718 nm, 987 nm and 681 nm respectively. The results show resolution enhancement by a factor of 1.96, 3.40 and 2.68 for 0.25 NA, 0.40 NA and 0.65 NA, respectively. Even though the signal to background ratio for 0.40 NA and 0.65 NA is 3.03 and 2.40, which is considered inadequate in case of conventional SIM, our approach offers resolution enhancement by a factor of 3.40 and 2.68. In Supplementary notes S6 and Supplementary Fig. S8 and S9, we have compared fluctuation based technique



such as MUSICAL, SOFI, SRRF and open source SIM reconstruction tool i.e., FairSIM in low signal to background SIM data. We addressed the artifacts of the methods under two different levels of signal to background ratio (SBR). The sample can be imaged with high SBR in SIM and low SBR in our system to still achieve comparable resolution enhancement. Therefore, the proposed approach could be considered as a simple solution to avoid the instrumentation and computational complexity of SIM. In addition, the entire framework can be applied to different imaging modalities by improving their resolution and field of view.

**Table 1. Experimental diffraction limited resolution and super-resolution using different collection objective lens. Factor of resolution enhancement of 2-3.5 with signal to background ratio of 2.4-5.9. Note that signal to background ratio of 2.4-5.9 considered inappropriate in conventional SIM reconstruction approach. Less resolution enhancement in case of 0.25NA is observed because of poor sampling of PSF.**

Collection lens ( $NA_{coll}$ )	Expt. diffraction limited resolution (nm)	Expt. super resolved (nm)	Resolution enhancement	Signal to background
0.25	1718	877	1.96	5.93
0.40	987	289	3.40	3.03
0.65	681	254	2.68	2.40

## 5. Conclusion

In this work, we have presented a structured illumination approach using linear optical system together with non-linear reconstruction algorithm to achieve super-resolution imaging. Our technique enables more than the conventional 2-fold resolution enhancement of linear SIM. The key enabling idea is to obviate the need of strictly clean peaks in the Fourier spectrum of the illumination by using illumination-blind super-resolution algorithm. The idea is realized through an experimental scheme which introduces an illumination pattern containing high orders of frequency peaks obfuscated by an illumination objective lens that projects binary periodic illumination patterns on the sample. We validate the resolution enhancement using beads sample and actin filaments in cell sample. We show the resolution enhancement using a variety of collection objective lenses, ranging from 10X, 0.25 NA to 50X, 0.55 NA. We used the same illumination objective lens while using the different collection objective lenses, indicating easy scalability of resolution and field-of-view. We use a fixed high NA lens for illumination in the present work therefore the field of view will be limited by the illumination objective lens. The main motivation for this choice was to demonstrate the technique's strength of being able to use different collection lenses. Nonetheless, low NA illumination objective lens can be used to improve the space-bandwidth product of the current approach if desired at almost no extra complication and with a resolution enhancement of only 2x. This can be achieved by replacing the excitation objective lens with a multi-objective lens turret such as used in the collection path.

In addition, the results consistently show resolution enhancement by a factor of 2.6 to 4 because of using high NA condenser lens for the excitation and low NA as a collection objective lens. On the other hand, non-linearity of the reconstruction algorithm supports better resilience to noise as well as better contrast enhancement, which further helps in achieving the said resolution enhancement in very low signal conditions. In the future, we would like to improve the signal strength while balancing the exposure time, light dose, system stability and photobleaching, thereby optimizing the system for living cells with resolution enhancement by a factor of more than 4. We believe that the proposed experimental+computational framework using linear optical system will find wide applicability to many biological applications where simple optical design and super-resolution is essential.



## Author contribution

KA conceived the idea and supervised the work. AB designed experimental system and planned the experiments. DHH and AB optimized and automate the system. SA optimized the computational part and reconstructed all the images. AB, SA and KA analysed the result and prepared the figure. AB and KA mainly wrote the manuscript and all author contributed revised the manuscript.

**Funding.** European Research Council (804233).

**Acknowledgement.** Authors would like to acknowledge Hong Mao and Biswajoy Ghosh, UiT Tromsø, for providing the samples. KA and AB acknowledges the European Research Council Starting grant (id 804233).

**Disclosures.** Authors declare no competing interest.

**Data availability.** Data may be obtained from the authors upon reasonable request.

**Supplemental document.** See [Supplement 1](#) for supporting content.

## References

1. J. W. Lichtman and J.-A. Conchello, "Fluorescence microscopy," *Nat. Methods* **2**(12), 910–919 (2005).
2. J. F. Dekkers, M. Alieva, L. M. Wellens, H. C. Ariese, P. R. Jamieson, A. M. Vonk, G. D. Amatngalim, H. Hu, K. C. Oost, H. J. Snippert, J. M. Beekman, E. J. Wehrens, J. E. Visvader, H. Clevers, and A. C. Rios, "High-resolution 3D imaging of fixed and cleared organoids," *Nat. Protoc.* **14**(6), 1756–1771 (2019).
3. G. Devitt, K. Howard, A. Mudher, and S. Mahajan, "Raman spectroscopy: an emerging tool in neurodegenerative disease research and diagnosis," *ACS Chem. Neurosci.* **9**(3), 404–420 (2018).
4. S. W. Hell and J. Wichmann, "Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy," *Opt. Lett.* **19**(11), 780–782 (1994).
5. H. Blom and J. Widengren, "Stimulated emission depletion microscopy," *Chem. Rev.* **117**(11), 7377–7427 (2017).
6. W. Moerner, "Microscopy beyond the diffraction limit using actively controlled single molecules," *J. Microsc.* **246**(3), 213–220 (2012).
7. J. Schnitzbauer, M. T. Strauss, T. Schlichthaerle, F. Schueder, and R. Jungmann, "Super-resolution microscopy with DNA-PAINT," *Nat. Protoc.* **12**(6), 1198–1228 (2017).
8. E. Nehme, L. E. Weiss, T. Michaeli, and Y. Shechtman, "Deep-STORM: super-resolution single-molecule microscopy by deep learning," *Optica* **5**(4), 458–464 (2018).
9. A. Butola, D. A. Coucheron, K. Szafranska, A. Ahmad, H. Mao, J.-C. Tinguely, P. McCourt, P. Senthilkumaran, D. S. Mehta, K. Agarwal, and B. S. Ahluwalia, "Multimodal on-chip nanoscopy and quantitative phase imaging reveals the nanoscale morphology of liver sinusoidal endothelial cells," *Proc. Natl. Acad. Sci.* **118**(47), e2115323118 (2021).
10. T. Dertinger, R. Colyer, G. Iyer, S. Weiss, and J. Enderlein, "Fast, background-free, 3D super-resolution optical fluctuation imaging (SOFI)," *Proc. Natl. Acad. Sci.* **106**(52), 22287–22292 (2009).
11. S. Culley, K. L. Tosheva, P. M. Pereira, and R. Henriques, "SRRF: Universal live-cell super-resolution microscopy," *Int. J. Biochem. Cell Biol.* **101**, 74–79 (2018).
12. O. Solomon, M. Mutzafi, M. Segev, and Y. C. Eldar, "Sparsity-based super-resolution microscopy from correlation information," *Opt. Express* **26**(14), 18238–18269 (2018).
13. I. Yahiatene, S. Hennig, M. Muller, and T. Huser, "Entropy-based super-resolution imaging (ESI): From disorder to fine detail," *ACS Photonics* **2**(8), 1049–1056 (2015).
14. K. Agarwal and R. Macháň, "Multiple signal classification algorithm for super-resolution fluorescence microscopy," *Nat. Commun.* **7**(1), 13752 (2016).
15. L.-H. Yeh, L. Tian, and L. Waller, "Structured illumination microscopy with unknown patterns and a statistical prior," *Biomed. Opt. Express* **8**(2), 695–711 (2017).
16. M. G. Gustafsson, "Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy," *J. Microsc.* **198**(2), 82–87 (2000).
17. E. Mudry, K. Belkebir, J. Girard, J. Savatier, E. Le Moal, C. Nicoletti, M. Allain, and A. Sentenac, "Structured illumination microscopy using unknown speckle patterns," *Nat. Photonics* **6**(5), 312–315 (2012).
18. M. G. Gustafsson, "Nonlinear structured-illumination microscopy: wide-field fluorescence imaging with theoretically unlimited resolution," *Proc. Natl. Acad. Sci.* **102**(37), 13081–13086 (2005).
19. R. Heintzmann and T. Huser, "Super-resolution structured illumination microscopy," *Chem. Rev.* **117**(23), 13890–13908 (2017).
20. E. H. Rego, L. Shao, J. J. Macklin, L. Winoto, G. A. Johansson, N. Kamps-Hughes, M. W. Davidson, and M. G. Gustafsson, "Nonlinear structured-illumination microscopy with a photoswitchable protein reveals cellular structures at 50-nm resolution," *Proc. Natl. Acad. Sci.* **109**(3), E135–E143 (2012).
21. G. Ball, J. Demmerle, R. Kaufmann, I. Davis, I. M. Dobbie, and L. Schermelleh, "SIMcheck: a toolbox for successful super-resolution structured illumination microscopy," *Sci. Rep.* **5**(1), 15915 (2015).
22. K. Samanta, S. Sarkar, S. Acuna, J. Joseph, B. S. Ahluwalia, and K. Agarwal, "Blind super-resolution approach for exploiting illumination variety in optical-lattice illumination microscopy," *ACS Photonics* **8**(9), 2626–2634 (2021).

23. S. Acuña, I. S. Opstad, F. Godtlielsen, B. S. Ahluwalia, and K. Agarwal, “Soft thresholding schemes for multiple signal classification algorithm,” *Opt. Express* **28**(23), 34434–34449 (2020).
24. I. S. Opstad, D. H. Hansen, S. Acuña, F. Ströhl, A. Priyadarshi, J.-C. Tinguely, F. T. Dullo, R. A. Dalmo, T. Seternes, B. S. Ahluwalia, and K. Agarwal, “Fluorescence fluctuation-based super-resolution microscopy using multimodal waveguided illumination,” *Opt. Express* **29**(15), 23368–23380 (2021).
25. A. Descloux, K. S. Grubmayer, and A. Radenovic, “Parameter-free image resolution estimation based on decorrelation analysis,” *Nat. Methods* **16**(9), 918–924 (2019).

Supplemental Document

**Optics** EXPRESS

## Scalable-resolution structured illumination microscopy: supplement

**ANKIT BUTOLA,<sup>1,2,\*</sup>  SEBASTIAN ACUNA,<sup>1,2</sup>  DANIEL HENRY HANSEN,<sup>1</sup> AND KRISHNA AGARWAL<sup>1</sup>**

<sup>1</sup>*Department of Physics and Technology, UiT The Arctic University of Norway, Norway*

<sup>2</sup>*Co-first authors*

\*[ankitbutola321@gmail.com](mailto:ankitbutola321@gmail.com)

---

This supplement published with Optica Publishing Group on 15 November 2022 by The Authors under the terms of the [Creative Commons Attribution 4.0 License](https://creativecommons.org/licenses/by/4.0/) in the format provided by the authors and unedited. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

Supplement DOI: <https://doi.org/10.6084/m9.figshare.21365490>

Parent Article DOI: <https://doi.org/10.1364/OE.465303>

## 1 Supplementary Material

### 2 Scalable-resolution structured illumination 3 microscopy

4 ANKIT BUTOLA,<sup>1,†,\*</sup> SEBASTIAN ACUNA,<sup>1,†</sup> DANIEL HENRY HANSEN,<sup>1</sup>  
5 AND KRISHNA AGARWAL<sup>1</sup>

6 <sup>1</sup>*Department of Physics and Technology, UiT The Arctic University of Norway, Norway*

7 <sup>†</sup>*Shared first authorship*

8 <sup>\*</sup>*ankitbutola321@gmail.com*

#### 9 Note S1. Periodicity of the illumination pattern

10 Actin filaments of the HeLa cells is used as a sample in this study. We chose  $MO_2$  to be 20X, 0.4  
11 NA for acquiring this dataset. The diffraction limited image in this experiment was acquired by  
12 using a uniform illumination. As discussed in the main article, a periodicity of  $2w$  is optimal  
13 where  $w$  is the null-to-null width of the point-spread function of the illumination objective lens.  
14 We consistently increased the illumination pattern's periodicity from  $\sim w$  to  $\sim 4w$ . We used only  
15 4 translations per period for each pattern. Rotation of patterns by  $2^\circ$  was employed, resulting into  
16 90 rotations. Therefore the number of images per stack is 360.

17 Figure S3 presents the results of using different periods of illumination patterns. The  
18 reconstructed images are show for the illumination modulated stacks. We see that the amount of  
19 details increases with the period. This is because of the increased flatness. However, the presence  
20 of spurious mesh background and the spottiness in the reconstructed image also increases with  
21 the increase in the period. On the other hand, actin filaments are more clearly visible and reliably  
22 imaged in the case of  $2w$  period as compared to the other cases.

#### 23 Note S2. Shifts of pattern

24 In order to increase the resolution we use multiple patterns of illuminations created through the  
25 SLM. We designed the periodic pattern to be similar to regular SIM. For a particular angle of  
26 illumination we consider the case where the pattern align perfectly with the array in the SLM.  
27 For a pattern with period of 14 pixels (7 pixels on, 7 pixels off) the maximum amount of shifts  
28 is 14. We performed a small simulation to check the effective illumination of the pattern after  
29 considering also the rotations. The results are displayed in Fig. S4.

30 For the case of 14 pixels, we created 90 different rotations of the pattern, and then each one  
31 was shifted. As a result, for 14 pixels pattern we generated 1260 different illumination patterns.  
32 Then, we indexed each shift by an integer number between 0 and 13, and picked different subsets  
33 of them and summed them to check uniformity. In Fig. S4, we show how the binary pattern is  
34 shifted in a single dimension for each of the cases and which shifts are considered for every case.

35 Figure S4(a) shows the case where the patterns were not shifted but only rotated. The result  
36 shows that the illumination is not homogeneous. Figure S4(b) corresponds to shifting the pattern  
37 half a period ( a total of two patterns per orientation such that the two patterns are the opposite of  
38 each other). This results in homogeneous illumination as expected. However, upon observation,  
39 we see that the pixels 2-6 and 8-13 of the pattern do not experience any mutual change in intensity,  
40 which also implies that the spatial fluctuations from them are absent. Since MUSICAL works  
41 on spatio-temporal fluctuations, this setting is not optimal despite presenting homogeneous  
42 illumination. Fig. S4c shows the case where all the even illuminations are selected resulting in a  
43 non-homogeneous pattern. Finally, selecting all the frames (Fig. S4d) results in homogeneous  
44 illumination as well.

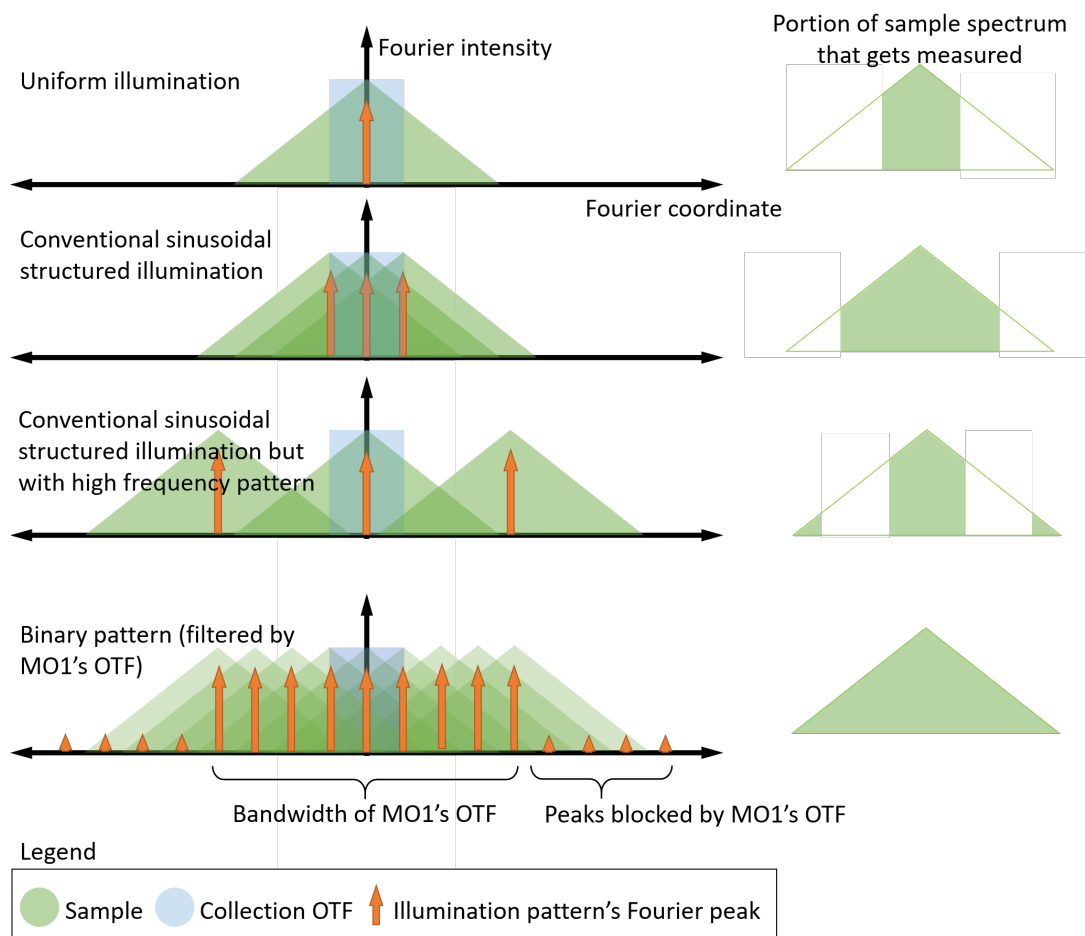


Fig. S1. Comparison with conventional SIM in 1-D Fourier domain is presented here. The multiple peaks of the binary patterns allow get filtered by the illumination objective lens and contribute to the sample frequency span that gets actually measured.

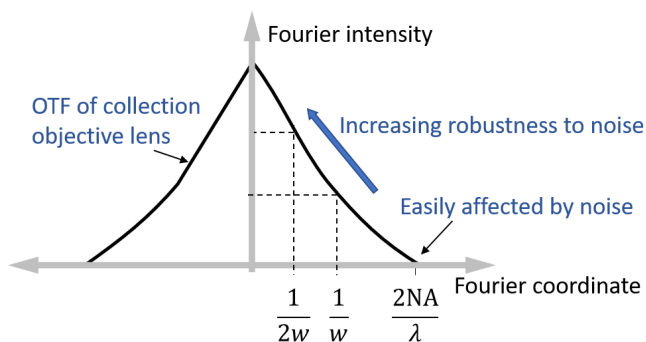


Fig. S2. The periodicity of patterns we used is  $2w$ , which corresponds to one-fourth of the cut-off frequency. Since the Fourier intensity of the OTF is significantly higher for patterns of period  $2w$ , our reconstructions are more robust than if patterns of period  $w$  were used.

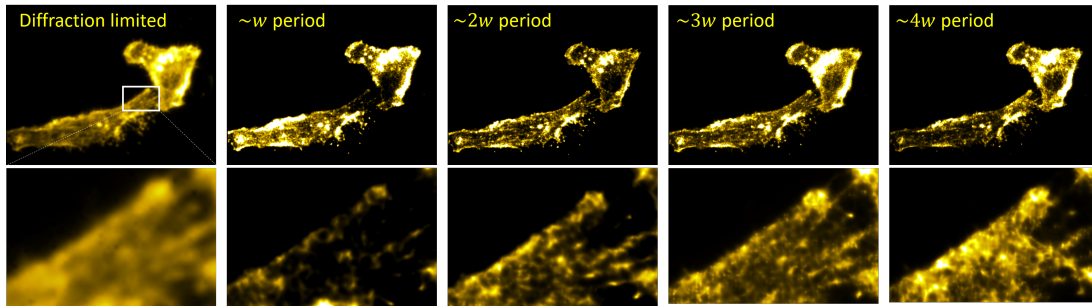


Fig. S3. **Effect of period of illumination patterns.** Actin filaments imaged and reconstructed using illumination patterns of different periods. The null-to-null width of the point spread function is denoted as  $w$ .

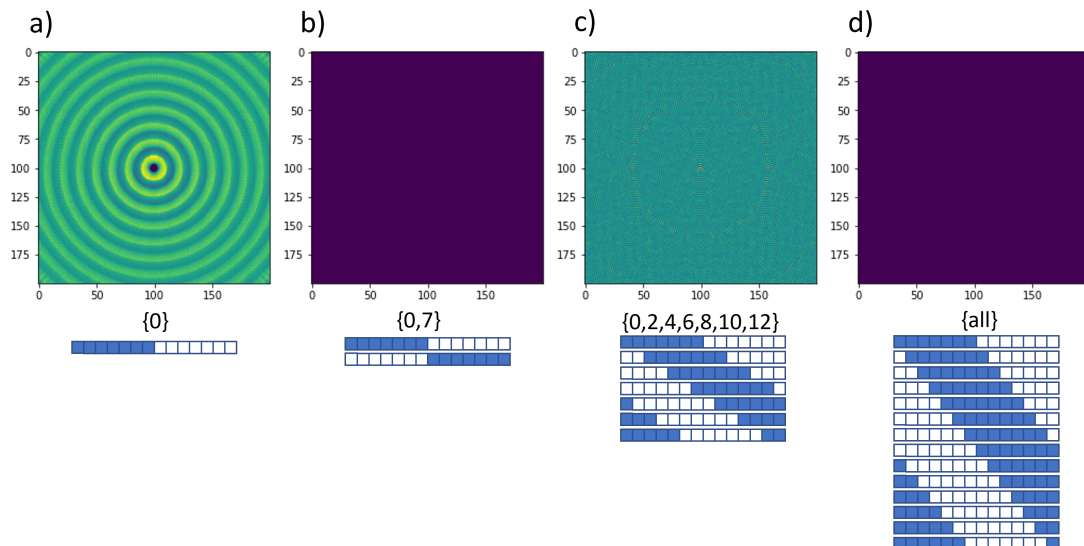


Fig. S4. **Summed sequence of masks using different subsets of shifts and binary patterns shown in one-dimension.** **a.** A single illumination for every angle. It results in a circular pattern. **b.** Two shifts per angle result in uniform illumination. **c.** Seven different shifts create a non-uniform illumination. **d.** All the shifts (14) create a uniform illumination.

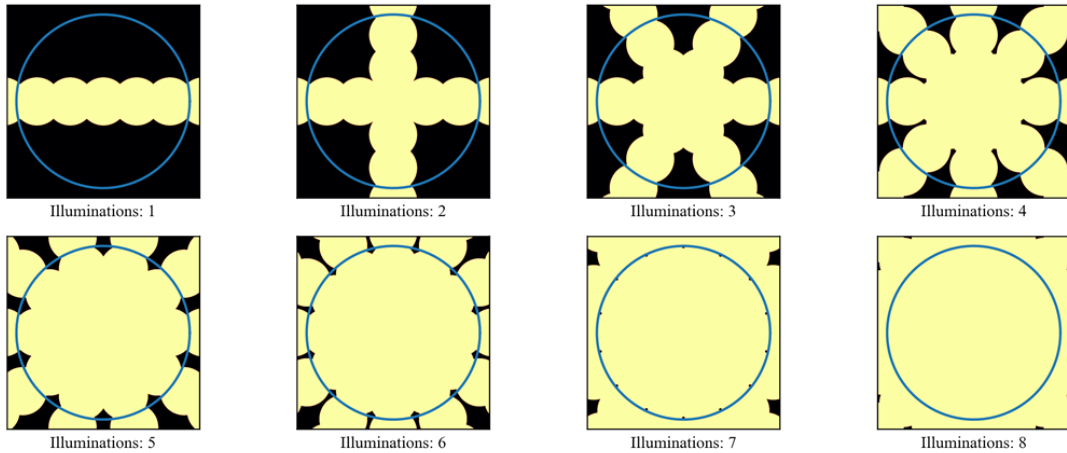


Fig. S5. Effect of different number of illumination angles to fill the Fourier spectrum of the super-resolved image.

45 Since using two shifts was not enough to create fluctuation we chose to use all the frames.

46 **Note S3. Spectral area coverage scheme for determining the number of orientations**  
47

48 In order to determine an optimal number of orientations of illuminations, we followed a  
49 geometrical approach. We compute the minimum number of angles such that the periodic  
50 illumination patterns fully cover a certain area of the 2D Fourier spectrum needed for supporting  
51 the desired resolution. For conventional microscopy (to which we refer as 1x), we assumed a disk  
52 in the Fourier spectrum of radius  $f_1$ . Such function is defined using  $\rho = \sqrt{f_x^2 + f_y^2}$  as follows:

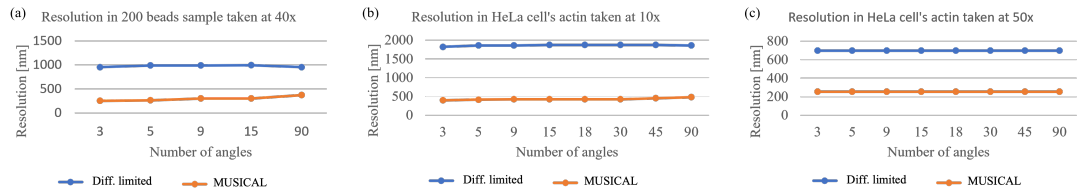
$$D(\rho) = \begin{cases} 1 & \rho \leq f_1 \\ 0 & \rho > f_1 \end{cases} \quad (1)$$

53 Here, it is practically fair to assume that the signal to background ratio of individual images  
54 defines the value of  $f_1$  and the signal outside the radius  $f_1$  is not separable from noise. Next,  
55 assuming an ideal solver, we should be able to retrieve a spectrum given by the superposition of  
56 the convolution of disk  $D(\rho)$  and an impulse train given for every different illumination angle.  
57 Using  $D(\rho/K)$  results into a binary mask of the desired minimum spectrum. Then, we measured  
58 the overlap between the desired mask and the pattern created by different number of angles. The  
59 angles were defined by segmenting the range  $[0, \pi)$  uniformly.

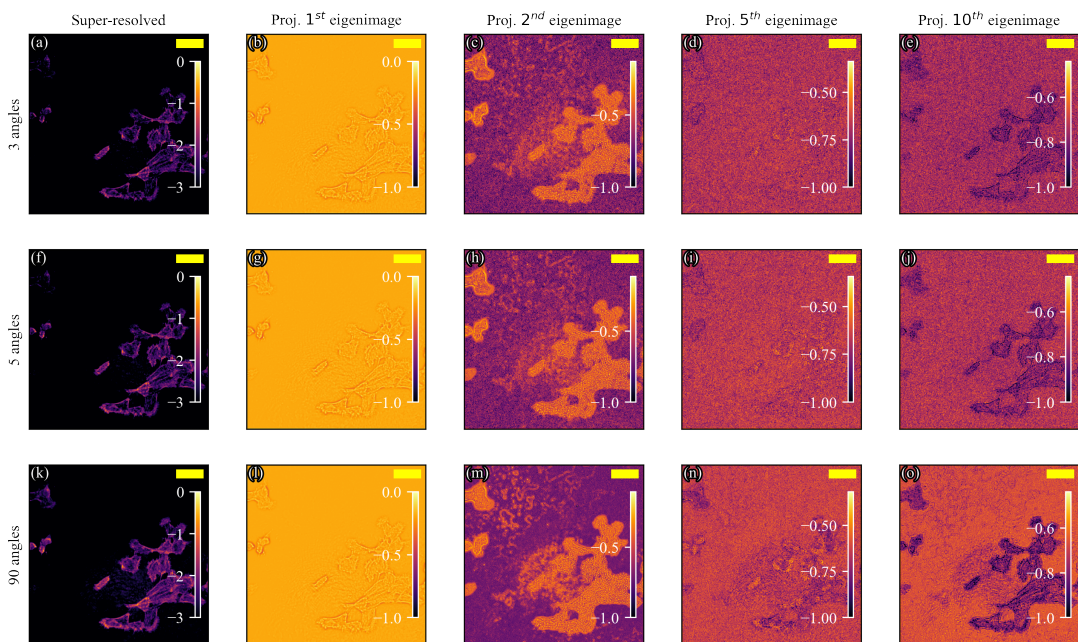
60 **Note S4. Illustration of the effect of using more orientations than necessary**

61 We used the actin filaments in HeLa cells for this study. The study uses the data acquired for  
62 collection objective lens MO<sub>2</sub> being 10X, 0.25 NA. We compared the reconstructions obtained  
63 by using different subsets of frames based in the number of angles, starting from 3 up to 90  
64 different illumination angles. To do this, we took the full stack with 90 illuminations and every  
65 translation and filter it depending on the configuration of interest. Since the stack contains all the  
66 even angles between 0 and 178 (90 in total), we divided the range uniformly using as parameter  
67 the number of angles to consider (i.e 3 or 90) and selected the corresponding subset by taking  
68 the angles (together with all its corresponding translations) where such division is possible. For  
69 instance, taking 3 angles means that the illumination angles taken are 0, 60 and 120 degrees. For





**Fig. S6. Resolution assessment using decorrelation analysis.** The resolution gained was estimated between 2.5 and 4 times the resolution obtained with the average of every stack. The blue colored line plots indicate the resolution obtained for the diffraction limited images and the orange colored line plots indicate the resolution obtained for super-resolved images.



**Fig. S7. Effect of the number of illuminations in the projections of the eigenimages for 3, 5 and 90 angles.** The effect is more noticeable in the projections of highest order. We show here the example of the number 10 in which the difference between structure and background is accentuated. The effect is less apparent for lower order projections.



$NA_{coll}$	Number of orientations	Number of images	Experimental diffraction-limited resolution (nm)	Experimental super-resolved (nm)	Resolution enhancement	Signal to background ratio
0.25	3	120	1592	480	3.31	5.89
	5	200	1677	877	1.91	5.96
	9	360	1718	877	1.96	5.93
	90	3600	1718	909	1.88	5.93
0.40	3	78	987	289	3.40	3.03
	5	130	1009	398	2.53	3.03
	9	234	899	289	3.10	3.01
	90	2340	1027	385	2.66	3
0.65	3	48	681	254	2.68	2.40
	5	80	689	265	2.59	2.40
	9	144	672	265	2.53	2.41
	90	1440	627	302	2.07	2.39

Table S1. Resolution calculation with number of images required at different orientations for 3 different collection lens. Factor of resolution enhancement with signal to background ratio is shown for each case.

70 5 angles we take 0, 36, 72, 108 and 144 degrees. While simple, this procedure does not allow for  
71 any combination of angles. For example, taking 4 different angles is not possible in this range as  
72 the it would require the illuminations at 0, 45, 90 and 135 degrees. Since 45 and 135 are not  
73 available, we cannot considerate this subset of illuminations. Next, we obtained different stacks,  
74 one per number of angles and computed its reconstruction using MUSICAL [1]. The resolution  
75 was then assessed using decorrelation analysis [2] implemented in MATLAB by its authors.

76 In Fig. S6 we report the maximum resolution given by the tool in nanometers. It shows  
77 that the resolution itself does not change significantly with the number of orientations. This is  
78 understandable from the fact that the noise in the measured data serves as the final determinant  
79 of the resolution in practice. However, we investigated that number of illumination angles plays  
80 an important role to fill the Fourier spectrum of the reconstructed super-resolution image. Figure  
81 S5 shows the effect of different number of illumination angle in the Fourier spectrum coverage.  
82 Blue circle represents the collection cone of illumination objective (0.9NA) and the yellow circle  
83 corresponds to 0.25 NA lens. Missing frequency components can be seen with 3 angle orientation  
84 and are barely fill with 6 angle illumination. Similar simulation study is done for 20X, 0.40 NA  
85 and 40X, 0.65 NA lens where 3 angle orientation is found optimum for symmetrically cover the  
86 Fourier spectrum of the specimen. Then details of number of frames at different angles and the  
87 factor of resolution enhancement is shown in Table S1.

88 Further, the effect of using large number of orientations by studying the MUSICAL images of  
89 different orders. These results are presented in S7In the figure we observe how as we increase the  
90 number of angles (and therefore the number of frames) the overall contrast improves. However,  
91 it was not possible to determine an important degree of improvement. On the other-hand, the  
92 projections of different eigenimages provides an interesting view of the relation between signal  
93 and background in the sample. While in general, the eigenimages associated to the signal  
94 (projections of the first and second order eigenimages, shown in the second and third column

95 respectively) do not show an important difference, the ones associated to the signal do (fifth  
 96 and tenth order, shown in fourth and fifth column respectively). From this analysis we conclude  
 97 that the contrast between noise and signal is proportional to the number of angles (and frames),  
 98 proving an easier separation between noise and signal sub-spaces. However, this is not explicitly  
 99 exploited by MUSICAL and more research is needed.

#### 100 **Note S5. MUSICAL for patterned illumination**

101 This theory was published originally in [3], and is reproduced for the sake of completeness here  
 102 with minor modifications. Consider a sample consisting of point-like sources corresponding to  
 103 fluorophores. These have the same properties and emit randomly provided that they are excited  
 104 by an excitation beam. Consider now a grid on the sample plane finer than the actual sensor  
 105 that represents the position of such emitters through its pixels. The grid is defined as  $E(m)$   
 106 where  $m$  is the index of the pixels. We consider too the point-spread-function of an optical  
 107 system  $G(p, m)$  where  $p$  is the index of the pixel in the camera. If the illumination pattern is  
 108  $L(m, n) = \sum_{q=1}^Q a_q \mathcal{L}(m, q)$  with  $q$  being one particular spatial frequency. Then, in the camera,  
 109 for the  $n^{th}$  illumination pattern the image is given by  $I(p, n) = \sum_m e G(p, m) E(m) L(m, n)$  with  
 110  $e$  being the emission coefficient of the emitter. This can be expressed in matrix notation as  
 111  $I = GEL$  where the  $n^{th}$  column of  $I$  contains the  $n^{th}$  image  $I(p, n)$  taken under illumination  
 112  $L(m, n)$ , the  $(p, m)^{th}$  element of  $G$  corresponds to  $G(p, m)$ ,  $E$  is a diagonal matrix with  $m^{th}$   
 113 element being  $E(m)$ , and finally,  $L$  contains the  $n^{th}$  illumination pattern  $L(m, n)$  in its columns.

114 MUSICAL decomposes the image stack  $I$  using SVD which is equivalent to perform eigen-  
 115 decomposition over the square matrix  $J = \Pi^T$ . This can be written as  $J = GELL^T E^T G^T$ .  
 116 Therefore, the patterns  $L$  contribute to MUSICAL through the mathematical support of  $LL^T$ . We  
 117 call the  $(m_1, m_2)^{th}$  element of the matrix  $LL^T$  as  $l(m_1, m_2)$  which can be fully expressed as:

$$l(m_1, m_2) = \sum_n L(m_1, n) L(m_2, n) = \sum_n \sum_{q_1=1}^Q \sum_{q_2=1}^Q a(q_1, n) a(q_2, n) \mathcal{L}(m_1, q_1) \mathcal{L}(m_2, q_2) \quad (2)$$

118 From this, the maximum spatial frequency support of the eigimages derived from  $J$  is given  
 119 by the multiplication of all the illuminations patterns with each other. This corresponds to the  
 120 convolutions of all frequency components with each other. As a result, if the maximum spatial  
 121 frequency in the illumination is  $k_{max}$  then the maximum frequency of  $LL^T$  is  $2k_{mmax}$ . Further,  
 122 multiplication with  $G$  on either side of  $J$  contributes additional  $2k_{obj}$  where  $k_{obj}$  is the maximum  
 123 frequency supported by the optical transfer function of the collection objective. This implies that  
 124 the best possible resolution achievable by MUSICAL for illumination engineered systems under  
 125 conducive conditions and selection of best algorithmic parameters is given by the frequency  
 126  $2(k_{max} + k_{obj})$ .

#### 127 **Note S6. Comparison of super-resolution methods when used for SIM data**

128 Here we attempt reconstruction of SIM data with different 3 methods designed for fluorescence  
 129 images stack displaying intensity fluctuations. The data is freely available in [4] and corresponds  
 130 to microtubules imaged under two different acquisition parameters. The images were acquired on  
 131 a DeltaVision OMX V3 Blaze using 60x/1.42 objective. The sample was stained with Alexa 488.

132 Here we use two sets that show distinctive signal to background intensity. The lowest  
 133 correspond to a ratio of 1.6, obtained for 1% laser power and 30 ms of exposure. The second one  
 134 has 5.2, taken at 31.3% and 30 ms. Each set correspond to a single plane with 3 orientations and  
 135 5 phases.

136 Signal to background (SBR) was estimated selecting arbitrary pairs of patches of 3x3 pixels in  
 137 the image and considering their respective average values (average of 9 pixels). On each pair,

138 one corresponds to high intensity value and another to a low intensity value. These points are  
139 close. The reported SBR is the average of the ratio (high over low) of the three points. Only the  
140 average image was considered in this estimation.

141 The reconstruction were made using different already established methods namely Super-  
142 resolution optical fluctuation imaing (SOFI) [5]and Super-Resolution Radial Fluctuations  
143 (SRRF) [6]. SOFI combines the information of multiple frames through cumulants. By doing  
144 so it reduces the effective size of the PSF by orders of 2 to 4, increasing the resolution of the  
145 image. SRRF follows a different approach. It computes radial maps on each frame. These maps  
146 estimate the degree of converging gradient on arbitrary points in a grid finer than the available in  
147 the original data. The PSF is assumed symmetrical and with high degree of radially on points  
148 where an actual source is located. After this step, the maps can be combined following a variety  
149 of mechanisms. Here we used cumulants, which follows the same argument as SOFI.

150 In addition to MUSICAL-S, we have also used FairSIM. The results for two different signal to  
151 background acquisitions are displayed in Fig. S8 and S9.

152 The results show that all the methods are prone to artifacts when the signal is low. However,  
153 the nature of them differs greatly. SRRF tends to create discontinuous lines that make resolving  
154 between structures challenging in addition to add artificial structures in the background. SOFI 2  
155 (SOFI of order 2) provides good results in both cases. However, challenging regions where there  
156 is also a gaze around the structure (right part) creates mesh-like structures whenever the signal  
157 is low. MUSICAL also creates meshes. However, the contrast it's better in regions where the  
158 original image display a haze. Finally, FairSIM creates artifacts in the background that merges  
159 with the actual structure. This improves greatly with higher signal levels as the parameters can  
160 be estimated in a better way. Nonetheless, these results establish the robustness of MUSICAL in  
161 comparison to the other methods compared here in low SBR conditions.

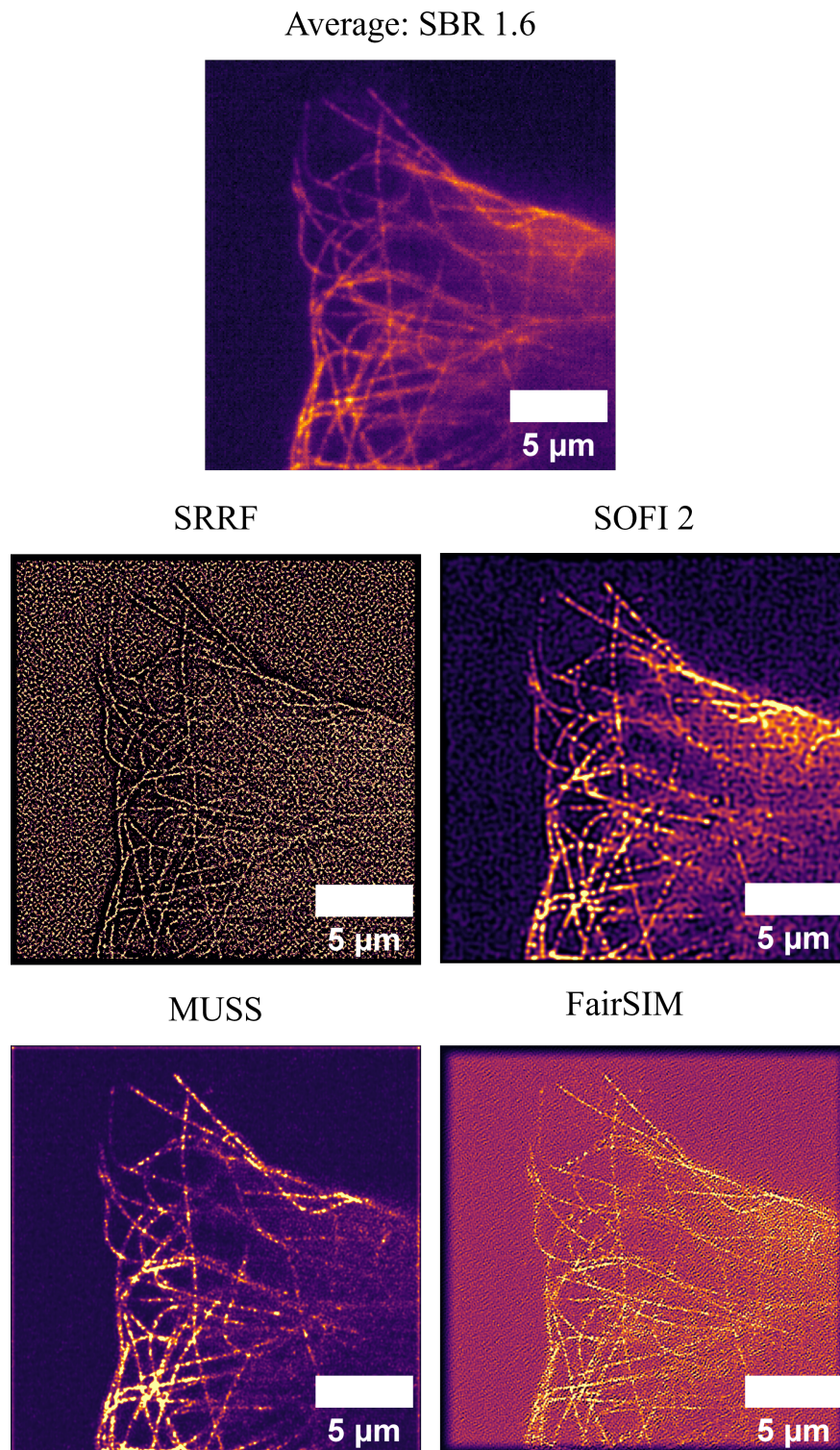


Fig. S8. **Results of applying super-resolution to SIM data with low signal-to-background ratio.** The data consist of microtubules taken at 3 different angles with 5 phases each (total of 15 images). The result show the different type of artifacts the methods create in low signal scenario. SRRF creates noise in the background and intermittent lines. SOFI creates mesh-like structures, specially in regions where the average shows haze. MUSS contrast is not uniform across the image but provide good reconstruction overall. FairSIM creates several artifacts in the background that even merge with the structure.



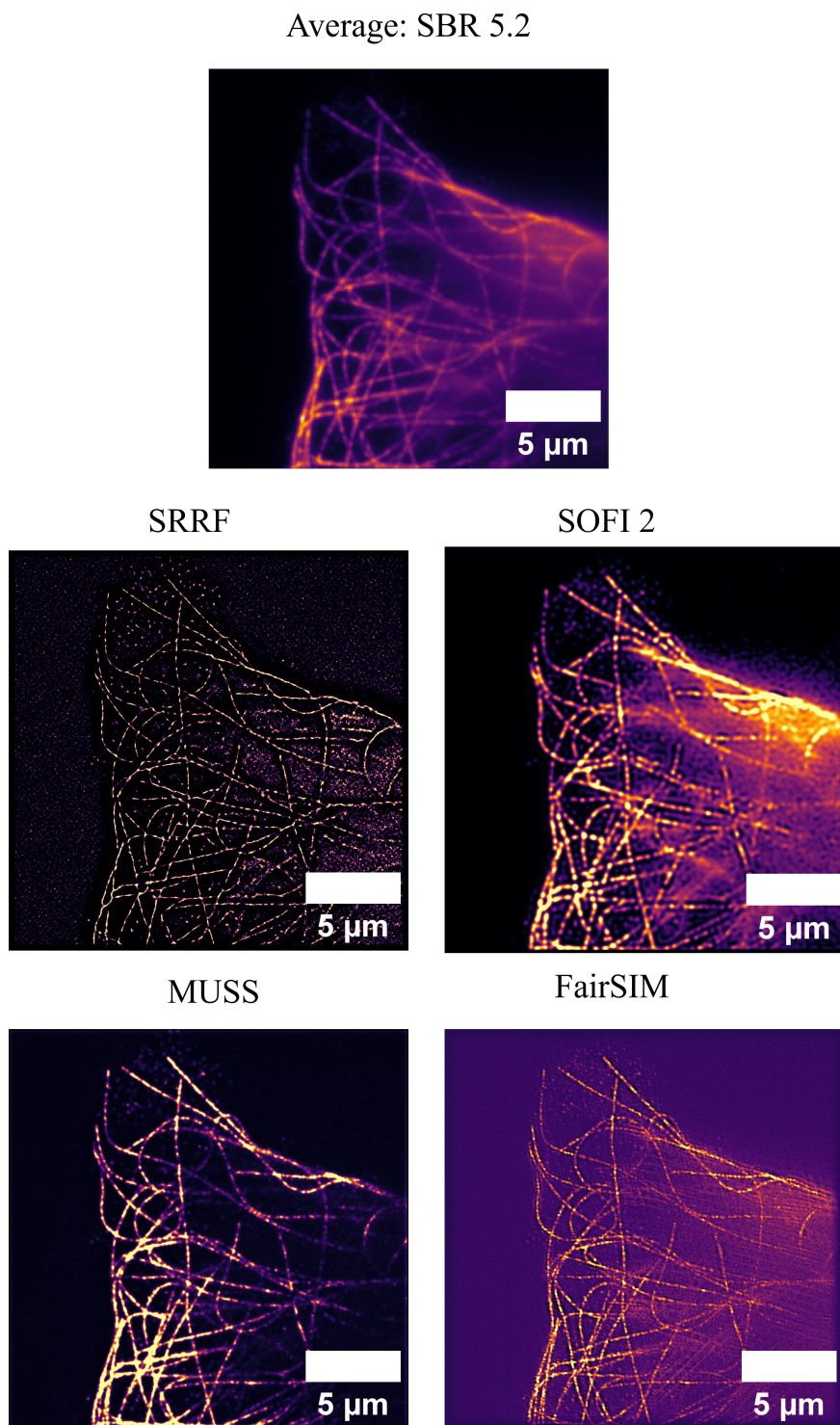


Fig. S9. **Results of applying super-resolution to SIM data with high signal-to-background ratio.** The data consist of microtubules taken at 3 different angles with 5 phases each (total of 15 images). The result show how all the methods benefit from higher signal levels. SRRF tends to creates finer lines which in this case match the underlying structure. SOFI's artifacts are less prominent and the contrast improves across the entire image. MUSS contrast improves and structures definition improves. FairSIM is able of estimating better the parameters which in turn improves the reconstruction.

162 **References**

- 163 1. K. Agarwal and R. Macháň, “Multiple signal classification algorithm for super-resolution fluorescence microscopy,”  
164 Nat. communications **7**, 1–9 (2016).
- 165 2. A. Descloux, K. S. Größmayer, and A. Radenovic, “Parameter-free image resolution estimation based on decorrelation  
166 analysis,” Nat. methods **16**, 918–924 (2019).
- 167 3. K. Samanta, S. Sarkar, S. Acuna, J. Joseph, B. S. Ahluwalia, and K. Agarwal, “Blind super-resolution approach for  
168 exploiting illumination variety in optical-lattice illumination microscopy,” ACS Photonics **8**, 2626–2634 (2021).
- 169 4. C. S. Smith, J. A. Slotman, L. Schermelleh, N. Chakrova, S. Hari, Y. Vos, C. W. Hagen, M. Müller, W. van Cappellen,  
170 A. B. Houtsmuller, J. P. Hoogenboom, and S. Stallinga, “Structured illumination microscopy with noise-controlled  
171 image reconstructions,” Nat. Methods **18**, 821–828 (2021).
- 172 5. T. Dertinger, R. Colyer, G. Iyer, S. Weiss, and J. Enderlein, “Fast, background-free, 3D super-resolution optical  
173 fluctuation imaging (SOFI),” Proc. Natl. Acad. Sci. **106**, 22287–22292 (2009).
- 174 6. S. Culley, K. L. Tosheva, P. M. Pereira, and R. Henriques, “SRRF: Universal live-cell super-resolution microscopy,”  
175 The international journal biochemistry & cell biology **101**, 74–79 (2018).



