FRIEDRICH-SCHILLER-
UNIVERSITÄT
JENA

Doctoral Thesis

Holistic Improvement of Image Acquisition and Reconstruction in Fluorescence Microscopy

A dissertation submitted in fulfillment of
the requirements to attain the degree of

Doctor rerum naturalium (Dr. rer. nat.) Physik

in the Faculty of Physics and Astronomy of

Friedrich-Schiller-Universität Jena

presented by

M.Sc. *René Lachmann*

born on 26 May 1988, Jena (Germany)

2022

Reviewer / Examiner

1. Prof. Dr. Rainer Heintzmann, Jena
2. Prof. Dr. Christian Eggeling, Jena
3. Prof. Dr. Jörg Enderlein, Göttingen

Disputation day: 17.11.2022

DECLARATION OF AUTHORSHIP

I hereby declare on my honor that I have prepared this thesis independently, without the unauthorized assistance of third parties and without the use of other aids and literature than those indicated. Data and concepts taken directly or indirectly from other sources are marked with the source.

In the selection and evaluation of the following material, the persons listed below have assisted me in the manner described in each case, against no payment :

- Dr. Benedict Diederich (Leibniz IPHT Jena) with the continuous development of the UC2 toolbox and providing data for focus-quality evaluation.
- Kenta Temma / Prof. Dr. Fujita Katsumasa (University Osaka) with continuing the work on the DSAX-ISM system and providing processable data.
- Hauke Rehr / Prof. Dr. Rainer Heintzmann (FSU Jena) with preliminary work on Zeiss Airy Data processing and providing the data for further evaluation.
- Dr. Daniel Bremer / Prof. Dr. Raluca Niesner (Charitee Berlin) with preparing and measuring mouse retina fluorescence samples.

Further persons were not involved in the contentwise-material production of the present work. In particular, I have not used the paid help of intermediary or consulting services (doctoral advisors or other persons) for this purpose. No one has received from me, directly or indirectly, any pecuniary benefits for work related to the content of the thesis presented.

The thesis has not been submitted to any other examination authority in the same or similar form, either in Germany or abroad. I am aware of the applicable doctoral regulations of the Faculty of Physics and Astronomy. I declare on my honor that to the best of my knowledge I have told the truth and have not concealed anything.

Jena, 7.12.2022 — Place, date

René Lachmann — René Lachmann

To Claudia.

ABSTRACT

Recent developments in microscopic imaging led to a better understanding of intra- and intercellular metabolic processes and, for example, to visualize structural properties of viral pathogens. The developments in microscopy have greatly enriched and accelerated the pace of development in many other scientific disciplines. While the number of fundamentally different microscopy techniques has changed little in recent years, many have been greatly enhanced by numerical methods. Even approaches whereby a theoretical imaging problem is first solved using an algorithm and then an existing microscopy system are adapted or a new microscopy system is developed. While the variety of microscopy techniques represented in laboratories is slowly expanding fluorescence widefield as well as fluorescence confocal imaging techniques are the most widely used.

Existing microscopy techniques do not have to be completely replaced, but can be readily improved and renewed by partial extension of the setup and/or appropriate processing of the data. In this thesis, the imaging process of the above mentioned fundamental microscopy techniques is treated holistically to highlight general strategies and maximise their information content. Poisson or shot noise is assumed to be the fundamental noise process for the given measurements.

A stable focus position is a basic condition for e. g. long-term measurements in order to provide reliable information about potential changes inside the field of view (FoV). While newer microscopy systems can be equipped with hardware autofocus, this is not yet the widespread standard. For image-based focus analysis, different metrics for ideal, noisy and aberrated, in case of spherical aberration and astigmatism, measurements are presented. The experience gained *in-silico* is evaluated on real measurement data acquired using a 3D-printed microscope. The measured samples are 1) inorganic patterns and 2) fixed, non-living Henrietta Lacks Cervical Cancer (HeLa) cells.

A stable focus position is also relevant in the example of 2-photon confocal imaging and at the same time the situation is aggravated in the given example, the measurement of the retina in the living mouse. In addition to the natural drift of the focal position, which can be evaluated by means of previously introduced metrics, rhythmic heartbeat, respiration,

unrhythmic muscle twitching and movement of the mouse kept in artificial sleep are added. A dejittering algorithm is presented for the measurement data obtained under these circumstances. Further, *in-silico* data relevant to this biological example are generated, the dejitter algorithm is evaluated on the simulated and later on real measurement data.

The step from confocal to Image Scanning Microscopy (ISM) is achieved by simply replacing the detection pinhole by a detector array, e. g. a camera. Instead of just measuring the intensity by signal integration through the detection pinhole, the brightness distribution at the detector is correlated with the location information of the different camera pixels. Using the additional information about the sample distribution obtained in this way, a method for reconstructing 3D from 2D image data is presented in the form of thick slice unmixing. This method can further be used for suppression of light generated outside the focal layer of 3D data stacks and is compared to selective layer multi-view deconvolution. To reduce phototoxicity and save valuable measurement time for a 3D stack, the method of zLEAP is presented, by which omitted (Nyquist sampled) Z-planes are subsequently calculated and inserted. The results are evaluated on in-silico as well as on measurement data acquired on the commercially available Zeiss Airyscan.

ZUSAMMENFASSUNG

Durch Entwicklungen in der mikroskopischen Bildgebung konnten intra- und interzelluläre Stoffwechselprozesse besser verstanden und z. B. strukturelle Eigenschaften von viruellen Krankheitserregern sichtbar gemacht werden. Die Entwicklungen der Mikroskopie haben die Entwicklungsgeschwindigkeit vieler anderer wissenschaftlichen Disziplinen stark bereichert und beschleunigt. Während sich die Anzahl grundlegend verschiedener Mikroskopietechniken in den letzten Jahren kaum geändert hat wurden viele Techniken stark mittels numerischer Methoden erweitert. Mittlerweile rückt sogar der Software-First Ansatz in den Vordergrund, wobei ein theoretisches Bildgebungsproblem zuerst mittels eines Algorithmus gelöst und anschließend ein existierendes Mikroskopiesystem angepasst oder ein neues Mikroskopiesystem entwickelt wird. Während sich die in den

Laboren vertretene Vielfalt an Mikroskopietechniken langsam erweitert hat sind insbesondere die grundlegenden Techniken der Weitfeld- sowie konfokalen Fluoreszenzbildgebung noch immer am weitesten verbreitet.

Bestehende Mikroskopietechniken müssen nicht vollständig ersetzt werden, sondern können durch eine partielle Erweiterung des Aufbaus und/oder eine geeignete Verarbeitung der Daten leicht verbessert und erneuert werden. In dieser Arbeit wird der Abbildungsprozess der oben genannten grundlegenden Mikroskopietechniken ganzheitlich behandelt, um allgemeine Strategien aufzuzeigen und ihren Informationsgehalt zu maximieren. Für die Modellierung des Messprozesses wird von Poisson- oder Schrotrauschen als fundamentalen Rauschprozess ausgegangen.

Eine stabile Fokusposition ist eine Grundbedingung für z. B. Langzeitmessungen um verlässliche Aussagen über die Veränderungen im Sichtfeld (field of view (FoV)) treffen zu können. Während neuere Mikroskopiesystem mit hardware Autofokus ausgestattet werden können ist dies noch nicht der überall verbreitete Standard. Zur bildbasierten Fokusanalyse werden verschiedene Metriken für ideale, verrauschte und spherisch sowie astigmatisch aberrierte Messungen vorgestellt. Die *in-silico* gewonnenen Erfahrungen werden an mittels 3D-gedruckten Mikroskops aufgenommenen, realen Messdaten evaluiert. Die vermessenen Proben sind hierbei 1) anorganische Muster und 2) fixierte, nicht mehr lebende (HeLa-) Zellen.

Die Frage einer stabilen Fokusposition ist auch am Beispiel einer konfokalen 2- Photonenbildgebung relevant und gleichzeitig ist die Situation am gegebenen Beispiel, der Vermessung der Retina in der lebenden Maus, verschärft. Zusätzlich zum natürlichen Drift der Fokusposition, der mittels vorher eingeführten Metriken evaluiert werden kann, kommen rhythmischer Herzschlag, Atmung, unrhythmisches Muskelzucken und Bewegung der im künstlichen Schlaf gehalten Maus hinzu. Für die unter diesen Umständen gewonnenen Messdaten wird ein Dejitteralgorithmus vorgestellt, für dieses biologische Beispiel relevante in-Silico Daten generiert und der vorgestellte Dejitteralgorithmus an Simulations- sowie realen Messdaten evaluiert.

Der Schritt von der konfokalen hin zur Bildrastermikroskopie (Image Scanning Microscopy (ISM)) gelingt durch simples Ersetzen des Detektionspinholes durch ein Detektorarray, z. B. eine Kamera. Dadurch wird statt bloßer Intensitätsmessung mittels Signalintegration durch das Detektionspinhole die am Detektoranliegende Helligkeitsverteilung mit der Ortsinformation der verschiedenen Kamerapixel korreliert. Mittels der so gewonnenen

zusätzlichen Information über die Probenverteilung wird mit dem Thickslice Unmixing ein Verfahren zur Rekonstruktion von 3D aus 2D Bilddaten vorgestellt. Dieses Verfahren kann weiterhin zur Unterdrückung von Licht, das außerhalb der fokalen Schicht erzeugt wurde, von 3D-Datenstacks verwendet werden und wird mit einer selektiven Schicht Multiview Entfaltung verglichen. Zur Reduktion von Phototoxizität und sparen von wertvoller Messzeit für einen 3D-Stack wird das Verfahren des zLEAP vorgestellt, durch welches ausgelassene (Nyquist gesampelte) Z-Ebenen nachträglich ausgerechnet und eingefügt werden. Die Ergebnisse werden an In-Silico sowie am Zeiss Airyscan aufgenommenen Messdaten evaluiert.

ACKNOWLEDGEMENTS

With these lines I would like to thank many of the people that supported me over all these years. I would like to thank Prof. Dr. Rainer Heintzmann for giving me the chance to work within the field of microscopy. Dr. Jan Becker, Dr. Ronny Förster, Dr. Christian Karras, Robert Kretschmer and Dr. Ondrej Stranik for being very fun room-mates whose interaction with always led to fruitful discussions. Dr. Benedict Diederich, Dr. Swen Carlstedt and Barbora Marsikova for enjoying working with me on my biggest side-project: *UC2* which is not part of this thesis. Prof. Dr. Katsumasa Fujita (Osaka), Dr. Ryosuke Oketani and Kenta Temma for making my stay in Japan while doing researching on differential Saturated Excitation Image Scanning Microscopy (dSAX-ISM) fun and possible at all. Hauke Rehr for working with me on a first basic deconvolution comparison and acquisition of the Airyscan data. Dr. Daniel Bremer and Prof. Dr. Raluca Niesner for working together with me on mouse-retina imaging and providing data for further use within my thesis. I would like to thank all co-workers within the FA8 and at the IPHT that made the time fly by very fast.

I would like to thank the Octavians for all the musical input that was a welcome distraction from research. My good friends Michael Reichmann, Takayoshi Yamashita, Philipp Kahn and Christoph Schnedermann who kept my motivation high in good times as well as in bad times. My family, my son Nelio and especially my beloved wife Claudia, who supported me through all these years and without whom this thesis would never have come into being.

Finally, I would like to thank you my dear reader for investing your precious lifetime and try to understand my thoughts and results. This very time will never come back and I appreciate your investment. Thank you!

LIST OF FIGURES

LIST OF TABLES

NOTATION

Abbreviation	Meaning
	Imaging Methods
CLSM	Confocal Laser Scanning Microscopy
dSAX	differential Saturated Excitation
dSAX-ISM	differential Saturated Excitation Image Scanning Microscopy
ISM	Image Scanning Microscopy
SAX	Saturated Excitation Microscopy
SIM	Structured Illumination Microscopy
SMLM	Single Molecule Localization Microscopy
SPIM	Selective-Plane Imaging Microscopy

WF	Widefield Microscopy
	Imaging Vocabulary
ATF	Amplitude Transfer Function
AU	Airy Unit
APSF	Amplitude Point Spread Function
BFP	Back Focal Plane
BS	Beam Splitting Element
FoV	field of view
FWHM	Full Width at Half Maximum
ooF	Out of Focus
OTF	Optical Transfer Function
PSF	Point Spread Function
ROI	Region of Interest
SNR	Signal-to-Noise Ratio
	Functions
CoM	Center of Mass
FFT	Fast Fourier Transform
GF	Gaussian Function
GR	Good's Roughness
GS	Gradient Square
SGF	Savitzky-Golay Filter
SV	Singular Values
SVD	Singular Value Decomposition
TV	Total Variation
	Numeric Methods
DDA	Discrete Dipole Approximation
FDTD	Finite Difference Time Domain
FI	Fixpoint Iteration
GD	Gradient Descent
HAF	Hardware Autofocus
HDF5	Hierarchical Data Format Version 5
L-BFGS-B	Limited-memory Broyden-Fletcher-Goldfarb-Shanno algorithm
MAL	Maximum Anterior Likelihood

MAP	Maximum A Posteriori likelihood
NN	Neural Networks
NCC	Normalized Cross Correlation
nWAVG	noise normalized Weighted Averaging in Fourier Space
PiRe	Pixel Reassignment
RL	Richardson-Lucy Deconvolution
SAF	Software Autofocus
SSUM	Sheppard Summation
WAVG	Weighted Averaging in Fourier Space
TU	Thick slice Unmixing
WD	Wiener Deconvolution
zLEAP	LEAve-out z-Planes
	Deconvolutions
DEC	single-view DEConvolution
cDEC	confocal single-view DEConvolution
prDEC	pixel reassigned single-view DEConvolution
mDEC	multi-view DEConvolution
amDEC	multi-view DEConvolution using all available images
2d3dDEC	multi-view DEConvolution with 2D to 3D reconstruction
	Metrics
BRE	Brenner
KEN	Kristans Entropy
MAX	Maximum
SEN	Shannon Entropy
TEN	Tenenbaum Gradient
TOV	Total Variation
VAR	Variance
VO5	Vollath F5
VS4	Symmetric Vollath F4
	Electronic Vocabulary
ADC	Analog-Digital-Conversion Unit
AOM	Acousto Optical Modulator
APD	Avalanche Photo Diode

CCD	Charge-Coupled Device
CPU	Central Processing Unit
emCCD	electron-multipliying Charge-Coupled Device
FPGA	Field Programmable Gate Array
fps	frames per second
GaAsP	Gallium-Arsenide Phosphate
GM	Galvanometric Mirror pair
GPU	Graphics Processing Unit
OS	Operating System
PD	Photodiode
PMT	Photo-Multiplying Tube
RAM	Random-Access Memory
RMS	Root-Mean-Square
sCMOS	Scientific-Metaloxide Semiconductor
SPAD	single photon avalanche diode
STO	Storage

Chemistry and Biology

FRET	Förster-Resonance Energy Transfer
HeLa	Henrietta Lacks Cervical Cancer
HOMO	Highest Occupied Molecular Orbital
IC	Internal Conversion
ISC	Intersystem Crossing
LUMO	Lowest Unoccupied Orbital
U2OS	Human Bone Osteosarcoma Epithelial Cells
VR	Vibrational Relaxation

NOMENCLATURE

Symbol	Example	Meaning
$(\cdot)$	$z^{(o)}$, $\widetilde{M}^{(\mathrm{ISM})}$, $d^{(\mathrm{SPL,ax})}$	mark different (spatial/frequency) positions or variable names

l, m, n	M_{lmn}, x_m	matrix, vector and general indices; for the 5 used data-axes typically chosen as: M_{nalmn} for n along the non-linearity (N-), a along the view (A-), l along the X-, m along the Y- and n along the Z-axis
$\tilde{l}, \tilde{m}, \tilde{n}$	$\tilde{M}_{\tilde{l}\tilde{m}}$, $\tilde{\hat{S}}_{\tilde{n}}$	indices for matrices in Fourier space, where $\tilde{l}$ is along k_x-, $\tilde{m}$ is along k_y- and $\tilde{n}$ is along k_z axis
α, β	$\alpha = \pi/2$	used as general angles in radians
$\chi = \lvert\vec{\chi}\rvert$	$\chi = \lvert\vec{\chi}\rvert$	to simplify display of calculations using the abs-value of a vector the vector-arrow is omitted together with the abs-operator

SYMBOLS

Symbol	Meaning
α	half-opening angle
NA	numerical aperture
$n^{(\mathrm{im})}$	refractive index on example of the immersion between objective and cover slip
a	amplitude point-spread function
h	intensity point-spread function
$\tilde{a}$	amplitude transfer function
$\tilde{h}$	optical transfer function
$\mathcal{F}$	(Forward) Fourier transformation
$\mathcal{F}^{-1}$	Inverse Fourier transformation
supp $\{f\}$	support, i. e. non-zero region of a function f in (Fourier Space)
$\vec{x}$	cartesian coordinates, $\vec{x} = (x, y, z) = (\vec{\chi}, z)$
$\vec{\chi} = (x, y)$	lateral (cartesian) coordinate-vector
$\vec{k}$	frequency vector in Fourier-Space; in analogy to $\vec{x}$ defined as $\vec{k} = (k_x, k_y, k_z) = (\vec{\kappa}, k_z)$; $\lvert\vec{k}\rvert = \frac{2\pi n}{\lambda}$
$\vec{\kappa} = (k_x, k_y)$	lateral frequency vector in Fourier-Space
$d^{(\mathrm{NAME})}$	pixel pitch (typically in nm) w. r. t. NAME, so if e. g. NAME=SPL,ax then $d^{(\mathrm{SPL,ax})}$ the axial sampling pitch is meant
I	Intensity

Φ	(Photon-) Flux
S	Sample
$\hat{S}$	reconstructed sample estimate
M	noisy measured (n-dimensional) image
C	NCC
N	number/amount of pixels, photons, iterations, ...; e. g. $N^{(\mathrm{Phot})}$ amount of photons or $N^{(\mathrm{x})}$ amount of pixels along X-direction
∂_{S_n}	$\partial_{S_n} = \frac{\partial}{\partial S_n}$ partial derivative with respect to variable S_n; $\partial_{\vec{x}} = \sum_i \partial x^{(\mathrm{i})}$ means the vectorial derivative w. r. t. basis vectors $x^{(\mathrm{i})}$

CONTENTS

1

GENERAL INTRODUCTION

Physicists are made of atoms. A physicist is an attempt by an atom to understand itself.

— Michio Kaku

1.1 MOTIVATION

Microscopy is a cornerstone of everyday research and production in a wide range of different disciplines. It can be used to resolve structures that lie far below the spatial resolution limit and spectral and dynamic sensitivity of the eye. It thus enables more detailed analysis of microscopic processes and opens up space for new questions and insights. For example, for structural analysis of cancer cells resolutions well below 500 nm are required, but field of view (FoV)s of at least 1 cell, i.e. $\geq 10\,\mu m$, are necessary [1]. Once the general mapping of such an object is achieved, functional and time-dependent properties can be investigated. The goal of this work is to assure imaging-quality of the recorded data, e. g. for structural analysis of the sample, and to optimize existing methods.

Abbe's mathematical analysis of diffraction-limited resolution of optical systems has allowed to predict and analyze limiting effects of optical components [2]. Remarkably, Abbe established a resolution limit, based on physical parameters of the optical system and assuming the absence of any aberrations, which surpassed all other existing approaches at the time and is still valid today. Abbe's resolution limit describes the support-boundaries of the transfer function of the analyzed optical system and can neither be exceeded nor worsened by aberrations. Hence, it is a fundamental limit that allowed to replace the hitherto common „trial and error"method by precise precalculations and led to the production of high-quality optical-imaging devices.

The development of new optical and electrical components led to the advent of microscopy systems using wavelengths beyond the spectral sensitivity (VIS) of the eye, 380 nm − 750 nm [3], e. g. electron microscope. However, only microscopy techniques with application in VIS are relevant to this work. The development of the confocal mi-

croscope by Marvin Minsky in 1957 revolutionized fluorescence widefield imaging by introducing optical sectioning [4]. Minsky used a divergent arc lamp and thus did not work in 4f configuration (see Section A.4), but already used dichroic mirrors, a detection pinhole and sample scanning. He illuminated a small volume of the sample partial coherently and imaged it onto a detection pinhole while using the same objective lens for illumination and detection. The discovery and subsequent commercial proliferation of the laser in 1960 led to changes in confocal design to 4f or 6f (Köhler illumination) systems. Since then, confocal systems are mostly operated in laser scanning mode, which is why they are also referred to as Confocal Laser Scanning Microscopy (CLSM).

In 1988 Sheppard describes how the detection pinhole together with the detector could be replaced by an array of (point-like) detectors and suitably summed [5], i. e. Pixel Reassignment (PiRe). Müller & Enderlein published the first working experimental implementation of Image Scanning Microscopy (ISM) in 2010 [6]. Steady advances of this technology brought parallelized approaches with analog processing (2013: instant Structured Illumination Microscopy (iSIM) [7], Re-Scan [8], Optical-Photon-Reassignment (OPRA) [9], 2015: spinning-disk ISM [10]) as well as digital processing (2013: spinning-disk ISM [11], 2016: REfocusing after SCanning using Helical phase engineering (RESCH) [12]). The major advantage of ISM over classical CLSM is that the pinholing effect is maintained by the individual (point) detectors while no light is blocked by a physical pinhole. By reassignment, the additionally detected photons can be suitably combined and this leads to a superconcentration effect, since more photons are reassigned to the same pixel position than would have been the case in widefield or CLSM [13].

Although the confocal microscope was established as the gold standard in the biological community quickly after its discovery, new solutions in the form of microscopy systems have evolved in parallel to address increasingly challenging questions. Methods of fluorescence microscopy have been (further) developed with particular success. They can be divided into linear and non-linear methods with respect to the sample response to the excitation field. Non-linearity is usually achieved by saturation or excitation of higher harmonics in the sample response, which is why almost any linear system can be directly converted into a non-linear system. Furthermore, the methods can be divided into wide-field and scanning methods, such as the CLSM. Most methods agree on the fact that the axial dimension is recorded as a (z-/axial-) stack slice by slice. Despite all further and

new developments, the CLSM is still a standard system in many laboratories. This can be explained, among other things, by its design, ease of calibration, ease of modification, comparatively low construction costs and wide availability of equipment and protocols.

In order to modify existing ISM setups and to use their data more efficiently through improved computational reconstruction, a deep understanding of the underlying theoretical approaches is necessary. This knowledge will be conveyed in the following sections and the nomenclature used will be introduced step by step.

For a better understanding of the advance made in this thesis a foundation of the theory of image formation and interaction of light with matter is important. The main part of the thesis uses predominantly high-level descriptions and abbreviations of imaging operators. A brief compilation of fundamental equations, derivations and assumptions will be presented. The aim of this section is to describe the imaging process in the case of fluorescent imaging. For this purpose, suitable field propagators are first introduced for propagation e. g. from the sample through the optical measurement system to the detector. Further, necessary technical terms and equations for the used setup geometries (e. g. ISM) as well as resolution criteria are presented.

For a more thorough derivation see Appendix A and the splendid work by Mertz [14], Goodman [15] and Born & Wolf [16].

1.2 USEFUL EQUATIONS AND FORMALISM

The Fourier-Transformation throughout this thesis and is defined as:

$$\text{forward:} \qquad F\left(\vec{k}\right) \equiv \quad \mathcal{F}\left\{f(\vec{x})\right\}\left(\vec{k}\right) \equiv (2\pi)^{-3/2} \int_{\mathbb{R}^3} f(\vec{x})\, e^{-i\vec{k}\vec{x}} d\vec{x} \tag{1.1}$$

$$\text{backward:} \qquad f(\vec{x}) \equiv \mathcal{F}^{-1}\left\{F\left(\vec{k}\right)\right\}(\vec{x}) \equiv (2\pi)^{-3/2} \int_{\mathbb{R}^3} F\left(\vec{k}\right) e^{i\vec{k}\vec{x}} d\vec{k} \tag{1.2}$$

and $i = \sqrt{-1}$ the imaginary unit. Further, the convolution operation $\otimes$ is defined as:

$$[f \otimes g](x) \equiv \int_{\mathbb{R}^3} f(\vec{x}')g(\vec{x} - \vec{x}')d\vec{x}' \tag{1.3}$$

The convolution is typically interpreted to span the whole 3D real space and hence $\otimes = \otimes_{\vec{x}}$. If the operation is limited to a different domain, e. g. the lateral 2D plane, then this is marked by a representative vector of this domain, hence e. g. by appending $\vec{\chi}$ on the operator to form $\otimes_{\vec{\chi}}$.

For an arbitrary set $x_1, x_2, \ldots, x_n$, $n \geq 3$ of realizations of a random variable X the symbol $\forall$ means that a statement holds *for all* elements (here: realizations of X) taken from this set.

For two discrete, jointly distributed random variables X, Y with random distributions P_X, P_Y and at least partially non-zero joint distribution P_{XY} the adopted nomenclature is [17]:

$$\textbf{Expectancy} \quad \mathbb{E}\{X\} = \sum_x X \cdot P_X(X = x) \tag{1.4}$$

$$\textbf{Variance} \quad \mathbb{V}\{X\} = (\mathbb{E}\{X\})^2 - \mathbb{E}\left\{X^2\right\} \tag{1.5}$$

$$\textbf{Standard Deviation} \quad \mathbb{S}\{X\} = \sqrt{\mathbb{V}\{X\}} \tag{1.6}$$

$$\textbf{Co-Variance} \quad \mathbb{W}\{X, Y\} = \mathbb{E}\{[X - \mathbb{E}\{X\}][Y - \mathbb{E}\{Y\}]\} \tag{1.7}$$

Within this thesis, the words *expectancy* and *mean* will be used interchangeably. The extension to complex random variables is straightforward [18].

To determine the distance $\Delta M = M^{(1)} - M^{(2)}$ between two images or matrices $M^{(1)}$, $M^{(2)}$ the l_p norm [19] is used in this work:

$$\|\Delta M\|_p^p \equiv \left[\sum_i^{N^{(\mathrm{i})}} |\Delta M_i|^p\right]^{1/p} \tag{1.8}$$

with $N^{(\mathrm{i})}$ the number of pixels of used. The distance Equation 1.8 thus determines the $1/p$ exponentiated sum of the p exponentiated amounts of all pixels of the difference matrix ΔM. Note, that even though M is a (potentially N-dimensional) matrix it is flattened in Equation 1.8 for brevity.

1.3 THE IMAGING PROCESS

In most of this thesis' analyses the sample as well as the imaging process will be treated as 3D while the detector axially cuts out one slice of the incident field, i. e. sharp in-focus image information together with Out of Focus (ooF) blur is measured.

The spreading of the response signal of an illuminated point-like sample through-out the optical system can be characterized by the Amplitude Point Spread Function (APSF) in case of coherent or Point Spread Function (PSF) h in case of incoherent, e. g. fluorescent, light. The Fourier transformed APSF is called Amplitude Transfer Function (ATF) and

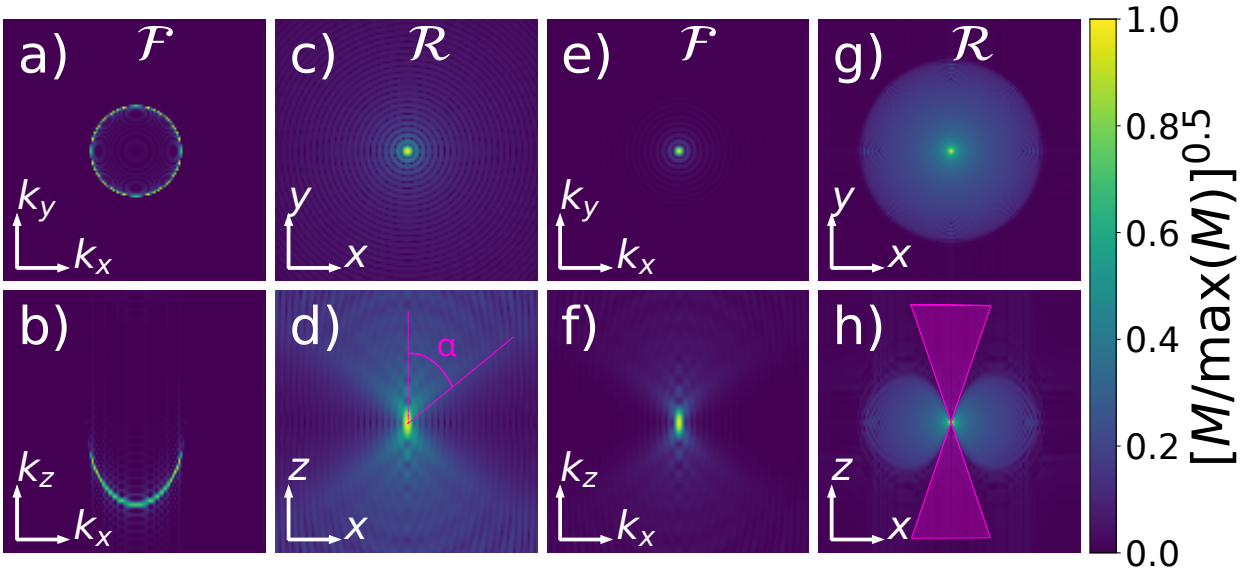


FIGURE 1.1: **Fourier-Relations in 3D.** a,b,g,h) in Fourier space and c-f) in real space coordinates. ATF at cut a) $k_z = 0$, b) $k_y = 0$; APSF at cut c) $z = 0$, d) $y = 0$; PSF at cut e) $z = 0$, f) $y = 0$; OTF at cut g) $k_z = 0$, h) $k_y = 0$. d) shows the half-opening angle α of the objective lens and h) the missing cone in magenta. All panels are normalized to individual maximum and have a gamma of 0.5. $\log_{10}(1 + \tilde{h})$ normalization is applied to g+h) beforehand.

the Fourier transformed PSF is called Optical Transfer Function (OTF) $\tilde{h}$. For analysis of transferable Fourier spatial frequencies by an optical system the extent of the non-zero region of the OTF, the so called OTF support $\operatorname{supp}\left\{\tilde{h}\right\}$, is of high-interest. For a more detailed introduction see Section A.2.

In the following $\vec{x} = [x, y, z]$ is the continuous 3D spatial coordinate vector, $\vec{\chi} = [x, y]$ is the lateral coordinate vector, $\vec{k} = [k^{(x)}, k^{(y)}, k^{(z)}]$ is the 3D spatial frequency vector and $\vec{\kappa} = [k^{(x)}, k^{(y)}]$ is the lateral spatial frequency vector.

Given a fluorescent density distribution $S\left(\vec{x}^{(o)}\right)$, typically called *sample*, residing at a position $z = z^{(o)}$ and being excited with an excitation intensity distribution $I^{(ex)}\left(\vec{x}^{(o)}\right)$ the emitted irradiance $I^{(o)}\left(\vec{x}^{(o)}\right) = I^{(ex)}\left(\vec{x}^{(o)}\right) \cdot S\left(\vec{x}^{(o)}\right)$ at $z = z^{(o)}$ can be measured if the sample is axially (Z-) scanned. The propagation of irradiance $I^{(o)}$ through the optical detection system can be modeled by convolution (Equation 1.3) of $I^{(o)}$ with the detection system's PSF $h^{(det)}$, thereby yielding the measurable Z-scanned irradiance distribution at $I^{(1)}$ $z = z^{(1)}$:

$$I^{(1)}\left(\vec{x}^{(1)}\right) = \left[h^{(det)} \otimes I^{(o)}\right]\left(\vec{x}^{(1)}\right) \tag{1.9}$$

Fourier transformation of Equation 1.9 yields the Fourier frequency spectrum of the irradiance distribution at the detector at $z = z^{(1)}$:

$$\tilde{I}^{(1)}\left(\vec{k}\right) = \tilde{h}^{(\mathrm{det})}\left(\vec{k}\right)\tilde{I}^{(0)}\left(\vec{k}\right) \tag{1.10}$$

Note, that in Equation 1.10 the spectrum at every $\vec{k}$ is calculated by a point-wise multiplication of the OTF $\tilde{h}^{(\mathrm{det})}$ with the emission spectrum $\tilde{I}^{(0)}\left(\vec{k}^{(0)}\right) = \left[\tilde{I}^{(\mathrm{ex})} \otimes_{\vec{k}} \tilde{S}\right]\left(\vec{k}^{(0)}\right)$.

Figure 1.1 displays the relation between 3D APSF (a+b), ATF (c+d), PSF (e+f) and OTF (g+h) in case of a single lens. The first and third column are defined in Fourier space while the second and forth column use real space. The half objective lens' opening angle α is visualized in panel d). Panel h) demonstrates the basic information transfer problem of a single lens at low lateral spatial frequencies, the so-called *missing cone*. The OTF support supp $\left\{\tilde{h}\right\}$ is axially not extended and hence these frequencies cannot be distinguished/resolved. This leads to ooF blur by low-frequency information (e. g. a fluorescent plane) not only of adjacent but of all axial slices of the Z-stack. Naively, every wave that can be detected by an optical system is automatically resolved. In fact, this statement is wrong in many respects. The core idea was already formulated by Ernst-Abbe [2]. He finds that to resolve a scattering object through a lens / optical system, at least two diffraction orders must be intercepted and interfered. Classically, the ballistic field (0th order) and the $\pm 1^{(\mathrm{st})}$ order of scattering are often used here. While Abbe in case of illuminating the sample without an angle (=central illumination) finds the lateral resolution limit $d^{(\text{Abbe central})} = \frac{\lambda}{NA}$, it is the case of „extreme oblique illumination" [2] that became famous and the lateral limit he found is:

$$d^{(\text{Abbe,coherent})} = \frac{\lambda}{NA^{(\mathrm{cond})} + NA^{(\mathrm{det})}} \tag{1.11}$$

$$d^{(\mathrm{Abbe})} = \frac{\lambda}{2NA} \tag{1.12}$$

where $NA = n\,\sin(\alpha)$, n is the immersion medium and α is the half aperture angle of the microscope objective. Equation 1.11 is also called „Abbe's diffraction limit" in case of coherent and Equation 1.12 in case of incoherent (e. g. fluorescent) imaging, because it describes the maximum resolvable spatial frequency of a diffraction limited optical system. Thus the maximum lateral frequency $k^{(\mathrm{Abbe})}$ or lateral cut-off frequency $k^{(\mathrm{C})}$ in Fourier space can be derived via:

$$k^{(\mathrm{C})} \equiv k^{(\mathrm{Abbe})} = \frac{2\pi}{d^{(\mathrm{Abbe})}} \tag{1.13}$$

The limit can be partially overcome using *a priori* assumptions about the sample type, distribution, or behavior (see Section 1.7). The systems analyzed in this thesis mainly use a highly coherent illuminating laser that is mostly focused on an (incoherently interacting) fluorescent sample (see Section 1.5).

Limit Name	Measure
Airy Unit	$d^{(1\mathrm{AU})} = 2.44\, d^{(\mathrm{Abbe})}$
FWHM	$d^{(\mathrm{FWHM})} \approx 1.02\, d^{(\mathrm{Abbe})}$
Sigma Gauss	$\sigma^{(\mathrm{cG})} \approx 0.42\, d^{(\mathrm{FWHM})}$

TABLE 1.1: **Resolution measures.** Set of resolution measures used within this thesis. Inspired by [20] and [21].

The exact formulation of the circumstances around the measuring apparatus as well as the sample are essential for the definition of a suitable resolution measure. The measures used in this thesis are given in Table 1.1 and based on the assumptions of a *simultaneous measurement*, no *statistical recalculation* and *unknown but equally bright sample structure*. One Airy unit $d^{(1\mathrm{AU})}$ is the disc-diameter of the central non-zero region, i. e. from the $-1^{(\mathrm{st})}$ to the $+1^{(\mathrm{st})}$ minimum, of the PSF. $d^{(\mathrm{FWHM})}$ is the diameter of the in-focus PSF central peak at half of its maximum value. $\sigma^{(\mathrm{cG})}$ is the standard deviation standard deviation $\mathbb{S}\left\{h^{(\mathrm{G})}\right\}$ of a Gaussian $h^{(\mathrm{G})}$ that fits the in-focus (i. e. 2D) PSF h in case of adaptive offset clipping w. r. t. the PSF maximum.

For an ideal or non-noise limited imaging process the Full Width at Half Maximum (FWHM) is an established measure for comparison and typically bead or line-samples are used for characterization. In case of noise-limited imaging the cut-off frequency is predominantly determined by the noise-floor. Here, Fourier-Ring-Correlation [22] for estimation of the cut-off frequency is a promising approach. If the underlying statistics for describing the photon nature can be represented by a Poisson-process the general noise-limitation can be overcome by simply collecting more photons (see Section 1.7).

1.4 WIDEFIELD- AND CONFOCAL IMAGING

Building on the findings in the previous sections, any optical system can be described for both coherent (Equation A.51) and incoherent imaging (Equation A.52) using the respective PSF or OTF. Since most imaging methods in the context of this work use fluorescent samples, henceforth incoherent imaging will be described.

During image recording, the intensity distribution $I^{(1)}$ (Equation 1.35) present in the detector plane is time- and space integrated (i. e. discretized) and converted into digital units. Thus, for a simplified description, the sample is viewed as consisting of a set of *discretized voxels*. The forward model appropriate for the imaging method of interest can straightforwardly be transferred to the discretized representations of the sample S, the system PSF h and the ideal (i. e. noise-free) measured image $\breve{M}$, which represents the discretized version of $I^{(1)}$ in case of an axial sample scan.

In case of fluorescence widefield microscopy (i. e. non-scanning) mode the ideal image $\breve{M}$ directly follows from Equation 1.9 and Equation 1.30 as:

$$\breve{M}^{(\mathrm{WF})}\left(\vec{x}^{(1)}\right) = \int d\vec{x}^{(0)}\, h^{(\mathrm{WF})}\left(\vec{x}^{(1)} - \vec{x}^{(0)}\right) S\left(\vec{x}^{(0)}\right) = \left[h^{(\mathrm{WF})} \otimes S\right]\left(\vec{x}^{(1)}\right) \tag{1.14}$$

This mode is called Widefield Microscopy (WF). The continuous convolution operation of the forward model (Equation 1.9) can be approximated as:

$$\overrightarrow{\breve{M}} = \mathbf{H} \cdot \vec{S} \tag{1.15}$$

$$\breve{M}_l = \sum_m H_{ml} \cdot S_m \qquad \forall l \tag{1.16}$$

where the 3D ideal image $\breve{M}$ and sample S have been flattenend into 1D-vectors $\overrightarrow{\breve{M}} \in \mathbb{R}^l$ and $\vec{S} \in \mathbb{R}^m$ such that the multiplication with the kernel matrix $\mathbf{H} \in \mathbb{R}^{m \times l}$ could be written more compact. Equation 1.16 is a component-wise representation of Equation 1.15.

The change from fluorescence WF to CLSM imaging is accomplished by focusing the illumination and scanning the sample laterally to store spatially resolved brightness information per scan position. For fluorescence CLSM, together with the assumption of the sample being a 3D-stack of *thin slices* Equation 1.9 can be used. The pinhole is located in the detector plane and assumed to have no axial extent. Thus the 3D pinhole function $B(\vec{\chi}^{(1)}, dz)$ of infinitesimal axial extent dz will be used as a 2D function $B(\vec{\chi}^{(1)})$. By changing the lateral extent and position of the pinhole, the actual imaged confocal focal

volume size and position, i. e. imaged volume of the sample, can be changed. A detector located directly behind the pinhole thus detects the spatially summed (integrated) flux:

$$\begin{aligned}\Phi^{(1)} &= \int d^2\vec{\chi}^{(1)}\, B\left(\vec{\chi}^{(1)}\right) I^{(1)}\left(\vec{x}^{(1)}\right)|_{z^{(1)}=0} \\ \Phi^{(1)} &= \Phi^{(\mathrm{ex})} \iiint d^2\vec{\chi}^{(1)} d^2\vec{\chi}^{(0)} dz^{(0)} B\left(\vec{\chi}^{(1)}\right) h^{(\mathrm{det})}\left(\vec{\chi}^{(1)} - \vec{\chi}^{(0)}, z^{(0)}\right) \\ &\quad \cdot h^{(\mathrm{ex})}\left(\vec{\chi}^{(0)}, z^{(0)}\right) S\left(\vec{\chi}^{(0)}, z^{(0)}\right) \end{aligned} \tag{1.17}$$

where $I^{(\mathrm{ex})}\left(\vec{x}^{(0)}\right) = \Phi^{(\mathrm{ex})} h^{(\mathrm{ex})}\left(\vec{x}^{(0)}\right)$ and $\Phi^{(\mathrm{ex})} = \int d^3\vec{x} I^{(\mathrm{ex})}(\vec{x})$. Since in this imaging method the spatial sample information is lost, due to the lateral integration, the sample must be scanned such that an image is gained by correlation of scanning positions with the measured fluxes. Through a scanning unit, the illumination is scanned and the detection is simultaneously descanned, thus with Equation 1.17 the ideal CLSM image $\breve{M}^{(\mathrm{CLSM})}$ and respective effective PSF $h^{(\mathrm{CLSM})}$ as well as OTF $\tilde{h}^{(\mathrm{CLSM})}$ are obtained according to:

$$\begin{aligned}\breve{M}^{(\mathrm{CLSM})}\left(\vec{x}^{(\mathrm{s})}\right) &= \Phi^{(\mathrm{ex})} \int d^3\vec{x}^{(0)}\, h^{(\mathrm{CLSM})}\left(\vec{x}^{(0)} - \vec{x}^{(\mathrm{s})}\right) S\left(\vec{x}^{(0)}\right) \\ &= \Phi^{(\mathrm{ex})}\left[h^{(\mathrm{CLSM})} \otimes S\right]\left(\vec{x}^{(\mathrm{s})}\right) \end{aligned} \tag{1.18}$$

$$\begin{aligned} h^{(\mathrm{CLSM})}(\vec{x}^{(0)} - \vec{x}^{(\mathrm{s})}) &= \int d^2\vec{\chi}^{(1)} B\left(\vec{\chi}^{(1)} - \vec{\chi}^{(\mathrm{s})}\right) h^{(\mathrm{det})}\left(\vec{\chi}^{(1)} - \vec{\chi}^{(0)} - \vec{\chi}^{(\mathrm{s})}, z^{(0)} - z^{(\mathrm{s})}\right) \\ &\quad \cdot h^{(\mathrm{ex})}\left(\vec{\chi}^{(0)} - \vec{\chi}^{(\mathrm{s})}, z^{(0)} - z^{(\mathrm{s})}\right) \\ &= \left[\left[B \otimes_{2D} h^{(\mathrm{det})}\right] h^{(\mathrm{ex})}\right]\left(\vec{x}^{(0)} - \vec{x}^{(\mathrm{s})}\right) \end{aligned} \tag{1.19}$$

$$\tilde{h}^{(\mathrm{CLSM})}(\vec{k}) = \left[\left(\tilde{B}\,\tilde{h}^{(\mathrm{det})}\right) \otimes_{3D} \tilde{h}^{(\mathrm{ex})}\right](\vec{k}) \tag{1.20}$$

Note that $h^{(\mathrm{ex})}$ and $h^{(\mathrm{det})}$ are completely sum normalized to 1 and can be interpreted as excitation and detection probability density functions. Hence, their product $h^{(\mathrm{CLSM})}$ resembles their joint probability density function which is the product of the two. While Equation 1.17, before the integration due to detection, still reveals information about the spatial distribution of the sample, Equation 1.18 only holds a brightness value at one scan position and needs to be correlated e. g. with the positions of a scanning process.

Unlike in case of the incoherent WF system the OTF is a convolution of excitation and detection OTF (Equation 1.20), which could even double the OTF-support $\mathrm{supp}\left\{\tilde{h}\right\}$ in the case of a (theoretically) Stokes shift free fluorescence. Similarly, it can be seen that a finite-size pinhole cannot change the maximum possible Fourier support, but can only manipulate the transfer efficiency of individual frequencies. Combining pinholes of different shapes/sizes can be found to enhance the relative contrast of individual frequency bands to each other thereby increasing perceived resolution [23]. A single

spatially extended pinhole can only degrade the overall transfer contrast, compared to the ideal transfer when using a δ-like pinhole, with the limiting case being a completely open pinhole which is the so-called *scanned Widefield* (sWF) case.

For the OTF support $\operatorname{supp}\left\{\tilde{h}^{(\mathrm{CLSM})}\right\}$, we find the lateral $\kappa^{(\mathrm{Abbe})}$ as well as axial $k_z^{(\mathrm{Abbe})}$ maximum frequency to be:

$$\kappa^{(\mathrm{Abbe,CLSM})} = \frac{4\pi NA}{\lambda^{(\mathrm{h})}} \qquad , \lambda^{(\mathrm{h})} = \frac{\lambda^{(\mathrm{ex})}\,\lambda^{(\mathrm{det})}}{\lambda^{(\mathrm{ex})} + \lambda^{(\mathrm{det})}} \tag{1.21}$$

$$k_z^{(\mathrm{Abbe,CLSM})} = \frac{4\pi NA\,(1-\cos\alpha)}{\lambda^{(\mathrm{h})}\sin\alpha} \tag{1.22}$$

where $\lambda^{(\mathrm{h})}$ is the harmonic sum of $\lambda^{(\mathrm{ex})}$ and $\lambda^{(\mathrm{det})}$ [24]. Unlike the WF 3D PSF $h^{(\mathrm{WD})}$ (Equation 1.14), whose axial intensity profile is very different from the lateral distribution, the CLSM 3D PSF $h^{(\mathrm{CLSM})}$ (Equation 1.19) can be analyzed fundamentally well by only using its central (1^{st}) Airy-unit and approximating it with a 3D Gaussian blob[1]. In doing so, one finds that the lateral standard deviation $\sigma^{(\mathrm{CLSM,lat})}$ of this approximation is at most $\sqrt{2}$ smaller than $\sigma^{(\mathrm{WF,lat})}$. Given, that σ can be converted into FWHM (Table 1.1) the same argument holds for the FWHM, even though the support of the OTF doubles.

By using the pinhole in the confocal microscope, high axial sectioning, due to ooF rejection, is exchanged for efficient use of the intensity distribution arriving at the detector. To be able to detect all photons, in Image Scanning Microscopy (ISM) the single detector is replaced by a detector array (e. g. a camera) and the physical pinhole is omitted. This changes Equation 1.18 to:

$$\begin{aligned} M_a^{(\mathrm{ISM})}\left(\vec{x}^{(\mathrm{s})},\vec{x}^{(\mathrm{a})}\right) &\equiv \Phi_a^{(1)}\left(\vec{x}^{(\mathrm{s})},\vec{x}^{(\mathrm{a})}\right) \\ \Phi_a^{(1)}\left(\vec{x}^{(\mathrm{s})},\vec{x}^{(\mathrm{a})}\right) &\equiv \Phi^{(\mathrm{ex})}\left[h_a^{(\mathrm{CLSM})} \otimes S\right]\left(\vec{x}^{(\mathrm{s})},\vec{x}^{(\mathrm{a})}\right) \qquad (1.23) \\ h_a^{(\mathrm{ISM})}\left(\vec{x}^{(\mathrm{s})},\vec{x}^{(\mathrm{a})}\right) &= \left[h_a^{(\mathrm{det})} \cdot h^{(\mathrm{ex})}\right]\left(\vec{x}^{(\mathrm{s})},\vec{x}^{(\mathrm{a})}\right) \qquad (1.24) \\ h_a^{(\mathrm{det})}\left(\vec{x}^{(\mathrm{s})},\vec{x}^{(\mathrm{a})}\right) &= \left[B_a \otimes h^{(\mathrm{det})}\right](\vec{x}^{(\mathrm{s})}) \qquad (1.25) \end{aligned}$$

Note, that Equation 1.18 can be derived from Equation 1.23 by restricting to the center pinhole only, e. g. in the case of $N^{(\mathrm{a})}$ pinholes thus choosing $a^{(\mathrm{c})} = \left\lfloor \frac{N^{(\mathrm{a})}}{2} \right\rfloor$ and $\vec{x}^{\left(\left\lfloor \frac{N^{(\mathrm{a})}}{2} \right\rfloor\right)} = 0$. Here $\lfloor\cdot\rfloor$ selects the lower rounded integer. In Equation 1.25, B_a was introduced to be able to model the differences of the individual camera pixels (e. g. shape and detection efficiency). For the simulations all camera-pixels where assumed to be of equal shape and efficiency.

1 A repeated convolution of a function $f^{(1)}$ with a second function $f^{(2)}$ converges towards a Gaussian shaped result $f^{(3)}$, see central limit theorem [17].

The OTF support supp $\left\{\tilde{h}^{(\mathrm{ISM})}\right\}$ of ISM is identical to the support supp $\left\{\tilde{h}^{(\mathrm{CLSM})}\right\}$ of the CLSM system, but so far we have only detected additional ooF information by means of neighboring pixels and thus rather reduced the foreground/background ratio. Interpreting $h^{(\mathrm{CLSM})}$ again as a joint-probability density function and asking where to find the most likely position of the fluorescent emitter given excitation $\vec{x}^{(\mathrm{s})}$ and detection position $\vec{x}^{(\mathrm{a})}$, this can be determined using a Center of Mass (CoM) (see Equation 3.5) approach. Once these positions are known, the information stored at the detector positions can be shifted back to the most likeli emission positions with a summation over all views afterwards. This processing strategy is called PiRe and the final image $M^{(\mathrm{rISM})}$ can be calculated to:

$$M^{(\mathrm{rISM})}\left(\vec{x}^{(\mathrm{s})}\right) = \sum_a M_a^{(\mathrm{ISM})}\left(\vec{x}^{(\mathrm{s})}, \vec{x}^{(\mathrm{a})}\right) \otimes \delta\left(\vec{x}^{(\mathrm{s})} - m^{(\mathrm{ISM})}\vec{x}^{(\mathrm{a})}\right) \tag{1.26}$$

$$\tilde{h}^{(\mathrm{rISM})}\left(\vec{k}^{(\mathrm{s})}\right) = \sum_a \left[\tilde{h}_a^{(\mathrm{det})} \cdot \tilde{h}^{(\mathrm{ex})} \cdot e^{ikm^{(\mathrm{ISM})}x^{(\mathrm{a})}}\right]\left(\vec{k}^{(\mathrm{s})}\right) \tag{1.27}$$

with the effective reassigned ISM OTF $\tilde{h}^{(\mathrm{ISM})}\left(\vec{k}^{(\mathrm{s})}\right)$ and $m^{(\mathrm{ISM})}$ the reassignment shift factor. Under the assumption, that the effective PSF $h_a^{(\mathrm{ISM})}\left(\vec{x}^{(\mathrm{s})}, \vec{x}^{(\mathrm{a})}\right)$ (Equation 1.24) of each view can be suitably approximated for a practical measurement processes by a Gaussian distribution with negligible errors, all $h_a^{(\mathrm{det})}$ and $h^{(\mathrm{ex})}$ can be defined by a Gaussian function. Then, if additionally $\lambda^{(\mathrm{ex})} = \lambda^{(\mathrm{em})}$ the most probable emission position is $\vec{x}^{(\mathrm{m,a})} = \vec{x}^{(\mathrm{a})}/2$, corresponding to a shift factor of $m^{(\mathrm{ISM})} = 0.5$. A brief demonstration of PiRe in case of simple transition from CLSM to ISM by using a camera instead of a Photo-Multiplying Tube (PMT) is given in Section 4.2.1.

The reassignment can be applied fully analog/optical [9] or digital [6] and leads to a densification of the photon distribution which is also called superconcentration [13]. Since the theoretical description by Sheppard in 1988 [5], many different ISM configurations have evolved, such as e. g. parallelization by multi-focal scanning and descanning with two Digital Micromirror Devices (DMDs, [25]), deeper sample penetration using 2-photon single-focus [26], video-framerate imaging [10], or 3D reconstruction using 2D measurements and detection-psf manipulation [12].

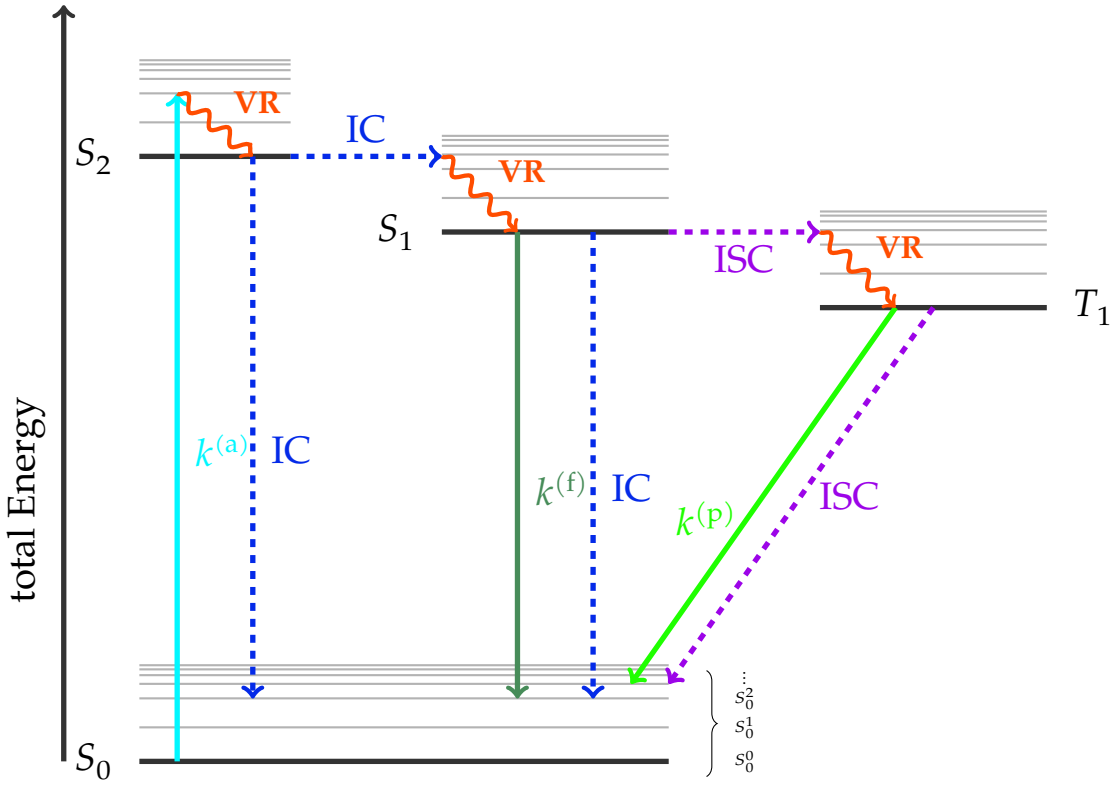


FIGURE 1.2: **Jablonski Diagram** for an exemplary set of transitions for 4 states of an excitable fluorophore. S_0, S_1, S_2 are the electronic singlet and T_1 triplett states. Radiationless Vibrational Relaxation (VR) and Internal Conversion (IC) are marked with red and blue, while the quantum-mechanically forbidden Intersystem Crossing (ISC) is displayed in purple. $k^{(a)}$ marks the absorption rate of incoming photons while $k^{(f)}$ and $k^{(p)}$ are the radiative de-excitation (fluoroscence and phosphorescence) processes. The vibronic levels are drawn in slight grey and examplarily named in case of S_0^0 to S_0^2. Inspired by [27].

1.5 FLUORESCENCE

For modelling the imaging process the Born approximation (Section A.8) is used, since its assumptions are approximately fulfilled in case that the emitted light is spatio-temporarily incoherent, meaning that the mutual interference of the sample response reduces to a delta peak (see Equation A.50). Fluorescence fulfills these requirements and thus the derivations made in Section 1.3 for Fourier-Optics can be directly applied to describe the imaging process of a fluorescent sample.

For describing light-matter interaction in case of photon absorption and emission of multi-atomic molecules the Jablonski diagram will be adopted as simplified model [27]. Section A.9 provides a more detailed discussion on the assumption and implications of Jablonski diagrams.

Here, a valence electron can be excited from the ground state S_0^0 to a higher energy state S_j^k by supplying the appropriate discrete energy $\Delta E_{0,j}^{0,k}$. The upper indices here describe the vibrational level - and are often not directly noted - while the lower indices denote the electronic level or general energy state. Rotational states are omitted.

For a 4-state system, the absoprtion (a) of a photon of suitable wavelength excites a valence electron from the ground state S_0^0 to an excited electronic state, S_2^4, with the absorption rate, $k^{(\mathrm{a})}$:

$$k^{(\mathrm{a})}(\vec{x}) = \eta^{(\mathrm{a})}(\vec{x}) I^{(\mathrm{ex})}(\vec{x}) \tag{1.28}$$

and $\eta^{(\mathrm{a})}(\vec{x})$ the fluorophore optical absorption cross-section. The spatial dependence is typically omitted, i. e. $k^{(\mathrm{a})}(\vec{x}) \rightarrow k^{(\mathrm{a})}$. The electron does not fall directly back to the ground level S_0^0, but the excited system lowers its energy through vibrational relaxation as well as interactions with the environment (often referred to as Stokes shift). The excited system will generally relax to the ground vibrational level in the lowest excited electronic state, from where other electronic transitions, e.g. fluorescence, can take place (Kasha's Rule) [27]. In this example, the electron initially falls back to the vibronic ground level of S_1 (i.e., S_1^0) without radiation within about $t = 10^{-11}$ s due to Vibrational Relaxation (VR) and Internal Conversion (IC), i. e. $S_2^4 \longrightarrow S_1^3 \longrightarrow S_1^0$. This state is stable for a short time $O(ns)$. The electron can relax to the ground state non-radiatively via fluorescence-quenching IC or via emission of a photon with $\Delta E_{1,0}^{0,l}$ as fluorescence (f) with the emission rate $\langle S_1 \rangle \cdot k^{(\mathrm{f})}$. Here, $k^{(\mathrm{f})} = 1/\tau^{(\mathrm{f})}$ is the emission rate constant and $\tau^{(\mathrm{f})}$ the fluorescent lifetime. Note, that the total radiative lifetime is anti-proportional to the sum of radiative, non-radiative and Intersystem Crossing (ISC), but will not be discussed further here. A forbidden transition via spin-flip into the energetically more favorable and slightly more stable triplet state T_1^m is also possible, although this process occurs much less frequent than the previous two processes and on longer time scales (ms). The transition from the triplet state T_1^0 to S_0^3 can be radiationless and is called ISC. ISC takes a comparably long time as the transition probability is low due to spin restrictions. In the case of a radiative transition, one speaks of phosphorescence (p). All these and further non-accounted-for

processes (e. g. bleaching or Förster-Resonance Energy Transfer (FRET)) finally result in a fluorescence flux $\Phi^{(\mathrm{F})}\,[W]$ which can be described by:

$$\Phi^{(\mathrm{F})} = Q^{(\mathrm{e})}\eta^{(\mathrm{F})}I^{(\mathrm{ex})} \tag{1.29}$$

with $Q^{(\mathrm{e})}$ the fluorophore quantum efficiency, $\eta^{(\mathrm{F})}/\left[cm^2\right]$ the absorption cross-section of the fluorophore and $I^{(\mathrm{ex})}/\left[W/cm^2\right]$ the local illumination intensity. Here, $I^{(\mathrm{ex})}$ was assumed to be small enough such that the fluorescent flux $\Phi^{(\mathrm{F})}$ can be modelled as a linear function in $I^{(\mathrm{ex})}$. $\Phi^{(\mathrm{F})}$ can be converted to the fluorescence emittance $I^{(\mathrm{F})}$ per unit-depth $dz/\,[m]$ by multiplication with the local concentration per Z-slice $C^{(\mathrm{F})}\,(\vec{\chi}, dz)\,/\left[cm^{-2}\right]$ to:

$$\begin{aligned} I^{(\mathrm{F})}\,(\vec{\chi}, dz) &= \Phi^{(\mathrm{F})}C^{(\mathrm{F})}\,(\vec{\chi}, dz) \\ &= I^{(\mathrm{ex})}\,(\vec{\chi}, dz)\,S^{(\mathrm{F})}\,(\vec{\chi}, dz) \end{aligned} \tag{1.30}$$

by using the voluminetric (thin-slice) local fluorescence response $S^{(\mathrm{F})} = Q^{(\mathrm{e})}\eta^{(\mathrm{F})}C^{(\mathrm{F})}$ with Equation 1.29.

According to the previous descriptions, an excitation spectrum consisting of many narrow peaks and an emission spectrum shifted towards longer wavelengths due to the Stokes shift can be expected. In most practical measurement cases, however, in the Region of Interest (ROI) not one molecule in isolation, but many molecules of the same kind are surrounded by solvents and other molecules, which determine the local and overall environment in near as well as in far distance. At rest and under constant external conditions, an equilibrium state is formed for this mixture of molecules to be analyzed, in which most of the similar fluorophores are in one of the possible vibronic levels of the ground state S_0, typically S_0^0. Local excitation of some molecules changes many environmental parameters and the dynamical system tries to return to equilibrium. In this process, different eigenstates (vibronic levels) are successively adopted until the system transitions locally to a more stable ground state, before the local reactivity is neutralized again and the system falls back to the ground state S^0. Time-scales and set of transition states chosen depend additionally on the solvent relaxation.

Quantum mechanically, the transition from singlet to triplet states is forbidden for an isolated molecule consisting of light atoms, but the pertubation of its own electron distribution by the interactions with the surrounding molecules allows for such a transition. If the molecule contains a heavy atom the spin-orbit coupling could enhance the probability of a transition. The spin flip back to the singlet state is again forbidden, which is why this

state is relatively long-lived [28]. Finally, this superposition of possible pathways, different local environments and especially the influence of the solvent / different neighboring particles results in the smearing and shifting of the single peaks to a continuous spectrum. If the fluorophore would have been one isolated molecule in vacuum and would have created an entangled state with the exciting field, then the phase and orientation-relation to the excitation-field would be continuous and a coherent description of the process would be possible. Due to the random interaction with and pertubations by the environment this phase-relation is lost and the fluorescent response to the incoming field can be approximated as incoherent or phase-unrelated. The average coherence time is in the range of $\tau_{\text{coh}}^{(\text{f})} \propto fs$ [14] whereas typical detector speeds are orders of magnitude slower. This allows for modelling the measurement process as a pure incoherent summation of non-temporally coherent fluorescent events. The absorption efficiency of an exciting electron by a fluorophore depends further on orientation of the fluorophore with respect to the incoming polarization of the exciting field [29]. Fluorescence (as well as phosphorescence) typically can be assumed to be spatially incoherent as well, as the polarized light sent out by each individual fluorophore (with its respective orientation) is not measured individually but as an ensemble of wave-fronts emerging from many different fluorophores at different spatial positions.

In this work, unless otherwise mentioned, the representative dye for simulations and further analyses will be Alexa-488 [30]. This dye is pH stable and exhibits low photobleaching. The dye has several relevant absorption bands, but in this work mainly 1-photon processes are analyzed whereby excitation wavelengths between $400\,nm - 800\,\text{nm}$ are relevant. Accordingly, excitation and emission around the maxima $\lambda^{(\text{ex})} = 490\,\text{nm}$ and $\lambda^{(\text{em})} = 525\,\text{nm}$ will be used.

1.6 SIGNAL, NOISE AND SIMILARITY

The quantum nature of light leads to fluctuations in the discrete conversion of a photon to electron during a detection process which is called *shot noise*. The measurement process can also be formulated as asking the question: „What is the probability that $K = k$ photons

were measured in the time interval $\Delta t = T$?". The Poisson statistic $\mathcal{P}$ of a random variable K answers this question:

$$\mathcal{P}\{\mu_K\} \equiv P_K(K = k|\mu_K) = \frac{(\mu_K)^k}{k!} e^{-\mu_K} \tag{1.31}$$

with $K = k$ realizations, i. e. measured photons, of the Poisson distributed random variable K with expectancy μ_K. The random variable K is discrete (i. e. realizations $k \in \mathbb{N}^+$), but the mean $\mu_K \in \mathbb{R}_0^+$ is continuous. The formulation $P_K\left(K = k \,|\, \mu_K\right)$ means: the probability of measuring $K = k > 0$ photons given the mean $\mu > 0$ of the random process takes the value μ_K.

While shot-noise can be viewed as multiplicative noise in the conversation from a continuous field distribution into a discrete charge distribution, construction-related additional multiplicative and additive noise sources are added in the processing of this charge. Provided that these additional noise sources or random variables $X^{(\mathrm{i})}$ are *mutually independent*, they can be holistically and suitably normalized as a random variable $X = \sum_i X^{(\mathrm{i})}$ whose probability distribution can be described by Gaussian probability density by means of the central limit theorem [17]:

$$\mathcal{G}\left\{\mu_X, \sigma_X^2\right\} \equiv p_X\left(X = x \,|\, \mu_X, \sigma_X^2\right) = \frac{1}{\sqrt{2\pi\sigma_X^2}} \exp\left\{-\frac{(x - \mu_X)^2}{2\sigma_X^2}\right\} \tag{1.32}$$

and hence $\mathbb{E}\left\{X\right\} = \mu_X$ and $\mathbb{V}\left\{X\right\} = \sigma_X^2$ are the mean and variance of this distribution.

For this thesis shot noise is assumed to be the significant noise source. For a deeper introduction see Section A.11 and for a comparison between Gaussian and shot noise see Section A.12.

The Signal-to-Noise Ratio (SNR) generally describes the relationship between the power of signal $L^{(\mathrm{signal})}$ and noise $L^{(\mathrm{noise})}$. The definition of *signal* and *noise* depends on the objective of the analysis process. In the context of this work, *signal* $L^{(\mathrm{signal})}$ is defined as the measured mean intensity $\mathbb{E}$, discretized and amplified as the count K (per pixel or scanning position) of the photon flux at the detector $\Phi^{(\mathrm{T})}(t)$ (see Equation A.67 for further details), but note that intensity and definition can change if ooF is taken into account. The *noise* $L^{(\mathrm{noise})}$ on the other hand describes the average fluctuation around the mean intensity (i. e. signal) and is measured by the standard deviation $\sqrt{\mathbb{V}}$. While in real space the mean also gives a real number, i. e. $\mathbb{E} \in \mathbb{R}$, in Fourier space it could be $\mathbb{E} \in \mathbb{C}$.

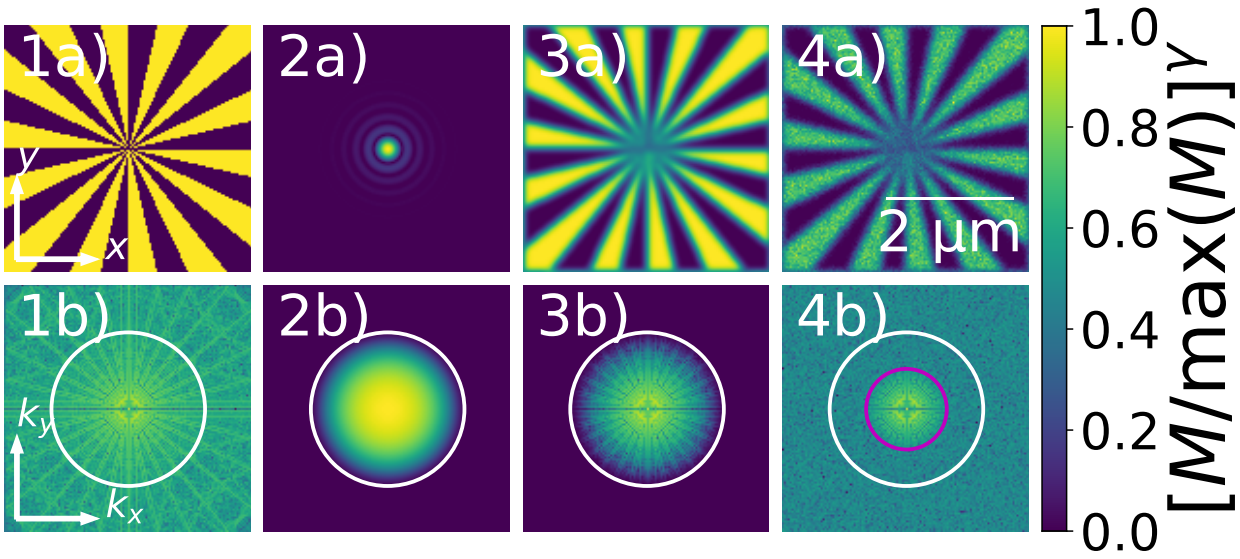


FIGURE 1.3: **Forward-Simulation Example** Upper row shows maximum normalized real space images with $\gamma = 0.2$ for 2a), else $\gamma = 1$. Lower row shows Fourier space magnitude of images with $\log_{10}(1 + \left|\widetilde{M}\right| / \left|\max \widetilde{M}\right|)^\gamma$ normalization and $\gamma = 0.5$. 1) *spokes2d* target. 2) WF PSF at $d^{(\mathrm{SPL,lat})} = 50\,\mathrm{nm}$, $NA = 1.4$, $n^{(\mathrm{im})} = 1.518$, $\lambda^{(\mathrm{em})} = 488\,\mathrm{nm}$, $N^{(\mathrm{pix})} = [181, 181]$. 3) WF image of 1a) (convolved with PSF 2a)). 4) Poisson noise added with $N^{(\mathrm{Phot})} = 100$. Theoretical OTF-support limit $k^{(\mathrm{Abbe,WF})}$ is shown as a white, the noise induced cut-off frequency $k^{(\mathrm{C})}$ as an magenta circle.

However, since the SNR is to be used as a real measure and for the variance $\mathbb{V} \in \mathbb{R}_+$ holds, a general real SNR of a random variable K can be defined as:

$$\mathrm{SNR}\,(K) = \frac{L^{(\mathrm{signal})}\,(K)}{L^{(\mathrm{noise})}\,(K)} = \frac{|\mathbb{E}\,\{K\}|}{\sqrt{\mathbb{V}\,\{K\}}} \in \mathbb{R}_+^0 \tag{1.33}$$

In real space, for a Poisson-distributed random variable K such an SNR can be increased by measuring more photons (Equation 1.34), i. e. increasing the mean:

$$\mathrm{SNR}^{(\mathcal{P})} = \frac{\mu_K}{\sqrt{\mu_K}} = \sqrt{\mu_K} \tag{1.34}$$

A brief 2D fluorescence WF simulation to display the difference between the theoretical maximum size of the OTF support $\mathrm{supp}\left\{\tilde{h}\right\}$ limited by $k^{(\mathrm{Abbe})}$ and the actual achieved cut-off frequency $k^{(\mathrm{C})}$, where for $k \geq k^{(\mathrm{C})}$ the signal is indistinguishable from the noise floor, follows. As simulation parameters a lateral pixel sampling of $d^{(\mathrm{SPL,lat})} = 50\,\mathrm{nm}$, objective lens $NA = 1.4$, oil immersion with refractive index $n^{(\mathrm{im})} = 1.518$, emission wavelength $\lambda^{(\mathrm{em})} = 515\,\mathrm{nm}$, measured pixel area $N^{(\mathrm{meas})} = [N^{(\mathrm{x})}, N^{(\mathrm{y})}] = [181, 181]$ and $N^{(\mathrm{Phot})} = 100$ as expected photon-count were used to simulate a noisy image of the

spokes2d[2] target. Figure 1.3 displays the simulation results. The *spokes2d* (Figure 1.31a) target is convolved with the simulated PSF (2a) to yield the ideal image (3a). Finally, Poisson noise is applied to the simulation (4a). The lower row of Figure 1.3 shows the respective Fourier space representation of the upper row image in the same column.

The theoretical limit of the OTF-support supp $\left\{\tilde{h}\right\}$, given by $k^{(\text{Abbe,WF})} = 1/d^{(\text{Abbe})} \approx 22/\,(\mu m)$, is highlighted with a white circle in the Fourier-images b) of Figure 1.3. The Fourier noise-floor $\sigma^{(\text{WF})}$ in case of the given WF imaging is calculated as an average of pixels where $k \geq k^{(\text{Abbe,WF})}$ holds. The effective noise-cutoff frequency $k^{(\text{C})}$ is calculated as the frequency for which on average $\left|\tilde{h}(k)\right| \approx \sigma^{(\text{WF})}$ for $k \geq k^{(\text{C})}$ holds and is marked with a magenta circle. In this simulation $k^{(\text{C})} = 11/\,(\mu m)$.

Thus noise leads to the fact that high frequencies of the image spectrum cannot be resolved, although they are measured with ideal imaging. However, reconstruction methods such as deconvolution can limit or even eliminate this degradation, see Section 1.7. The information collected from the sample is low-pass filtered by the transfer function of the system before it can be detected. During measurement the spatio-temporal continuous intensity distribution at the detector is time-averaged into temporal bins of length T and measured according Equation A.52 as $I^{(1)}\,(\vec{\chi})$, which is subject to further manipulations by the analog-digital-processing units until the final image M is stored. In case of a multi-pixel detector, each pixel has its own shape and detection efficiency and hence the discretized photon-count at the k-th pixel can be described by [31]:

$$M_k = \int_{\mathbb{R}^2} d\vec{\chi}\; I^{(1)}\,(\vec{\chi}) \cdot D_k\,(\vec{\chi}) \tag{1.35}$$

$$D_k\,(\vec{\chi}) = \int d\vec{\chi}_k \eta_k\, A_k\,(\vec{\chi}_k)\;\delta\,(\vec{\chi}_k - \vec{\chi}) \tag{1.36}$$

where η_k is the photon-conversion factor, $A_k\,(\vec{\chi})$ is the pixel-shape and $\delta\,(\vec{\chi} - \vec{\chi}_k)$ the position of the k's pixel within the detector coordinates. Analog to Equation 1.25 in this thesis all pixels are assumed to have the same form- and photon-conversion factor, hence $\eta_k \equiv \eta$ and $A_k \equiv A$.

The precision of the discrete representation of $I(\vec{x})$ increases with the amount of sampling pixels used making the use of higher and higher polynomials for data-fitting possible. It is even more remarkable that a bandlimited signal, as is the typical case in microscopy,

2 The *spokes2d* target is a binarized version of a radial-sinusoidal in 2D.

can be reconstructed perfectly if sampled at twice the frequency of its band-limit. This is called Shannon-Nyquist limit and can be written as:

$$k^{(\mathrm{Ny})} = \frac{k^{(\mathrm{SPL,lat})}}{2} \geq k^{(\mathrm{Abbe})} \tag{1.37}$$

where $k^{(\mathrm{Abbe})}$ is the Abbe or maximum possible frequency of the linear bandlimited signal, $k^{(\mathrm{SPL,lat})}$ is the sampling frequency and $k^{(\mathrm{Ny})}$ is the Nyquist limit. Especially it predicts that under the assumption of a bandlimited signal a too low sampling leads to aliasing, meaning that high-frequencies are mirrored back into lower frequencies. See Section A.13 for a one and a two dimensional example.

In order to find an objective measure of measurement data comparability, the first question to ask is what kind of transformation is preferred. That is, should dimensionality of the (potentially n dimensional) measurement data be reduced (e. g. 1D or 2D), keep its dimensionality or even increase to n+m dimensions? In the case of this work, for convergence testing of deconvolution algorithms or reconstruction quality comparisons, a one-dimensional image representation was chosen, i. e. a 2D/nD image is thus mapped to a real number.

For the comparison of the similarity of two images $M^{(1)}$ and $M^{(2)}$, the *Pearson correlation coefficient*, in this work denoted as Normalized Cross Correlation (NCC) C, was chosen [32]:

$$C\left(M^{(1)}, M^{(2)}\right) = \sum_l M_l^{(1,\mathrm{c})} \cdot M_l^{(2,\mathrm{c})} \qquad \in \mathbb{R}_{[-1,1]} \tag{1.38}$$

$$M_l^{(\mathrm{i,c})} = \frac{1}{\sqrt{N^{(\mathrm{l})} - 1}} \frac{M_l^{(\mathrm{i})} - \mathbb{E}\left\{M^{(\mathrm{i})}\right\}}{\sqrt{\mathbb{V}\left\{M^{(\mathrm{i})}\right\}}} \tag{1.39}$$

and $N^{(\mathrm{l})}$ the numbers of pixels of the image. Here, two images/stacks are reduced by their mean value, normalized by their repective standard deviation before the (pixel-wise) correlation is calculated. Two images are said to be NCC close if $C \to 1$, because then not only high-similarity but positive correlation is given. $C \to -1$ stands for high similarity but anti-correlation, thus the two images only differ by sign. Images are called *NCC* independend if $C \to 0$. The sensitivity to (sub-) pixel shifts can be both good and insufficient. Good, if the images should overlap pixel-wise exactly and relative intensity changes should be compared. Insufficient, if the two very identical images are shifted by n (sub-) pixels w. r. t. to each other. If the images have no inherent symmetry w. r. t. to

n (sub-) pixel shift-operations, then the NCC measure might tend towards zero $C \to 0$ even though they are similar. Within this thesis, a reconstruction that achieves a NCC of $C \geq 0.9$ will be accepted as a reconstruction of high quality.

1.7 INVERSE MODELLING

The last chapters have laid the foundation for the description of the imaging process, conditions on detectors as well as the generation of noise. In this chapter, an abstracted forward model is established and basic approaches and limits of invertibility are presented. With Weighted Averaging in Fourier Space (WAVG) an approximately SNR-optimal recombination of multiple images and with $1/OTF$, Wiener Deconvolution (WD) and Richardson-Lucy Deconvolution (RL) different deconvolution schemes are outlined.
The imaging equation found for a fluorescent ISM (Equation 1.23) can be reformulated as:

$$M_a = \mathcal{P}\{h_a \otimes S\} = h_a \otimes S + \mathcal{N}_a \tag{1.40}$$

$$\widetilde{M}_a = \tilde{h}_a \, \tilde{S}_{\tilde{l}} + \widetilde{\mathcal{N}}_a \tag{1.41}$$

$$M_{al} = \sum_m H_{alm} \, S_m + \mathcal{N}_{al} \tag{1.42}$$

where $\mathcal{N}_a$ is a specific value drawn from the specific noise distribution given a sample S, $\widetilde{\mathcal{N}}_a$ is the respective value drawn from the according noise distribution in Fourier space and $\mathcal{N}_{al}$ is the spatially-vectorized 1D-representation of $\mathcal{N}_a$. All functional dependencies have been omitted. While $h \otimes S$ is an ideal imaging process without any variance, the real measured values are resembled by a draw from a Poisson-distributed ($\mathcal{P}$) random variable with mean $\mu_{al} = \sum_m H_{alm} \, S_m$ per pixel and detector (Equation 1.42). Hence, every image yields only limited and subjective information about a sample. The question is: „How much information about S can be recovered from the measured image M given a set of assumptions?“ In case only Poisson-noise is present in the image M the $\mathcal{N}_{a\tilde{l}}$ holds the same constant for all frequencies.

Note, that the partial derivative of Equation 1.42 with respect to the parameters S_n is:

$$\partial_{S_n} \mu_l = \partial_{S_n} \sum_m H_{lm} S_m = \sum_m H_{lm} \left(\partial_{S_n} S_m \right) = \sum_m H_{lm} \, \delta_{nm} = H_{ln} \tag{1.43}$$

Every imaging process can be viewed as taking a sample from the ground-truth hence representing a subjective view on reality. While $P_{K_l}(K_l = k_l \,|\, \mu_l)$ asks the question: „How likely is it to measure the realization k_l of a random variable K_l given an unknown parameter μ_l?“, an atomic likelihood function

$\mathcal{L}_l(\mu_l \mid k_l) = P_{K_l}(K_l = k_l \,|\, \mu_l)$ can be introduced that asks: „How likely was a parameter-realization μ_l under a given measurement k_l?“. Note the subtle difference that P_{K_l} is a probability distribution of the random variable K_l and thus describes the probability of a measurement while the likelihood function describes the (basic) parameters μ_l [33]. Hence, $\mathcal{L}_l(\mu_l \mid k_l)$ is not derived from $P_{K_l}(K_l = k_l \,|\, \mu_l)$ via Bayes Theorem (Equation 1.54), but is the same w. r. t. notation while emphasizing the parameters instead of the measurements.

The joint likelihood-function $\mathcal{L}$ for a set of individual measurements can then be described by:

$$\mathcal{L}\left(\vec{\mu} \mid \vec{k}\right) = P_{\vec{k}}\left(\vec{k} \mid \vec{\mu}\right) \tag{1.44}$$

$$= \prod_{l=0}^{N-1} P_{K_l}(K_l = k_l \,|\, \mu_l) \tag{1.45}$$

with $\vec{\mu} = \left(\mu_0, \mu_1, \dots, \mu_{N_\mu}\right)$ the parameter vector and $\vec{k} = (k_0, k_1, \dots, k_{N-1})$ the measurement vector (=pixel-values of measured image). By variation of parameters $\vec{\mu}$ a (local) maximum of Equation 1.45 and thereby a most-likely estimate $\hat{\vec{\mu}}$ that might have given rise to the measured $\vec{k}$ can be found. The found estimate $\hat{\vec{\mu}}$ is not necessarily connected to the real ground truth which gives rise to further support the calculations by using e. g. regularizers (Equation 1.55). Note that while Equation 1.44 still includes the case of dependent random variables K_l Equation 1.45 uses independent variables. Given that Equation 1.45 is a positive function, its negative natural logarithm ln can be minimized instead:

$$\mathfrak{L}\left(\vec{\mu} \mid \vec{k}\right) = -\ln \mathcal{L}\left(\vec{\mu} \mid \vec{k}\right) \tag{1.46}$$

From the assumed model Equation 1.40 we know that the actual parameter S_m is hidden in $\mu_l = \sum_m H_{lm} S_m$ (Equation 1.16) and hence we need to find the optimal estimator $\hat{S}_m$ that minimizes $\mathfrak{L}$. In case an analytic relation between the optimization function $\mathfrak{L}$ and the estimator S_m is given, the minimum can be found by analyzing its first and second derivative.

As explained above and in Section A.11 the object information degrades through the imaging process. According to the imaging model given in Equation 1.41, it is reasonable to think of reconstruction by direct inversion of the forward model, but this so-called 1/OTF deconvolution either amplifies high spatial frequencies and particular noise strongly or strongly low-pass filters the image. An in-depth explanation of these issues is provided in Section A.18.

Instead, a better approach is to assume an information-limited measurement process where it can be conjectured that the best possible estimator $\hat{S}_l$ of each individually measured image pixel l is Gaussian distributed around the original object distribution S_l. Solving the log-likelihood problem Equation 1.46 for this conjecture leads to the filter-function

$$\widetilde{W} = \frac{\tilde{h}^*}{\left|\tilde{h}\right|^2 + \Gamma} \qquad , \Gamma = \frac{\mathbb{E}\left\{\left|\widetilde{\mathcal{N}}\right|^2\right\}}{\left|\tilde{S}\right|^2} \tag{1.47}$$

which reconstructs an estimate of $\hat{S}$ of the sample distribution according to $\hat{S} = \mathcal{F}^{-1}\left\{\widetilde{W}\widetilde{M}\right\}$ from the measured image M. The found filter Equation 1.47 is an optimal solution within the assumptions and is called WD [34]. The complex conjugation of the OTF in the numerator inverts any PSF conditional object shifts and reverts a potential asymmetry. Generally problematic is the determination of the regularization parameter Γ for which both original object spectrum $\tilde{S}$ and the noise spectrum $\left|\widetilde{\mathcal{N}}\right|^2$ are necessary. In the case of the so-called generalized Wiener filter $\Gamma = const.$ is set. Setting $\Gamma/\tilde{h}^* \equiv \varepsilon$, the filter found is equal to the regularized $1/OTF$ solution Equation A.83 in the domains $\left|\tilde{h}\right| \neq 0$. For the derivation see Equation A.87.

Imaging methods with parallel acquisition of multiple viewing angles on the sample distribution offer the possibility of individual as well as combined processing. In the context of the processing methods presented so far, the different views of the SNR-optimal weighted OTFs are subsequently simply summed together into one resulting view. A suitable combination of all images with weights $\widetilde{\omega}_a$ for all Fourier pixels $\tilde{l}$ can be found according to Equation 1.41 as:

$$\widetilde{M}_{\tilde{l}}^{(\mathrm{wa})} = \sum_a \widetilde{\omega}_{a\tilde{l}} \widetilde{M}_{a\tilde{l}} \tag{1.48}$$

The processing that yields the image $M_l^{(\mathrm{wa})} = \mathcal{F}^{-1}\left(\widetilde{M}_{\tilde{l}}^{(\mathrm{wa})}\right)$ is called Weighted Averaging in Fourier Space (WAVG) [35]. The weights $\widetilde{\omega}_a$ are found to be:

$$\widetilde{\omega}_c = \tilde{h}_c \tilde{\sigma}_c^{-2} \tag{1.49}$$

with the variance $\mathbb{V}\left\{\tilde{h}_c\right\} = \tilde{\sigma}_c^{-2}$. Note the difference of Equation 1.49 to Equation 1.47.

While the weights are mainly influenced by the PSF as before in case of Wiener-Filtering the normalization is with respect to their individual variance. The SNR-optimal recombination $\widetilde{M}^{(\mathrm{wa})}$ has a non-uniform noise-spectrum as opposed to the assumed noise-spectrum according to the forward model Equation 1.40. By utilization of Equation A.92 and Equation 1.49 noise normalization can be achieved via:

$$\widetilde{M}^{(\mathrm{nwa})} = \frac{\widetilde{M}^{(\mathrm{wa})}}{\sqrt{\mathbb{V}\left\{\widetilde{M}^{(\mathrm{wa})}\right\}}} \tag{1.50}$$

$$\mathbb{V}\left\{\widetilde{M}^{(\mathrm{wa})}\right\} = \sum_a \tilde{h}_a^2 \tilde{\sigma}_a^{-2} \tag{1.51}$$

This processing is called noise normalized Weighted Averaging in Fourier Space (nWAVG). With Equation 1.48 a noise-normalized OTF $\tilde{h}^{(\mathrm{nwa})}$ can now be introduced (for more details on the derivation see Section A.20):

$$\tilde{h}^{(\mathrm{nwa})} = \sum_a \widetilde{\omega}_a \tilde{h}_a = \sum_a \left[\tilde{h}_a \tilde{\sigma}_a^{-1}\right]^2 \tag{1.52}$$

Assuming a set of independent-Poisson distributed random variables, i. e. each measured pixel, with distribution $P_{K_l}\left(K_l = k_l \,|\, \mu_l\right)$ (Equation A.65) an extremum of the log-likelihood $\mathfrak{L}$ of the Maximum Anterior Likelihood (MAL) problem w. r. t. the parameters S_n can be found according to:

$$\partial_{S_n} \mathfrak{L} = -\partial_{S_n} \ln\left[\prod_{l=0}^{N-1} P_{K_l}\left(K_l = k_l \,|\, \mu_l\right)\right] \stackrel{!}{=} \mathbb{O}_n \tag{1.53}$$

with the partial derivative according to Equation 1.43 and the zero-matrix $\mathbb{O} \in \mathbb{N}^{N^{(\mathrm{l})}/2 \times N^{(\mathrm{l})}/2}$. Using the solution of Equation 1.53 iterative schemes can be used to reconstruct an estimate $\hat{S}$ of the original object distribution. The derivation of the RL update scheme is given in Section A.21 as an example.

In the Bayesian world-view, all events are always assumed possible even though their probability is very close to zero. The belief about the stochastic system can be updated by using Bayesian inference:

$$P\left(\vec{\mu} \mid \vec{k}\right) = \frac{P\left(\vec{k} \mid \vec{\mu}\right) P\left(\vec{\mu}\right)}{P\left(\vec{k}\right)} \tag{1.54}$$

where $P\left(\vec{\mu} \mid \vec{k}\right)$ is the posterior (conditional) probability distribution and $P\left(\vec{k} \mid \vec{\mu}\right)$ the anterior (conditional) probability distribution. Note that $P\left(\vec{k}\right) = \int d\vec{\mu} P\left(\vec{k} \mid \vec{\mu}\right) P(\vec{\mu})$ in case of continuous random variables. The principle can be transferred to likelihood functions and thus a posterior likelihood $\mathcal{L}^{(\text{post})}$ can be derived from an anterior likelihood $\mathcal{L}^{(\text{ant})}$ by multiplication with a prior $P(\vec{\mu})$ and normalization to the measurement $P(\vec{k})$. Hence, the Maximum A Posteriori likelihood (MAP) is the same as the MAL (Equation 1.44) for assumption of uniform prior and ignoring the normalization factor (denominator).

Depending on the shape of the anterior likelihood-function the imaging problem might not have a global minimum, might not converge towards the global minimum due to $\mathcal{L}$ having a rough likelihood-landscape or will only find a locally optimal solution that is by far not the best solution available. For an ill-posed problem a uniform prior seems like an unlikely assumption and thus more biased beliefs about the reality of the forward-problem are typically implemented. By this, the ill-posed problem tends to become more convexified and convergence towards a more stable solution seems possible. It follows for log-posterior likelihood $\mathfrak{L}^{(\text{post})}$:

$$\begin{aligned}
\mathfrak{L}^{(\text{post})}\left(\vec{\mu} \mid \vec{k}\right) &= -\ln \mathcal{L}^{(\text{post})} \\
&= -\ln \frac{\mathcal{L}^{(\text{ant})}\left(\vec{k} \mid \vec{\mu}\right) P(\vec{\mu})}{P(\vec{k})} \\
&= \ln P(\vec{k}) - \ln \mathcal{L}^{(\text{ant})}\left(\vec{k} \mid \vec{\mu}\right) - \ln P(\vec{\mu})
\end{aligned} \tag{1.55}$$

While many priors $P'(\vec{\mu}) \equiv \ln P(\vec{\mu})$ exist, Total Variation (TV) [36], Good's Roughness (GR) [37] and Gradient Square (GS) will be mainly used within this thesis. With $\partial_{\vec{e}_d}$ being the derivative along the dth spatial unit-vector[3] $\vec{e}_d$ of the sample estimate $\hat{S}(\vec{\mu})$ the chosen priors are defined according to:

$$\textbf{TV:} \qquad P'^{(\text{TV})}(\vec{\mu}) = \sqrt{\sum_d \left|\partial_{\vec{e}_d} \hat{S}(\vec{\mu})\right|^2} \tag{1.56}$$

$$\textbf{GR:} \qquad P'^{(\text{GR})}(\vec{\mu}) = \sum_d \frac{\left(\partial_{\vec{e}_d} \hat{S}(\vec{\mu})\right)^2}{\hat{S}(\vec{\mu})} \tag{1.57}$$

$$\textbf{GS:} \qquad P'^{(\text{GS})}(\vec{\mu}) = \left[P'^{(\text{TV})}(\vec{\mu})\right]^2 \tag{1.58}$$

Note, that $\vec{\mu}$ is space depended and thus an implicit function of $\vec{e}_d$. While TV uses the square-root of the l^2-norm and enforces sparsity of the gradient GR favors a smooth / curved, hence rather non-sparse, result. GS favors smoother changes than TV, because

3 For a 2D image $d \in \{0, 1\}$ and e. g. $\partial_{\vec{e}_{d=0}}$ is typically denoted as ∂_x.

the squaring leads to a stronger amplification of bigger derivatives (i. e. bigger local changes) and hence a bigger loss. During the imaging process the sample-information is low-pass filtered by the PSF and thus information about higher frequencies $k \geq k^{(\mathrm{C})}$ is lost. The imaging process further degrades due to the influence of noise. While the sample distribution has a (potentially infinite) frequency support, the imaging process (in the best case) is diffraction limited and thus has a limit support. Hence, the size of the domain of the inverse problem, i. e. non-zero frequencies of $\hat{S}$, might be fundamentally bigger than the support of the calculated solution, i. e. M' after application of the forward model on $\hat{S}$, leading to a potentially vast set of frequencies (i. e. $k \geq k^{(\mathrm{C})}$) that do not influence the calculated solution at all. Any of the three presented regularizers deforms the non-influencing high-dimensional sub-space of the domain such that it forms a rather convex shape of which a (potentially global and stable) minimum could be found.

In this thesis, only one of the 3 presented regularizers is used at a time, and the choice is made based on the analyzed sample.

For all deconvolution operations in this work, the multi-view 3D forward model Equation 1.40 was used, including the cases 1) single-view as limiting case on 1 angle (i. e. $N^{(\mathrm{a})} = 1$) and 2) single Z-slice (i. e. $N^{(\mathrm{z})} = 1$). The Poisson distribution (Equation A.65) was applied as noise model. The Poisson distribution is defined only for positive real expectancy values $\mu_K \in \mathbb{R}^+$, which is why a positivity constraint was used. To ensure positivity, the forward model is extended by a preobject $\hat{S}^{(\mathrm{pre})}$ which is updated in each iteration step of the deconvolution and from which the reconstructed object $\hat{S}$ is obtained via squaring $\hat{S} = \left[\hat{S}^{(\mathrm{pre})}\right]^2$. Thus, the domain of the object $\hat{S}$ is restricted to the non-negative half-space and can even be restricted to the positive half-space by adding an $\epsilon \in \mathbb{R}$, preserving all conditions for the Poisson noise model over the iterative update process. The gradients of constraints and priors required for this model were taken directly from the *InverseModeling* [38] toolbox. The Limited-memory Broyden-Fletcher-Goldfarb-Shanno algorithm (L-BFGS-B) [39] from the *Scipy* [40] package was used for minimization since it can be used to perform parameter determination efficiently even on limited hardware (see Section 1.8).

The deconvolutions used in this work is assigned the naming shown in Table 1.2. In the last column of Table 1.2 the used data axes are represented with symbols for a simple visualization of the dimensionality of the reconstruction problem. The 3D spatial axes are named X, Y, Z, the one dimensional flattened representation of the detection pixel axis

Abbr.	Result name	Deconvolution dimensionality
DEC	$M^{(\mathrm{DEC})}$	single-view $XYZ \longrightarrow XYZ$
cDEC	$M^{(\mathrm{cDEC})}$	single-view with CLSM preprocessing $AXYZ \longrightarrow XYZ$
prDEC	$M^{(\mathrm{prDEC})}$	single-view with PiRe preprocessing $AXYZ \longrightarrow XYZ$
2d3dDEC	$M^{(\mathrm{2d3dDEC})}$	multi-view $AXY \longrightarrow XYZ$
mDEC	$M^{(\mathrm{mDEC})}$	multi-view $AXYZ \longrightarrow XYZ$
amDEC	$M^{(\mathrm{amDEC})}$	multi-view $NAXYZ \longrightarrow XYZ$

TABLE 1.2: **Deconvolution Abbreviations.** Overview of naming given to the used deconvolution approaches throughout this thesis.

is named A and the axis of non-linear fluorescence is named N to account for a different saturation levels. For example, $XYZA \longrightarrow XYZ$ means that a 4D multi-view dataset of three spatial dimensions XYZ and multi-view axis A is combined to reconstruct a 3D sample distribution with axes XYZ. The 2D case $XYA \longrightarrow XY$ is always included as limiting case and is therefore not mentioned further. In this thesis typically the following parameters will be adjusted to tune reconstruction quality of the single-view DEConvolution (DEC): iteration number $N^{(\mathrm{iter})}$ and regularization weight γ.

For data sets that cannot be processed at once on the Graphics Processing Unit (GPU), a tiling algorithm has been included. The stack is first divided into matching tiles of size $N^{(\mathrm{tile})}$ with a mutual overlap of $\Delta\vec{x}^{(\mathrm{olap})}$. After sequential processing of the tiles, they are reassembled in a weighted manner using a Hann windowing function [41]. For use in this thesis, the existing online package *tiler* [42] was jointly developed.

1.8 HARDWARE AND TOOLBOXES

A laptop (T470p, Lenovo) serves for all data processing of the presented thesis. Relevant specifications are given in Table 1.3. The used abbreviations are: Central Processing Unit (CPU), Random-Access Memory (RAM), Storage (STO), Graphics Processing Unit (GPU) and Operating System (OS).

Component	Specification	Product Name, Vendor
CPU	4x 2.90 GHz	Core i7-7820HQ, Intel
RAM	32 GB DDR4 at 2400MHz	M471A2K43CB1-CRC, Samsung
STO	512GB SSD M.2 NVMe	MZVLW512HMJP-000L7, Samsung
GPU	2GB GDDR3 64bit	GeForce 940MX, NVIDIA
OS	Manjaro Linux	Manjaro 21.2.5, Manjaro Community

TABLE 1.3: **Processing System Specifications.** Overview of hardware specifications of the used laptop (T470p, Lenovo).

All code relevant to this thesis was written in the higher-level programming language *Python* [43] and is largely contained in the toolbox *MicroPy* [44], which was created as part of my PhD. Simulations were realized with *Numpy* [45] as well as the toolbox *NanoImagingPack* [38] based on it, which was jointly developed within our working group. For deconvolutions, the toolbox *InverseModelling* [38] is used, which uses *tensorflow* [46] as interface for data processing on the GPU. As part of my PhD, *InverseModelling* was also advanced. Since the hardware resources of the laptop used are very limited, some algorithms were equipped with tiled processing, for which the Python package *tiler* [42] was collaboratively advanced.

2

METRICS FOR SOFTWARE AUTOFOCUS

The first principle is that you must not fool yourself and you are the easiest person to fool.

— Richard P. Feynman

2.1 MOTIVATION

Since the achievable resolution of a microscope, in the case of diffraction limited imaging, depends on the *NA* of the objective lens (Equation 1.12), it is important to use to the full extent. If the plane of fluorescent emission is outside the first focal length of the lens, it will not be the optimal *NA*. Even though by definition the *NA* should increase for emission in the region between the lens and the first focal plane, propagation of the focal planes within and the influence of the effective appertures of the optical system are difficult to predict across the complete setup and on average the effective *NA* of the lens is reduced. Due to the inherent Out of Focus (ooF) rejection of Confocal Laser Scanning Microscopy (CLSM) setups, the signal obtained here is more sensitive to an accurate focus position than e. g. in the case of Widefield Microscopy (WF) systems. Although sample signals can still be measured using WF in the presence of focus drift, the Point Spread Function (PSF) degrades quadratically[1] with *NA* with increasing axial distance from the focal position, making suitable processing increasingly difficult for highly noisy images. Thus, the position of the focus should be appropriately chosen and held stable over the length of the measurement interval or, in the case of a longitudinal study, over multiple measurement intervals.

As part of the *You.See.Too.*(UC2, [47]) project, a WF incubator microscope was built (among others) using a plug-in, 3D printed cube system. The built microscope is inexpensive and can be operated in (LED white light) transmission and (fluorescence) reflection mode. For a longitudinal long term study cell growth should be analyzed for which a stable field of view (FoV) and focus position is required. In order to be able to implement

1 Use trigonometric identities on Equation 1.22 for verification.

these requirements as easily as possible without changing the existing setup, the implementation by means of a software solution was chosen. In preparation for potential use in the printed microscope [47], various metrics were analyzed in Section 2.2, a scoring system was established, and verified using an inorganic as well as organic sample.

2.2 DATASET, METRICS AND SCORING MODEL

In this section, for a 3D-printed fluorescence WF microscope a manually chosen Z-slice $z = z^{(0)}$ of an axially thin 3D sample shall be kept in focus over a long time-period. Therefore an algorithm to evaluate the focus quality will be presented and useful metrics will be compared for their applicability under different circumstances.

While the derivation of Equation 1.9 as a model for the 3D-imaging process is based around multiple assumptions the translation invariance of the system PSF h is particularly crucial. Even though the condition might be fulfilled for the analog imaging system, it can be violated again during the acquisition by the electro-mechanical apparatus. Assuming that the signal-conversion behaviour is sufficiently described by Section A.11 the finally measured PSF depends on the apparatus inherent distortions.

The precision requirements of the optical components always depend on the aimed resolution for a given wavelength and imaging mode. Diffraction-limited imaging means the effective Optical Transfer Function (OTF) cutoff-frequency is governed by the limits due to the wave nature of light and not by the noise-level or system-aberrations. If diffraction-limited imaging is secured, the best imaging quality of the given optical system can be achieved.

Analogous to the requirement for the optical elements, requirements can thus also be placed on the electromechanical components such that the diffraction-limited imaging can be recorded reproducibly. Especially in the case of a 3D scanning geometry, reproducible voxel distances that fulfill the Nyquist-Shannon limit (Equation 1.37) are mandatory. Both the lateral/axial step size and the temporal position stability of the X-Y(-Z) stage represent the essential basic requirement for a reliable record of the in-focus field distribution [48].

Both Hardware Autofocus (HAF) and Software Autofocus (SAF) modules can be used to implement an autofocus mechanism. The review by Bian *et al.* [49] summarizes various approaches and a brief summary of their findings is given in Section C.2. In this thesis a set of metrics that could be used with SAF approaches will be analyzed, since these can be implemented into any existing setup without additional costs.

Of all analysed metrics Tenenbaum Gradient (TEN) and Variance (VAR) are two easy to implement filters that provide stable focus measures for a broad range of high- and low Poisson noise as well as amount of aberration. The following steps will be used to proof this conjecture. 1) a set of useful *in-silico* datasets is introduced. 2) The metrics used for evaluation are introduced and the general scoring model is demonstrated. 3) The metrics are compared with respect to three different parameters - optimization independent speed, noise-level, level of spherical aberration and astigmatism - using the introduced scoring model in Section 2.3. 4) The results are applied to experimental data on example of a test target and fluorescent Henrietta Lacks Cervical Cancer (HeLa) cells in a 3D printed microscope in Section 2.4.

The simulated optical system follows a epi-fluorescence WF microscope setup. Basic parameters are lateral $d^{(\mathrm{SPL,lat})} = 80nm$ and axial $d^{(\mathrm{SPL,ax})} = 80nm$ sampling, $NA = 1.4$, $\lambda^{(\mathrm{em})} = 520nm$ and stack size $N = [N^{(\mathrm{x})}, N^{(\mathrm{y})}, N^{(\mathrm{z})}] = [64, 64, 64]$. The maximum expected photon count of each stack was normalized to $\max\{M\} = 1000$ and then used as mean μ_K for application of the Poisson distribution. The entire simulation is based on scalar considerations, which means that e. g. further influences on the focal position due to polarization changes were not considered.

For *in-silico* analysis two different targets were generated, see Figure 2.1. The panel-columns are a) a Z-slice at the in-center slice $z = 32$, b) a Y-slice at $y = 32$, c) the noisy fluorescence WF image of a) and d) the noisy fluorescence WF image of b). In the first row, the *spokes2d* target is demonstrated. It is very suitable for resolution and reconstruction testing due to its radial geometry. This 2D target is placed in the center of an empty 3D volume to simulate a very thin slice object. In the second row, a 3D target called *obj3d* is used, which is included in the *NanoImagingPack* [38]. The 3D target *obj3d* can be understood as a spherical (inflated) cell with nucleus intersected by a rod. *obj3d* is sufficiently sparse and thus can be used even to evaluate methodologies without sufficient ooF rejection.

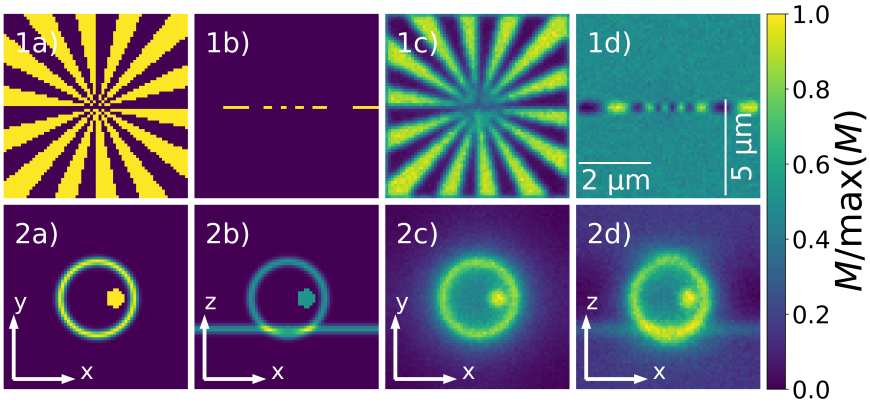


FIGURE 2.1: **Autofocus In-Silico Datasets** Data-sets displayed row-wise. columns from left to right: a) object in-focus cut at $z = 32$ (center), b) object lateral cut at $y = 48$ (lower quarter) for upper and $y = 32$ (center) for lower row, c) noisy-, but aberration-free image in-focus cut of the WF image of a) at $z = 32$ and d) image lateral cut at $y = 32$. Rows: 1) *spokes2d* target, 2) *obj3d*. All images are normalized to their individual maximum.

Two of the most common (Seidel-) aberrations were chosen as representatives for further analysis. Namely, spherical and astigmatic aberrations, represented by Zernicke Polynomials [50] Z_4^0 and Z_2^{-2}, where the upper index marks the degree of the azimuthal components and the lower index the degree of the radial components.The aberrated 3D-PSF $h^{(\text{aber})}(\vec{r})$ is calculated via Equation 2.3 by calculating the 2D Amplitude Transfer Function (ATF) $\tilde{a}$ (Equation 2.1), application of the aberration phase-map W (Equation 2.2), skalar Fourier-Slice-Propagation to calculate ooF slices via application of the Fourier-Shift-Theorem along z, inverse Fourier-transformation and finally taking the absolute square:

$$\tilde{a}^{(\text{aber})}(\vec{\kappa}) = e^{iW(\vec{\kappa})}\tilde{a}(\vec{\kappa}) \tag{2.1}$$

$$W(\vec{\kappa}) = c^{(\text{spher})}Z_4^0(\vec{\kappa}) + c^{(\text{astig})}Z_2^{-2}(\vec{\kappa}) \tag{2.2}$$

$$h^{(\text{aber})}(\vec{r}) = \left|\mathcal{F}\left\{e^{ik_z\Delta z}\tilde{a}^{(\text{aber})}(\vec{\kappa})\right\}\right|^2 \tag{2.3}$$

where $c^{(\text{spher})}, c^{(\text{astig})} \in \mathbb{R}$. It will be demonstrated that $c^{(\text{spher})} = 1$ already marks a rather severe degree of aberration. While spherical aberrations are laterally symmetric but axially asymmetric, astigmatism is rather both laterally and axially asymmetric and hence a comparison is of value.

The final goal of this section (Section 2.2) is obtaining an automated evaluation of image sharpness of a 3D stack M, consisting of $N^{(z)}$ slices around the in-focus Z-slice, of a thin object. Therefore, a set of (2D) sharpness metrics Ψ that reduce a 2D image to a single

number, i. e. the 3D stack is projected onto a 1D Z-dependend curve, and sub-sequent quality measures $q^{(i)}$ that evaluate the results of $\Psi(M)$ further. Finally, the results of the $q^{(i)}$ are weighted and combined to a total score Q.

The sharpness metrics Ψ used for calculation can be divided into:

- differential: Brenner (BRE), Tenenbaum Gradient (TEN) and Total Variation (TOV)
- correlative: Symmetric Vollath F4 (VS_4), Vollath F5 (VO_5)
- statistical: Maximum (MAX) and Variance (VAR)
- spectral: Kristans Entropy (KEN) and Shannon Entropy (SEN)

Function definitions are given in Section B.1. To distinguish the different metrics the naming: $\Psi^{(TEN)}$, on example of TEN, will be adopted.

In best case, for a 3D focus stack M around a thin object the application of the sharpness metrics leads to a high value for the in-focus slice and low value for ooF slices. Hence, the resulting Z-dependend curve $\Psi(M)$ is expected to exhibit 1 maximum (at the in-focus Z-position) and falling edges towards both ends of the Z-scan range. For automatic evaluation of this behavior a set of quality measures $q^{(i)}$ is selected, see [51–53]:

- **accuracy** Distance of calculated to actual in-focus plane; $q^{(accu)} \equiv q^{(0)}$
- $\mathbf{N}^{(falseM)}$ Number of additional (false) in-focus planes; $q^{(Nfalse)} \equiv q^{(1)}$
- **range** Region of monotonic falloff around the central in-focus plane; $q^{(range)} \equiv q^{(2)}$
- **Full Width at Half Maximum (FWHM)** of the sharpness curve; $q^{(FWHM)} \equiv q^{(3)}$

The results of the quality measures $q^{(i)}$ are thresholded by factor u with $M >= u \cdot \max(M)$ to be in the ranges given in Equation 2.4:

$$q^{(0)} \in \left[0, N^{(z)}/2\right], \quad q^{(1)} \in \left[0, N^{(z)}/4\right], \quad q^{(2)} \in \left[0, (N^{(z)}-1)/2\right], \quad q^{(3)} \in \left[0, 4N^{(z)}\right) \tag{2.4}$$

where the upper bound will be called $q^{(lim)}$ and thus the normalization :

$$\bar{q}_i = 1 - \frac{q_i}{q^{(lim)}} \quad , i \in \{0,1,2,3\} \quad , \bar{q}_i \in [0,1] \tag{2.5}$$

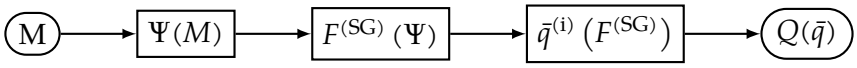


FIGURE 2.2: **Autofocus Processing flow** The Input 3D-stack M is sharpness filtered by a sharpness metric Ψ and the resulting 1D curve afterwards smoothed by a SGF $F^{(SG)}$. The filtered curve is then evaluated by a normalized evaluation measure $\bar{q}^{(i)}$ (Equation 2.5) and finally a total score Q is calculated as a weighted combination of all $\bar{q}^{(i)}$.

can be applied. Note, that the notation $[\ldots]$ means lower and upper limit are contained, while $[\ldots)$ does not contain the upper limit. The automated score-evaluation Q based on the sharpness-metrics Ψ relies on the reliable determination of the in-focus slice within the measured 3D stack M, i. e. the global maximum in the 1D sharpness curve. To avoid erroneous maxima-pickup due to a high noise-content in the input Z-stack an as-good-as-possible maximum-position preserving filter is applied before the normalized quality measures $\bar{q}^{(i)}$ are applied and the score Q is calculated. While the simplest noise reduction can be achieved by low-pass filtering, eg via convolution with a Gaussian Function (GF) (Equation 1.32), this operation tends to shift local maxima. The Savitzky-Golay Filter (SGF) $F^{(SGF)}$ approximates a local sample distribution within a window of uneven size by a polynomial function. Even though the first implementations used least-squares algorithm to find the optimal weighting prefactors of the fitting-polynomial [54] a convolution-kernel based implementation is used. The crucial benefit of SGF as compared to GF are its multiple passbands, the high band-flatness and the very slowly rising attenuation towards high frequencies. This allows for input-smoothing while preserving position and amplitude of data-peaks (e. g. edges) [55]. In this thesis, two parameter of the filter are used: the window size $w^{(SGF)}$ and the polynomial order $p^{(SGF)}$. While the former defines the kernelsize of the filter the polynomial order defines the degree of the polynomial to be fitted to the data. In this work, $w^{(SGF)} \in \{5, 6, \ldots, 13\}$ and $p^{(SGF)} \in \{3, 4. \ldots, 7\}$ are found to be appropriate.

The final score Q is calculated as a weighted combination of the quality measures $\bar{q}^{(i)}$

$$Q\left(\bar{q}\right) = \sum_{i=0}^{3} \omega_i \bar{q}^{(i)} \tag{2.6}$$

with the weights ω_i. The weights are chosen to represent the importance of the quality measures $\bar{q}^{(\mathrm{i})}$ with respect to the characteristics to be displayed by the sharpness-metrics and are normalized to one with respect to the l^1 norm. Within this work they were chosen as:

$$\omega \equiv [\omega_0, \dots, \omega_3] = \left[\frac{1}{2}, \frac{1}{6}, \frac{1}{4}, \frac{1}{12}\right] \tag{2.7}$$
$$\|\omega\|_1^1 = 1$$

to emphasize the importance of finding the correct maximum position ($= 1/2$) over the accuracy ($= 1/6$) and free range ($= 1/4$) while considering the FWHM ($= 1/12$) only peripherally. A flow-graph of the final score calculation is displayed in Figure 2.2.

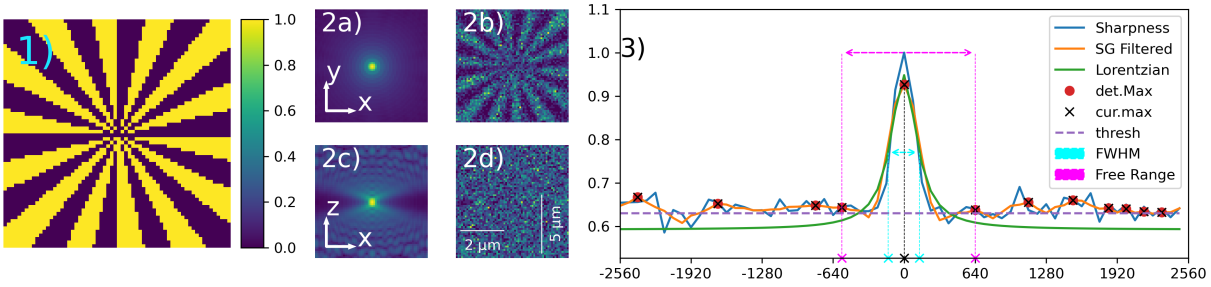
The calculation process of Q is carried out exemplary in case of the $\Psi^{(\mathrm{TEN})}$ for 3 different imaging scenarios on the *spokes2d* target (Figure 2.1) and displayed in Figure 2.3. In the following, the 3 exemplary simulations will be abbreviated as (I),(II), or (III), just as in Figure 2.3. The 3 simulations can be described as testing for a high noise level (I), near ideal imaging (II), and a mixture of aberrated imaging as well as increased noise level (III). The 3D-stacks 2b)+d) in Figure 2.3 are calculated by convolving the system-PSF 2a)+c) with the target-object 1) according to Equation 1.40 while having only 1 view (hence: $N^{(\mathrm{a})} = 1$). The $\Psi^{(\mathrm{TEN})}$ transforms the 3D-stack into a 1D Z-curve and thereby associating one sharpness-value with each Z-position as displayed in 3) (blue line). This 1D-curve is then $F^{(\mathrm{SGF})}$-filtered (orange line) and fitted with an offset-extended Lorentzian function according to Equation 2.8:

$$f^{(\mathrm{Lor})}(x, \mu, \sigma, A_0, A) = A\frac{\sigma}{\sigma^2 + (x-\mu)^2} + A_0 \tag{2.8}$$

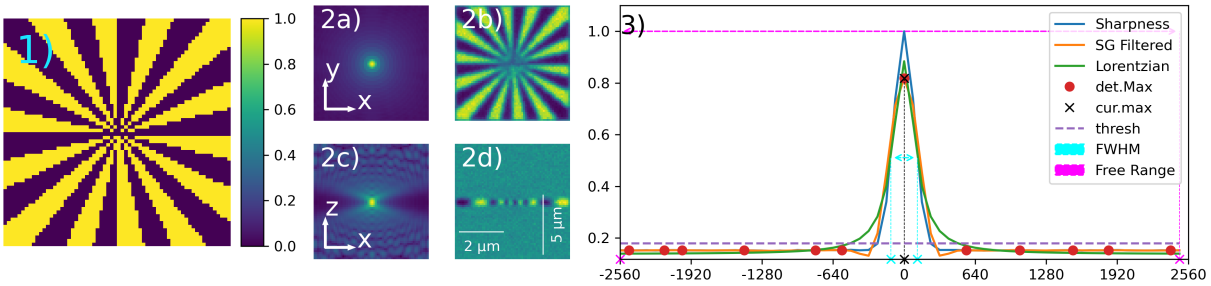
where μ is the center of this symmetric function, $2 \cdot \sigma$ is the FWHM, A the amplitude, A_0 the offset and x the 1D-coordinate. Note that with this definition $\left\|f^{(\mathrm{Lor})}\right\|_1^1 \neq 1$ follows. The FWHM (cyan dashed lines) is used as a measure of maximum emphasis because a steep slope and other invariant areas allow for finding the maximum easier.

The scoring algorithm (Figure 2.2) automatically determines major and minor maxima based on the $F^{(\mathrm{SGF})}$ filtered data in order to calculate the number of false maxima as well as the unique/free area around the major maxima. For this, a generic maxima-searching routine is used in conjunction with a relative threshold (purple) to exclude practically irrelevant minor maxima (red dot without x). All maxima found (red dot with black x) are used for further calculation. The threshold-factor was chosen as $u = 0.07$, yielding

(I) $NPhot = 5$, $c^{(\text{spher})} = 0$, thresholded with $u = 0.07$.



(II) $NPhot = 1000$, $c^{(\text{spher})} = 0$, thresholded with $u = 0.07$.



(III) $NPhot = 100$, $c^{(\text{spher})} = 0.5$, thresholded with $u = 0.07$.

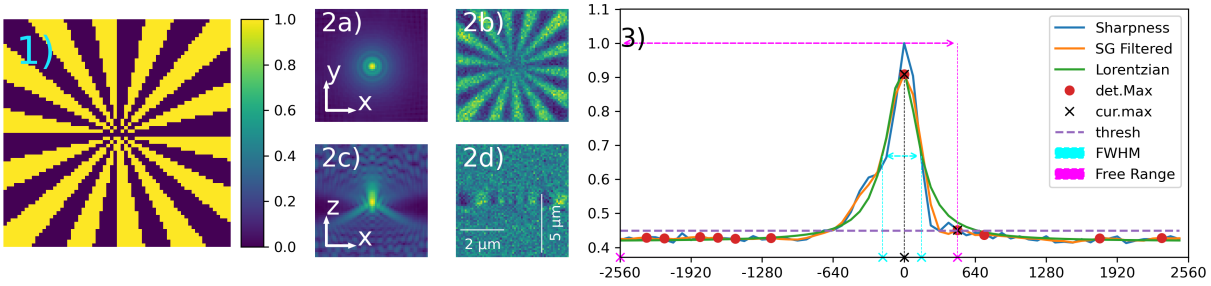


FIGURE 2.3: **Autofocus Quality Assessment Data-Flow** for 3 different PSF-scenarios (I)-(III) using TEN metric. 1) the object, a *spokes2d* target. 2a)+b) X-Y-slice ($z = N^{(z)}/2$) and 2c)+d) X-Z-slice ($y = N_y/2$) of the PSF (left) and resulting image (right). Images of 1-2) use the same colorbar showing the normalized magnitude $M/\max(M)$. 3) calculated sharpness measure (blue), $F^{(\text{SGF})}$-filtered sharpness measure (orange) and Lorentzian-fitted $F^{(\text{SGF})}$-filtered measure (green) from 2b)+d). The main-maxima (central black x), detected Maxima (red dot), selected maxima (black x), calculated FWHM (blue arrow and vertical lines) and free-range (magenta line) are displayed as well as the rejection threshold limit (violet dashed line). From these results the score-values displayed in Table 2.1 are calculated. Horizontal axis shows axial $Z/(nm)$ position and vertical axis the normalized magnitude $M/\max(M)$.

	Accuracy	#False Max	Free Range	FWHM	Total
(I)nn	0.0394	11.0000	7.5000	3.5098	-
(II)nn	0.0004	0.0000	31.5000	2.9493	-
(III)nn	0.2249	1.0000	19.0000	4.3649	-
(I)	0.9988	0.3125	0.2381	0.9726	0.6920
(II)	1.0000	1.0000	1.0000	0.9770	0.9981
(III)	0.9930	0.9375	0.6032	0.9659	0.8840

TABLE 2.1: **Scoring results.** Calculated scores for the 3 different scenarious in Figure 2.3 for TEN-metric. Indices with appended „nn“mark the non-normalized scores and hence no total-score is calculated for these rows.

all side maxima being above the boundary line in the case of (I), no side maxima in the case of (II) and 1 side maxima in case of (III). The distance from the main maxima to the nearest side maxima on both sides is called the free range (magenta dashed line) and may well be asymmetric, as seen in (I). If there are no side maxima, the global limits (size of the image stack) are used. For automated evaluation, the main maxima of the simulated PSF $h^{(\text{aber})}$ are calculated using $\Psi^{(\text{VO5})}$. Wherever possible these maxima- or focus-positions are then curated manually for reasonability.

The scores calculated from the filtered data are listed in Table 2.1. The first 3 rows have an „nn“in the identifier and are not yet normalized according to Equation 2.5, so no total score can be calculated. The last 3 lines are already normalized and a score according to Equation 2.6 was calculated. In the case of (I),(II),(III) $11, 0$ and 1 minor maxima were found, yielding a free range of $7.5, 31.5$ and 19.0. The accuracy is comparably solid for all examples due to the extremely thin sample (=2D). It is noticeable that the aberrations have a larger influence than the number of photons on the determination of the FWHM, since it is clearly larger in the case of (III) than (I). Finally, weighted according to Equation 2.7, we obtain the respective total score Q where (II) as an almost ideal case is also close to the maximum value 1, while the strongly noisy case (I) is even clearly below (III).

Especially the strong weighting of the false maxima and the free range lead to the weak evaluation of (I) compared to (III). This can be influenced by the choice of the threshold, strength and type of the smoothing function (here $F^{(\text{SGF})}$) as well as the type of extreme search. Since the ground truth is not known in the practical measurement

process, a reliable but strict measure should already be chosen in the simulation. Since the configuration (III) is very close to the real experiment, the chosen parameter configuration is to be accepted as suitable.

2.3 METRICS EVALUATION ON *IN-SILICO* DATA

In this section the performance of the presented sharpness metrics will be characterized by 1) computational speed as well as 2) variation of noise, spherical aberration and astigmatism level.

For temporal evaluation (1), even and odd images with side length factor v were evaluated according to:

$$\vec{N}^{(\text{even})} = [N^{(\text{x})}, N^{(\text{y})}, N^{(\text{z})}] = [2^v, 2^v, 2^v] \tag{2.9}$$

$$\vec{N}^{(\text{uneven})} = [2^v, 2^v - (v-1), 2^v - 2 \cdot (v-1)] \tag{2.10}$$

with uniform fill (=non-sparse) generated. The code written in Python was measured with the standard „timeit“package. 1 repetition consists of $N^{(\text{eval})} = 1$ evaluations of the data with the same sharpness-metric and the (python) garbage collector turned off. The resulting time $t/(s)$ was divided by $N^{(\text{eval})}$ to regain the time per evaluation. This process was repeated $N^{(\text{repeat})} = 7$ times. According to the developer side, only the minimum value of the resulting vector of $N^{(\text{repeat})}$ entries should be used as lower-bound to the possible performance of the algorithm on the system [56]. Still, the mean value is used as system-offsets are rather in the region of the system clock-speed, but the relevant times are multiple orders above this threshold. The results are displayed in Figure 2.4. The two panels display the $\log_{10}$ of the minimum time min (t) needed for the processing of the stack of size N generated for the left panel with v according to Equation 2.9 and for the right panel with v according to Equation 2.10. The two metrics $\Psi^{(\text{TEN})}$ and $\Psi^{(\text{VAR})}$ of the conjecture (Section 2.2) are shown in bold. The region highlighted in gray is dominated by general function-inherent processes of data preparation, i.e., functional overhead. These processes do not scale with the amount of data, or only to a very limited extent. From $v = 4$, i.e. images with e. g. of size $[16, 16, 16]$ and thus $N = 2^{12} = 4096$ in the case of cubic data stacks, the algorithm relevant processes dominate and thus the speed is data limited. Although all implemented metrics require different preprocessing, yielding a different offset in the gray region, they all follow roughly the same trajectory and are roughly in the

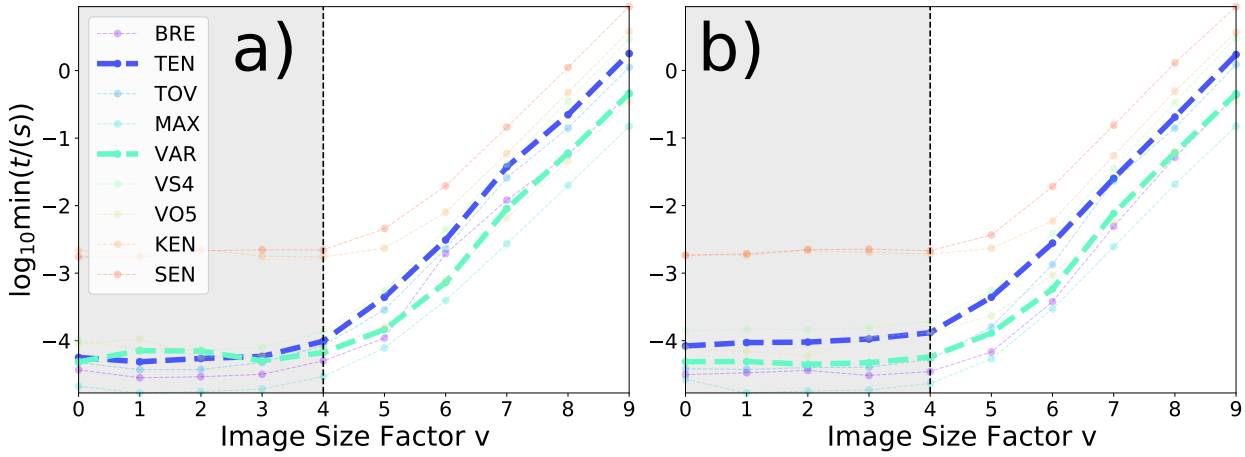


FIGURE 2.4: **Computational Speed Quantification** of the sharpness-metrics for a) cubic and b) non-cubic input. Each data-point resembles the calculated mean-value for $N^{(\mathrm{eval})} = 1$ evaluations and $N^{(\mathrm{repeat})} = 7$ repetitions for each sharpness metric. The grey region marks the preparation-overhead dominant region while the bright region marks the calculation-dominant region.

range $O(N\log(N))$. Stack sizes comparable to further *in-silico* simulations below lie in the range $v = 6$, while stack sizes of practical realizations are even in the range $v = 9$, so e. g. $N = [2^5, 2^{11}, 2^{11}] = [64, 2048, 2048]$. For the practically relevant range $v = 9$, the $\Psi^{(\mathrm{MAX})}$ with $t^{(\mathrm{MAX})} = (0.149 \pm 0.002)s$ is the fastest and the $\Psi^{(\mathrm{SEN})}$ with $t^{(\mathrm{SHA})} = (8.848 \pm 0.137)s$ is the slowest. The thickly marked $\Psi^{(\mathrm{TEN})}$ (blue) and $\Psi^{(\mathrm{VAR})}$ (cyan) metrics are around the mean of all metrics, with $\Psi^{(\mathrm{VAR})}$ with $t^{(\mathrm{VAR})} = (0.457 \pm 0.023)s$ about 3.91× faster than the $\Psi^{(\mathrm{TEN})}$ metric with $t^{(\mathrm{TEN})} = (1.787 \pm 0.027)s$.

Using the *spokes2d* target as an example, the scoring performance of the metrics w. r. t. I) variation of the noise and spherical aberration level simultaneously (Figure 2.5I) as well as II) spherical aberration and astigmatism level (Figure 2.5II) follows. On example of Figure 2.5I the graphic will be explained. The Y-axis shows the mean value $N^{(\mathrm{Phot})} = 1.5^b$ with $b \in \{0, 1, \ldots, 19\}$ used for application of the Poisson noise in the maximum pixel of the simulated image-stack M. On the X-axis the pre-factor $c^{(\mathrm{spher})} \in \{0, 0.05, \ldots, 1.0\}$ used for the phase-map calculation of the ATF in terms of spherical aberration (Equation 2.2) is given. Scoring results are color coded and thus a result of $Q = 1$ is colored yellow while a score of $Q = 0$ is dark blue. In case of II) pre-factor $c^{(\mathrm{spher})}$ is on the Y-axis while on the X-axis the astigmatism pre-factor $c^{(\mathrm{astig})}$ (Equation 2.2) is displayed with $c^{(\mathrm{astig})} \in \{0.0, 0.1, \ldots, 2.0\}$.

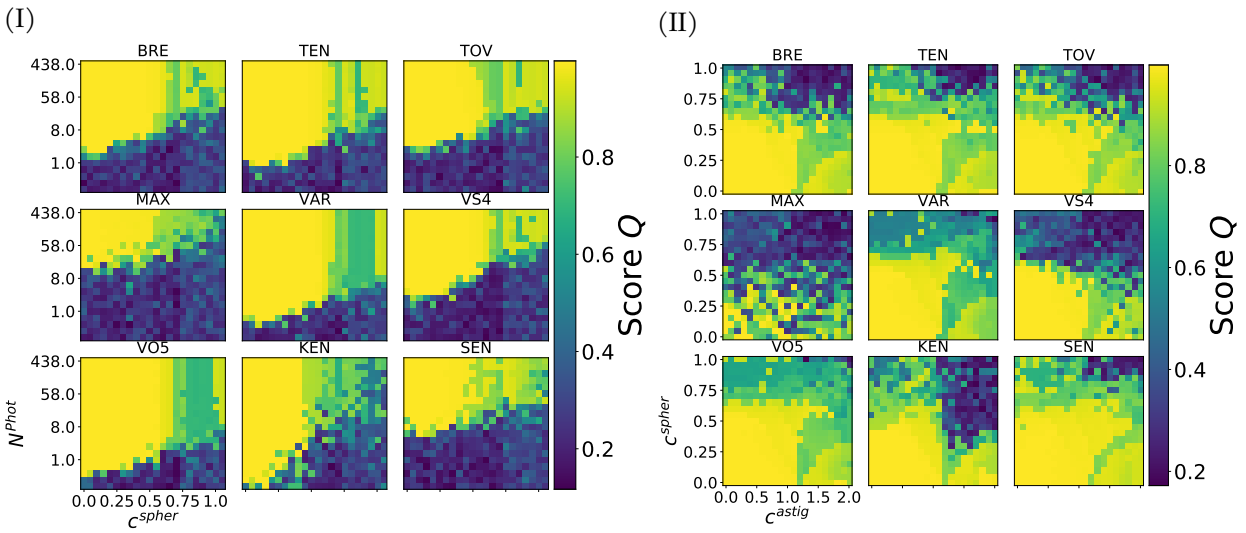


FIGURE 2.5: **Metric Scoring with Combined Simulation Parameter Variation.** Scoring results of sharpness metrics using *spokes2d* thin object (Figure 2.1 $0a) - d)$) for a) different noise levels $b \in [0, 20]$ with $\Delta b = 1$ and $c^{(\text{spher})} \in [0.0, 1.0]$ with $\Delta c^{(\text{spher})} = 0.05$. b) Different spherical $c^{(\text{spher})} \in [0.0, 1.0]$ with $\Delta c^{(\text{spher})} = 0.05$ and $c^{(\text{astig})} \in [0.0, 2.0]$ with $\Delta c^{(\text{astig})} = 0.1$ thresholded with $u = 0.14$.

The simulation Figure 2.5I demonstrates the broad reliability of $\Psi^{(\text{VO5})}$ and $\Psi^{(\text{VAR})}$ for relatively low photon levels ($N^{(\text{Phot})} \approx 1$) and yet high spherical aberrations ($c^{(\text{spher})} \geq 0.6$) with scores $Q \geq 0.6$. Comparable scoring is achieved by $\Psi^{(\text{TEN})}$, $\Psi^{(\text{TOV})}$, $\Psi^{(\text{VS4})}$, and $\Psi^{(\text{KEN})}$ only in the range $c^{(\text{spher})} \leq 0.5$ while $\Psi^{(\text{SEN})}$ achieves high scoring $Q \geq 0.7$ only at high photon levels ($N^{(\text{Phot})} \geq 15$).

Analysis of the simultaneous use of spherical aberrations and astigmatism shows reliable usability ($Q \geq 0.7$) of $\Psi^{(\text{VO5})}$ and $\Psi^{(\text{VAR})}$ over almost the entire simulation parameter range, see Figure 2.5II. $\Psi^{(\text{TEN})}$, $\Psi^{(\text{TOV})}$, and $\Psi^{(\text{SEN})}$ only achieve $Q \geq 0.6$ for $c^{(\text{spher})} \geq 0.7$ and $c^{(\text{astig})} \geq 1.1$, respectively. All metrics, except $\Psi^{(\text{MAX})}$, achieve a scoring around $Q \geq 0.9$ for $c^{(\text{spher})} \leq 0.5$ and $c^{(\text{astig})} \leq 1.1$. As noted previously, the scoring of $\Psi^{(\text{KEN})}$ suffers from the local sensitivity of the metric already for $c^{(\text{astig})} \geq 1.3$.

Note, that $\Psi^{(\text{KEN})}$ and $\Psi^{(\text{SEN})}$ probably detect the main maximum (sub-pixel) more accurately than e. g. $\Psi^{(\text{VO5})}$, but their high sensitivity to side maxima lowers the overall scoring achieved and thus their reliability for automated focus measurements.

The scoring-algorithm was tested for susceptibility to a selection of image degradation. It can be reliably used for keeping the sample in-focus or just tracking of the focus. Focus tracking is directly possible, because the scoring algorithm is invariant under Z-translation. For an in-depth individual analysis of the influence of the noise, spherical aberration and astigmatism on the scoring of the sharpness metrics see Section B.2.

Based on the tests, it was found that $\Psi^{(\mathrm{VAR})}$, $\Psi^{(\mathrm{TEN})}$ and $\Psi^{(\mathrm{VO}_5)}$ provide a good mix of temporal performance and robust implementation at low noise as well as high aberration levels.

2.4 METRICS EVALUATION ON EXPERIMENTAL DATA

The experimental setup is a slight variation of the UC2 fluorescence microscope (infinity) setup [47] and shown in Figure 2.6I. A fiber coupled laser-diode at $\lambda^{(\mathrm{ex})} = 635\,nm$ (150mW, Micost, China) is expanded with a telescope system using a iPhone lens of $f^{(\mathrm{L1})} = 3\,mm$ and an achromatic lens with $f^{(\mathrm{L2})} = 26.5\,mm$ therewith leading to a magnification of the beam-diameter by $M = f^{(\mathrm{L2})}/f^{(\mathrm{L1})} \approx 9\times$. The collimated beam is then focused onto the Back Focal Plane (BFP) of the 100×, 1.25NA oil objective (Zeiss, Germany) by a plano-convex tube lens $f^{(\mathrm{TL})} = 75\,mm$ (Thorlabs, UK). On its way the beam passes a fluorescent filter cube using 740 IY long-pass filter (Comar, UK) and 650/25 (Chroma, Germany) bandpass filter. Alternatively the setup can be used in transmission with an overhead LED-array or desk-lamp. For lateral positioning the UC2-Micronstage and UC2 Z-Stage ([47]) are used. The emitted fluorescence (or transmitted led/bright-light) is finally focused with another tube lens onto an CSI Alvium 1800 C-158 (Allied Vision, Germany) mono-chrome CMOS sensor controlled by a SBC Jetson Nano (Nvidia, USA).

As biological test sample for fluorescence fixed HeLa-cells with Alexa Fluor 647 (AF647) antibody-stained micro-tubules are analysed. Figure 2.6II1-2) show two different ROI within the sample. The images are Z-slice-wise max-normalized for visibility. Images in the upper row show an X-Y-slice at individual $z = z^{(\mathrm{cut})}$. The used positions are shown as gray lines per panel image in the row below. Vise versa: all images in the lower row are shown as X-Z-slice at $y = y^{(\mathrm{cut})}$ where the used cut position is shown as gray lines per panel image in the row above. For testing the filters on an edgy, anorganic but supposedly known sample two different ROI of the non-transparent USAF target (Thorlabs, UK;

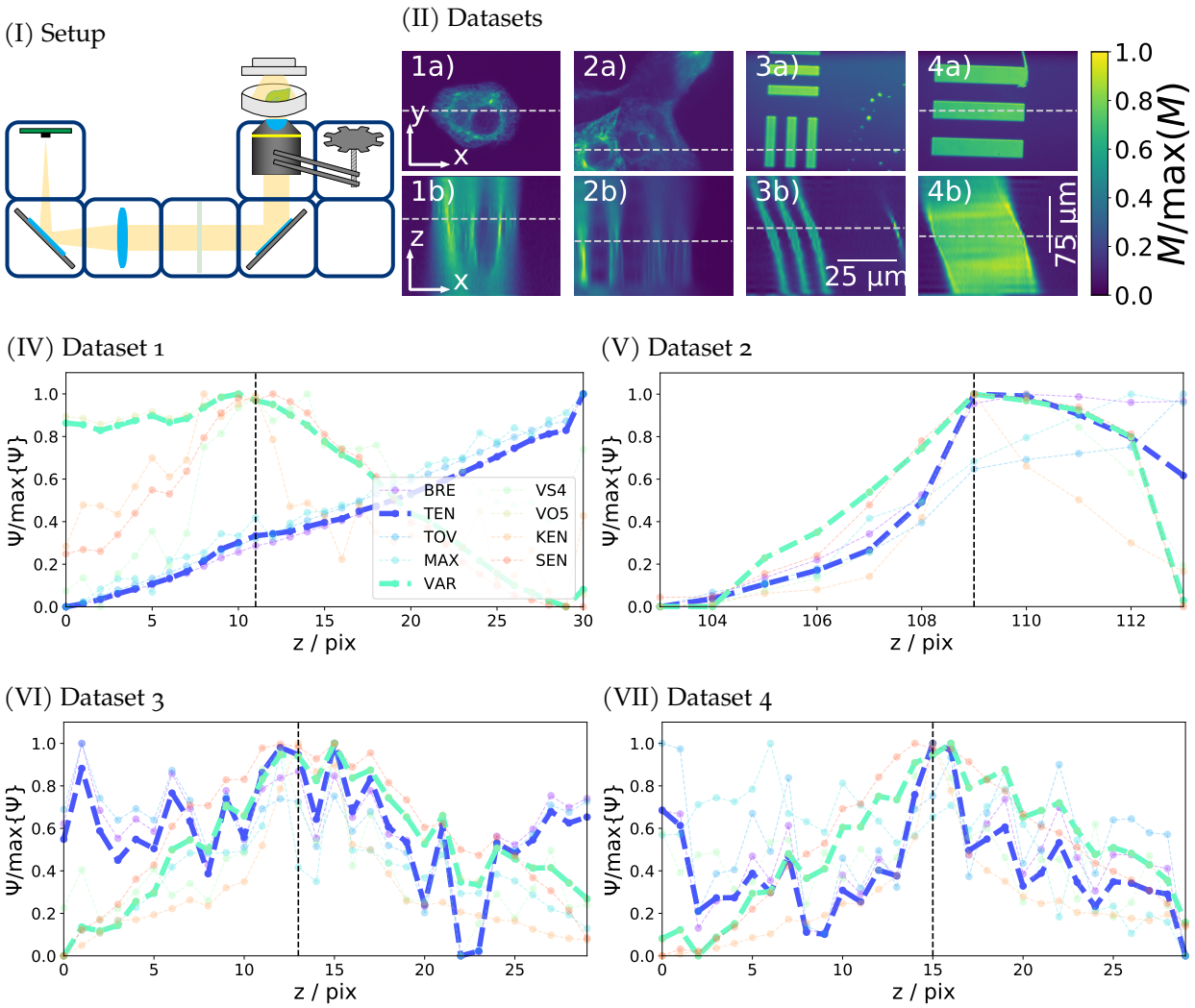


FIGURE 2.6: **Experimental Evaluation of Focus Metrics.** (I) Schematic drawing of the experimental setup using a 3D printed epi-fluorescence WF microscope. (II) Measured datasets: 1+2) HeLa cells at different Region of Interest (ROI)s, 3+4) USAF target at different ROIs. Upper row: X-Y-slice at individual $z = z^{(\mathrm{cut})}$ (gray lines per panel image in the lower row), lower row: X-Z-slice at $y = y^{(\mathrm{cut})}$ (gray lines per panel image in upper row). The images are Z-slice-wise max-normalized for better visibility. (III)-(VI) Sharpness metrics results of the 4 3D-stacks II1-4) .

Figure 2.6II3-4) are used in transmission mode, i. e. is illuminated with the LED from above and the light that propagates through the holes of the target is filtered and detected. The HeLa-cells on the other hand are excited with a laser in epi-fluorescence mode from below. During measurement the illuminating LED was slightly tilted w. r. t. the optical axis

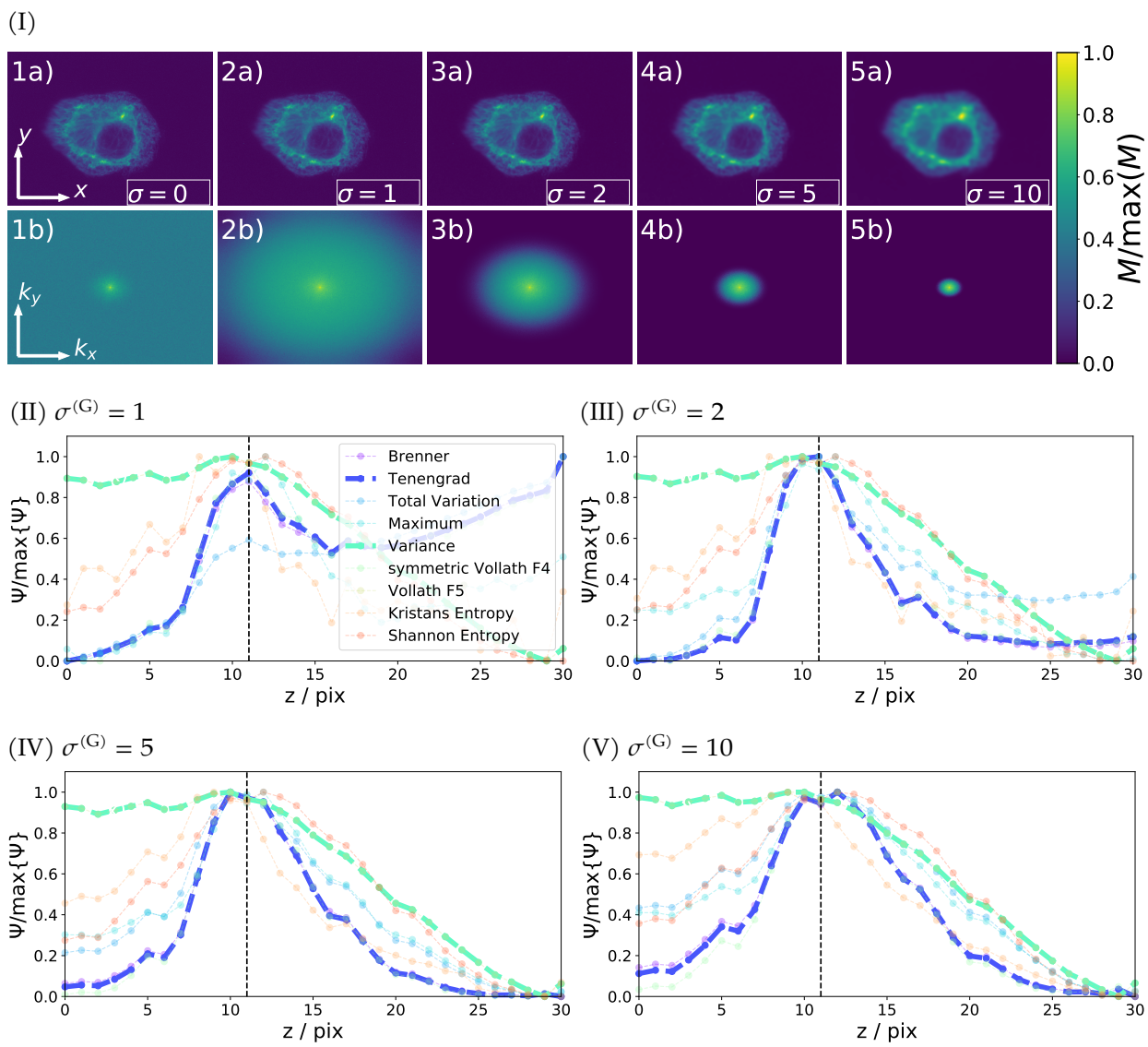


FIGURE 2.7: **Prefiltering of Experimental Data.** (I) Gaussian-prefiltered data using increasing (lateral) Sigmas 2) – 5) of $\sigma^{(G)} \in \{1, 2, 5, 10\}$ pixels. 1) is reprint of non filtered data. Upper row: real space images, lower row: Fourier space representation. (II)-(V) Sharpness metrics results of Gaussian-prefiltered data 2-5).

of the microscope to demonstrate the metrics behavior, see object-tilt in Figure 2.6II3+4b). For both sample types two Z-stacks are acquired at the respective two different (lateral) ROIs.

Even though the samples can be assumed to be thick (w. r. t. the analyzed volume) a position where the sample exhibits the most features over a rather big ROI can be (manually) found at $z^{(mf)} \in 11, 109, 13, 15$ for the individual stacks. Before applying the metrics, the background was subtracted from the stacks and the individual Z-slices were normalized to their individual sums.

The slices at $z^{(\mathrm{mf})}$ are chosen to represent the in-focus slice of interest that shall be found by the sharpness metrics. The panels Figure 2.6IV to Figure 2.6VII display the results after application of the sharpness metrics on the individual Datasets 1-4. Again, the metrics $\Psi^{(\mathrm{TEN})}$ and $\Psi^{(\mathrm{VAR})}$ are shown in bold to highlight them according to the conjecture Section 2.2. Surprisingly, for dataset 1 ($Q^{(\mathrm{SEN})} = 0.95$ and 1.86σ above average, Figure 2.6IV,see note[2]) and 3 ($Q^{(\mathrm{SEN})} = 0.88$, about 1.42σ away from the average) the sharpness metric $\Psi^{(\mathrm{SEN})}$ finds $z^{(\mathrm{mf})}$ achieves the highest score, while $\Psi^{(\mathrm{VO5})}$ or $\Psi^{(\mathrm{VAR})}$ deviate by at least about two Z-slices. For dataset 2, $\Psi^{(\mathrm{KEN})}$ performs best with being about 1.13σ above the scoring average and finds the correct maximum position, while $\Psi^{(\mathrm{TEN})}$ or $\Psi^{(\mathrm{VO5})}$ deviate by two Z-slices. In dataset 4, especially $\Psi^{(\mathrm{TEN})}$ and $\Psi^{(\mathrm{VAR})}$ find the correct maxima positions, where $\Psi^{(\mathrm{TEN})}$ scores about 1.01σ above the average and therewith is highest. The susceptibility of the filters to edges caused by noise is particularly evident in the ooF (or low contrast) of dataset 3 and 4.

In the case of Figure 2.6IIa, the influence of prefiltering will be demonstrated with a spatial 2D GF. The filter is applied on every Z-slice individually and the lateral isotrope kernel sizes are chosen to be $\sigma^{(\mathrm{G})} \in [0, 1, 2, 5, 10]$ pixels, 0 meaning that for the first image no GF was applied, and their influence on the in-focus Z-slice (top row) of Figure 2.7I as well as its Fourier transform (bottom row) is demonstrated. The sharpness metrics with the highest scores when being applied to the GF filtered images, i. e. Figure 2.7I2-5), are: $Q^{(\mathrm{SEN})}$ is 0.98σ, $Q^{(\mathrm{MAX})}$ is 0.79σ, $Q^{(\mathrm{BRE})}$ is 0.71σ and $Q^{(\mathrm{BRE})}$ is 0.86σ above the respective filtered images scoring averages. No clear correlation could be found between GF filter kernel size and standard deviation of the scores Q.

Since especially high frequency information not or only very slightly present in the image can strongly influence the results of edge-based filters, a simple denoising prior to sharpness estimation leads to more reliable results of the differential filters or simplistic filters like $\Psi^{(\mathrm{MAX})}$. For all other filters (Figure 2.7II-V), these pre-filtering processes have little influence until the kernel size significantly reduces the usable frequency information ($\sigma^{(\mathrm{G})} \approx 10$). Thus, in context of the spatial dimensions of the presented experiment a

2 The way of writing means $Q^{(\mathrm{SEN})} = 0.95$, but can be expressed in terms of mean $\mathbb{E}\left\{Q^{(\mathrm{i})}\right\}$ and standard deviation $\sigma^{(\mathrm{Q})}$ over all scoring results $Q^{(\mathrm{i})}$ with $Q^{(\mathrm{SEN})} = \mathbb{E}\left\{Q^{(\mathrm{i})}\right\} + b \cdot \sigma^{(\mathrm{Q})}$. For the scoring results in Figure 2.6IV $\mathbb{E}\left\{Q^{(\mathrm{i})}\right\} = 0.62$ and $\mathbb{S}\left\{Q^{(\mathrm{i})}\right\} = \sigma^{(\mathrm{Q})} = 0.18$, thus $Q^{(\mathrm{SEN})}$ is 1.86× the standard deviation above the average.

simple denoising by means of Gaussian pre-filtering with kernel size $\sigma^{(G)} = 2$ is suggested to still allow for a sharpness curve Ψ of small FWHM while more reliably finding the correct maximum.

2.5 CONCLUSION AND OUTLOOK

In this chapter, different sharpness metrics were implemented in Python and their behavior with respect to different maximal amount of photons as well as aberration levels was investigated. For the evaluation, a scoring function Q was established, which is based on a weighted average of the distance between found focal plane and ground truth, the number of erroneously determined maxima, the distance between major as well as minor maxima and the Full Width at Half Maximum (FWHM) of the Lorentzian that is fitted to the sharpness curve.

For thin samples (2D layer), in the case of spherical and astigmatic aberrations, especially Tenenbaum Gradient (TEN), Variance (VAR) and Vollath F5 (VO5) perform reliably. The analysis of the combination of spherical aberrations and different photon levels shows that Kristans Entropy (KEN) is the best choice at very low photon levels, but VAR, VO5 and TEN achieve the best overall performance.

The temporal performance analysis shows that KEN is about 2 orders of magnitude slower compared to VO5 or VAR. This is due to the necessary DCTs whereby the spectral methods generally entail higher hardware requirements. While the determination of VO5 and VAR is based on the execution of native C code, TEN is not yet optimized for this, yielding a slight performance offset to e. g. VAR.

The presented code is available in a Python and C mixed implementation and could still be fully transferred to C in the future to be adapted and optimized to the target hardware [44].

A justification of the choice of threshold parameters u and weighting factors ω was given, but an analysis of further influencing parameters as well as deeper investigation of the calculation formula of these would be interesting to enable an even more stable evaluation.

For thin samples, in case of unfavorable sampling (due to e. g. too coarse sampling) or two nearly identical slices the focus lies in-between two axial slices. This could in principle lead to erroneous determination of two (or more) instead of one focus layer. The use of axial subpixel sampling or using the center of the fitted Lorentzian solves this problem and thus the scoring model can be used reliably.

The implementation of a part of the presented scoring algorithm was tested for a 3D-printed LED-based transmission microscope in a cell incubator [57]. Although the samples studied in the publication, living differentiating macrophages, cannot generally be considered to be thin samples, this analysis shows that the term *thin sample* in the case of sharpness metrics refers to the relationship between sample thickness and axial resolution. Again, good results were obtained with TEN and VAR, but the Z-stage was sufficiently stable (invariant) for experiment duration such that refocusing was not necessary. Thus, while extending the presented scoring algorithm into a complete auto-focusing routine is straightforward it was not further investigated.

With the presented scoring model a 3D sample object can now be converted into a 1D set of metric-numbers using a chosen sharpness metric Ψ. A potential 3D Drift of the sample object over time can be tracked by Normalized Cross Correlation (NCC) comparison of the calculated sharpness numbers of each time point. This approach requires little storage and hence empowers limited hardware such as a Raspberry PI especially when using methods like TEN, VAR and VO5, whereby TEN can be used to reduce the necessary RAM further while keeping CPU load low.

As a global resumee taking $Q \geq 0.7$ can be assumed to be a reasonable and justifyable level for keeping the focus position and therewith choosing a useful metric for a given situation. The demonstrated analysis can thus be used as a guide on which metric to be used when.

3

JITTER CORRECTION

If you only read the books that everyone else is reading, you can only think what everyone else is thinking.

— Haruki Murakami

3.1 MOTIVATION

For this section it is assumed that while the focus is kept reliable stable the quality of the recording is reduced by spontaneous sample position changes. The goal is to be able to reconstruct few pixel size fluorescent structures on the retina of a living mouse using 2-photon Confocal Laser Scanning Microscopy (CLSM). Due to the comatose state of the mouse this is generally possible, but spontaneous muscle twitching, breathing and general drifting movements cause random pixel and line offsets (jitter). The present jitter problem is modeled as a one-dimensional problem for the case of no or small gaze angle changes of the mouse, simulated and confirmed by means of available experimental data. The case of a large change in viewing angle is excluded from the definition of jitter and classified as global drift. While comparatively large fluorescently stained structures (e. g. blood vessels) can be tracked even despite jitter or change of viewing angle and thus motion-induced distortions are potentially easier to fix, the reconstruction of few pixel-sized structures (excitable at a different wavelength than the blood vessels) such as e. g. mouse retina infiltrating T helper cells becomes impossible. In this section, a method to counteract the jitter using trackable structures (e. g. blood vessels) and afterwards applying deconvolution to allow for a reliable observation of the few pixel sized structures of interest (i. e. T helper cells) is presented. To measure the mouse-retina *in-vivo* using a 2-photon fluorescence CLSM[1] the mouse-body not only needs to be fixed and the imaging optics to be adapted to the refractive properties of the mouse eye, but the residual movement of the anesthetized mouse needs to be taken care of.

1 For the purposes of this thesis the imaging model of 2-photon is not further discussed, because it is simply implemented as CLSM with a squared excitation Point Spread Function (PSF) $h^{(\mathrm{ex,2P})} = \left[h^{(\mathrm{ex,CLSM})}\right]^2$.

Despite suitable sample preparation and first Z-stack preprocessing motion artifacts can still occur depending on the object being analyzed as well as the imaging methodology. In Widefield Microscopy (WF) imaging, e. g. in the case of Structured Illumination Microscopy (SIM), such artifacts can occur in feature-sized image sections (e. g. an entire cell) if the motion velocity of the object being analyzed is greater than a limiting temporal sampling frequency of the imaging method (including reconstruction) [58]. In the case of CLSM or Image Scanning Microscopy (ISM) imaging geometries objects can thus be distorted on a point-by-point and line-by-line basis. Distortions in the measured image can further occur due to spontaneous vibrations or errors in the scanning mechanics.

While spontaneous compressions of an object along a line (and without a temporal series of measurements) can be corrected only with difficulty without previous knowledge, the offset of individual lines as well as of blocks can be corrected.

The *spokes2d* (Figure 2.1a+b) target will be used as *in-silico* simulation target, because it can be thought as a strong abstraction of the blood vessels emanating from the optic nerve on the retina and further exhibits a broad range of spatial frequencies which can be used to demonstrate the relative influence of jitter w. r. t. the underlying structures size. In particular, over the axial extend of the imaging PSF the retina blood vessels change their shape only slightly which will be exploited in the algorithm to use them for drift correction and basic dejittering.

In this section, line-by-line offsets in the case of CLSM geometry will be modeled and simulated. For the correction of these offsets an algorithm is presented and evaluated on experimental *in-vivo* mouse-retina data. Finally, the data will be deconvolved and compared to uncorrected deconvolved data.

3.2 IMAGING MODEL FOR MOVEMENT CORRECTION

For modeling purposes, it is assumed that the spontaneous movements of the object can occur at any time, but only along the slow scanning axis, i. e. from line m to $m + 1$. The distortions of the object happen erratically but slower than the line-by-line scan speed, and the object undergoes a uniform offset and return to the initial position over a set of lines. This excludes e. g. whole image jumps, which however can also be detected with cross correlation approaches and lead to a reduction of the usable field of view (FoV).

Compressions/Distortions (along X) within the lines are not modeled, since under usual CLSM (non-spinning disk) pixel-dwell times of a few μs and scan geometries of e. g. 1024×1024 the scan of a line is in the range of ms. Assuming Nyquist-correct sampling for e. g. a 2-photon fluorescent CLSM system at $NA = 0.28$, $\lambda^{(\mathrm{ex})} = 900\,\mathrm{nm}$, $\lambda^{(\mathrm{em})} = 608\,\mathrm{nm}$ and $n^{(\mathrm{im})} = 1.33$ water gel, the maximum allowed samling is $d^{(\mathrm{SPL,lat})} \approx 402\,\mathrm{nm}$, $d^{(\mathrm{SPL,ax})} \approx 7548\,\mathrm{nm}$ and a typical scan-volume is about $N^{(\mathrm{z})} = 30$ pix meaning $FOV^{(\mathrm{z})} \approx 226\,\mu\mathrm{m}$ and $N^{(\mathrm{y})} = N^{(\mathrm{x})} = 997$ pix meaning $FOV^{(\mathrm{x})} = FOV^{(\mathrm{y})} \approx 401\,\mu\mathrm{m}$. Thus the pixel scan speed is about $v^{(\mathrm{scan})} \approx 402\,\mathrm{nm}/5\,\mu s \approx 80\,\mu m/ms$. Mammalian monocytes are used as a small, structured reference object that is close to e. g. T helper cells. They average about $\approx 10\,\mu\mathrm{m}$ ([59]) in size and move e. g. in the case of rapid infiltration during wound closure of a mouse skin at about $v^{(\mathrm{mono})} \geq 12\,\mu m/min \approx 0.2\,\mathrm{nm}/ms$ [60]. Thus, local distortions arising from the motion of these cells are negligible in the context of scan speed ($v^{(\mathrm{scan})}/v^{(\mathrm{mono})} \approx 4 \cdot 10^5$) and achieveable resolution ($d^{(\mathrm{SPL})}/2$). Drifts along several lines, on the other hand, can be generated by movement of the whole object or errors in the scan arrangement and shall be considered further here. The proposed model focuses on lateral distortions because the axial extent of the PSF $\approx 18\times$ is larger than the lateral extent and thus the considered distortions are negligible at the axial scale. Row-wise distortions, i. e. along the X-axis from l to $l+1$, are potentially relevant and can be represented via a two-dimensional jitter model. However, in the context of the experimental data analyzed in this work, simulation and correction using a one-dimensional model proved sufficient.

The above distortions occur randomly and are called *jitter*. Under the above assumptions, the forward model for this process can be described by Equation 3.1

$$M = \mathcal{P}\left\{D^{(1\mathrm{D})} \otimes (h \otimes S)\right\} \tag{3.1}$$

$$D^{(1\mathrm{D})}(x) = \sum_{n=0}^{N^{(\mathrm{z})}} \sum_{m=0}^{N^{(\mathrm{y})}} \delta_{mn}(x - \hat{x}_{mn}) \tag{3.2}$$

where $D^{(1\mathrm{D})}(x)$ is a one-dimensional convolution kernel describing row-wise shifts of a 3D stack and $\hat{x}_m$ are randomly drawn from a uniform distribution in $[0, \hat{x}^{(\mathrm{lim})}]$. The sum over Z and Y resembles that jitter can be distributed over all lines (Y) of each slice (Z) of the whole Z-stack. Note, that the sample is scanned pixel-wise and thereby convolved with the system PSF. If the imaging process is thought as imaging every scanning position individually (Equation 1.17) the jitter has to be added as a line-wise shift before convolution with the system PSF to include the random sample movement. Adding this additional shift to each

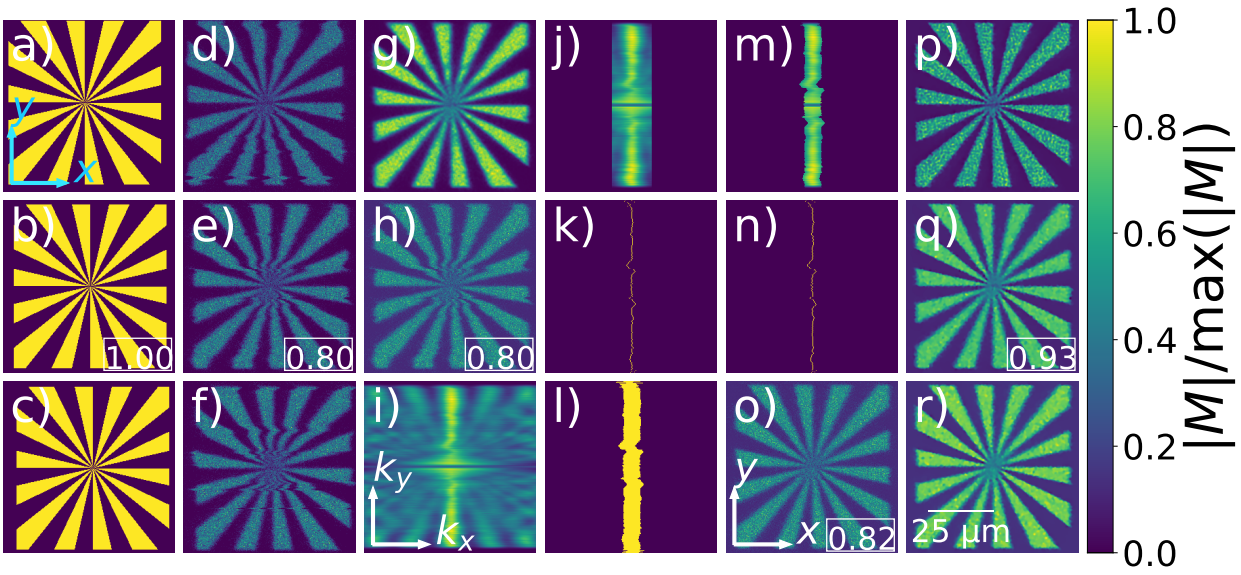


FIGURE 3.1: **Forward Data-Flow** of movement-correction on example of a $N^{(z)} = 16$ thick set of copies of *spokes2d* object. a-c) linearly Z-drifted raw-object at $z_{-4}, z^{(o)}, z_2$. d-f) 2-photon CLSM, Z-drifted, jittered image of a-c. g) mean-projection along z of the jittered stack. h-o) display of processing steps on example of slice $z^{(o)}$. h) image after drift correction $M^{(\mathrm{dcor})}$, i) correlation map in X-direction $L_{l'm'n'}$, j) rectangular windowed correlation map $L^{(\mathrm{win})}_{l'm'n'}$, k) mask from linewise maxima $B_{l'm'n'}$ found from windowed correlation map, l) dilated maxima $B^{(\mathrm{dil})}_{l'm'n'}$, m) multiplied with correlation map $L^{(\mathrm{dil})}_{l'm'n'}$, n) center of mass $L^{(\mathrm{1D})}_{m'n'}$, o) dejittered image $M^{(\mathrm{dej})}_{l'm'n'}$. p-r) dejittered and deconvolved results $\hat{S}_{lmn}$ of d-f). NCC results printed in lower right box of panels b,e,h,o and q.

scanning step and scanning over the whole image, as was demonstrated for derivation of the CLSM forward model (Equation 1.18), the integrals can be rewritten such that the jitter can be included as an additional convolution after the sample was convolved with the system PSF, but before measurement and thus application of the noise model.

3.3 ALGORITHM AND ITS EVALUATION ON *IN-SILICO* DATA

For the reconstruction it is assumed that the image data remain in the same focus volume (w. r. t. the object) over the entire measurement stack according to Section 2.2 and have been processed according to Section A.14. Flat-fielding and sum normalization of the individual images allow pixel-value bias and thus systematic error in shift determination to be avoided. Based on this, the novel movement correction algorithm is applied.

On example of a linear lateral drifted Z-stack of *spokes2d* targets the algorithm will be explained. For a pseudocode-implementation see Algorithm 2. The description is supported by Figure 3.1. In brief, Figure 3.1 a-g) display a set of steps to create the *in-silico* dataset with: a-c) the drifted dataset at different slices $z \in \left\{z^{(-4)}, z^{(0)}, z^{(2)}\right\}$, d-f) the jiterred noisy imaged Z-stack of the target at the same Z-positions and g) the mean projection along Z of the jittered Z-stack. From h-o) steps of the movement correction algorithm on example of the central Z-slice $z = z^{(0)}$ are shown. The stack is h) lateral drift corrected, i) a 1D correlation map along X of the image Z-slice with the mean projection from g) is calculated and j) a rectangular window is applied to the 1D correlation map. Then k) the linewise maxima of j) are determined and extended using binary dilation thereby l) creating a dilated-maxima mask of k). The mask is then m) multiplied with the correlation map i) and the n) a Center of Mass (CoM) (see Equation 3.5) is calculated linewise along X of l). Finally, the CoM positions are used to shift back each line individually yielding o) the dejittered Z-slice. The last column p-r) shows the deconvolution results of the dejittered Z-stack at the same Z-slices as a-c). In the following, the individual processing steps will be explained in detail.

For data generation, each Z-slice of the sample is shifted along X for $s^{(\mathrm{lin,x})}$ pixels and Y for $s^{(\mathrm{lin,y})}$ pixels according to Equation 3.3, i. e. for the e. g. 5th Z-slice the sample is shifted by $s^{(\mathrm{lin,x})} = 10$ and $s^{(\mathrm{lin,y})} = 5$ pixels from its original position at Z-slice $z^{(0)}$. Afterwards, jitter is generated from the *spokes2d* (Figure 3.1a-f) to demonstrate the algorithm. A 2-photon fluorescence CLSM system with the above optical parameters is simulated. The imaging volume is $N = [N^{(\mathrm{x})}, N^{(\mathrm{y})}, N^{(\mathrm{z})}] = [256, 256, 16]$ pix with a pixelpitch of $d^{(\mathrm{scan})} = [200, 200, 3500]\,\mathrm{nm}$. There will be Poisson noise with $N^{(\mathrm{Phot})} = 10$ expected photons in the maximum voxel, a linear drift according to Equation 3.3 as well as random jitter. The jitter is added by the following procedure: 1) the amount $N^{(\mathrm{y,jit})}$ of Y-lines affected by the jitter is randomly drawn from a uniform distribution over the range $[0, 0.1] \cdot N^{(\mathrm{y})}$ of all available Y-lines. 2) $N^{(\mathrm{y,jit})}$ affected lines $y_m^{(\mathrm{jit})}$ are drawn from the uniform integer distribution $[0, N^{(\mathrm{y})} - 1]$ and 3) the shifted distance $s_m^{(\mathrm{jit,x})}$ of each $y_m^{(\mathrm{jit})}$ is drawn randomly from the uniform distribution $\hat{x} = [0, 0.05] \cdot N^{(\mathrm{x})}$. A local retardation, i. e. a 1D blur-kernel $\sigma_m^{(\mathrm{jit,y})}$ along Y-direction, is applied and its size randomly drawn from a uniform distribution $[-4, 4]$ pixels. See Algorithm 1 for a pseudo-code like notation of the described algorithm.

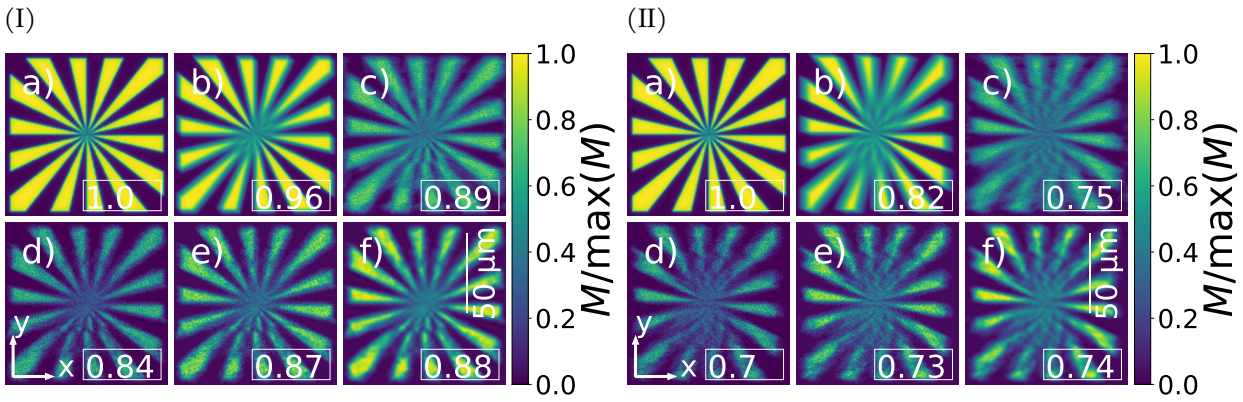


FIGURE 3.2: **Comparison of Reference Projection** for (I) linear ($s \in [-16, -14, \ldots, 14]$) and (II) non-linear sample drift of reference *spokes2d* stack. a) mean-projection of non-jittered and non-drifted image. b) Mean-projection of non-jittered but drifted image, c) mean- d) median- e) 2D-sequential Savitzky-Golay Filter (SGF) ($w^{(\mathrm{SGF})} = 5$) prefiltered median- f) SGF ($w^{(\mathrm{SGF})} = 7$) prefiltered median-projection of jittered, drifted and preprocessed image M_{lmn}. NCC results are displayed in lower right box in each panel.

From the Z-stack M_{lmn} an axial reference projection $M_{lm}^{(\mathrm{ref})}$ (Figure 3.1g) is determined and used to correct the axial drift by finding the shift value between the reference projection and each Z-slice using the positional offset of the maximum of the 2D image correlation between the two compared images. From the drift-corrected data stack $M_{lmn}^{(\mathrm{dcor})}$ (Figure 3.1h), an updated reference $M_{mn}^{(\mathrm{ref}')}$ is determined. Since based on the reference $M_{mn}^{(\mathrm{ref}')}$ a line in $M_{lmn}^{(\mathrm{dcor})}$ can be detected as jitter, special care has to be taken in determining a representative reference. Two different scenarios will be demonstrated. 1) linear drift (see Figure 3.2I) and 2) non-linear drift (see Figure 3.2II). The individual lateral shifts w. r. t. to the axial position z are chosen as:

$$s^{(\mathrm{lin})}(z) = [s^{(\mathrm{lin,x})}, s^{(\mathrm{lin,y})}] = [2 \cdot (z - z^{(\mathrm{o})}), z - z^{(\mathrm{o})}] \tag{3.3}$$

$$s^{(\mathrm{nlin})}(z) = [[s^{(\mathrm{nlin,x})}, s^{(\mathrm{nlin,y})}] = [0.9 \cdot (z - z^{(\mathrm{o})})^{1.3}, 0.2 \cdot (z - z^{(\mathrm{o})})^{1.5}] \tag{3.4}$$

for the linear $s^{(\mathrm{lin})}$ and non-linear $s^{(\mathrm{nlin})}$ case. The effect of the different shift on the reference projection calculation is demonstrated in Figure 3.2. The processing-steps of subfigures (I) and (II) are identical and will be explained on example of (I). a) shows the mean-projection of the non-jittered and non-drifted 3D image-stack and b) the mean-projection of non-jittered but linear drifted image (Equation 3.3). c-e) show the Z-projection $M^{(\mathrm{ref})}$ of the jittered and drifted Z-stack M using a c) mean- d) median- e) 2D-sequential SGF

with window size $w^{(\mathrm{SGF})} = 5$ prefiltered median- f) SGF with $w^{(\mathrm{SGF})} = 7$ prefiltered median-projection for calculation of the reference $M^{(\mathrm{ref})}$. (II) uses a Z-stack of *spokes2d* targets drifted non-linearly according to Equation 3.4.

Edges and jitters are more uniform in the mean projection than in the median projection, but are more blurred and thus are more difficult to determine for correlation. Preliminary processing with an SGF of chosen window size $w^{(\mathrm{SGF})}$ reduces the amount of noise in the projection, reducing random noise correlations between individual slices and the projection. For symmetric lateral drifts about an axial center position of a axially slowly changing structure, like is the case for blood vessels abstractly demonstrated by a Z-stack generated from a slice-by-slice laterally shifted *spokes2d* target, median-projection is favored. The same procedure nevertheless leads to comparable findings for nonlinear, axially asymmetric drifts according to Equation 3.4 in Figure 3.2II. Note that the offset of Figure 3.2IIb is asymmetric due to shift-dependent security padding in contrast to e. g. Figure 3.2Ib.

In the following dejittering step a bounding-shield of a few multiples of the lateral *FWHM* is added to the outside of the image to reduce wrap-around artifacts by Fourier-based operations. The correlation map $L_{l'm'n'}$ is calculated by computing a 1D correlation of each drift-corrected Z-slice of $M^{(\mathrm{dcor})}$ individually with the calculated 2D reference image $M^{(\mathrm{ref})}$ along X-direction. The correlation map L may have multiple, highly extended correlation maxima in pathological cases (e. g. periodic objects) or large maxima-shifts due to highly offset lines (e. g. due to noise). To reduce this effect, local 1D windowing along X around the center, i. e. a 1D rect-function (Figure 3.1j), is multiplied thus limiting the influence of maxima at far stretched X-positions from the center. Optionally, instead of the rect-function a 1D Gaussian-function can be used to select the central region with an attenuated window rather then just cutting it out. Within the central region of the correlation map L line-wise correlation maxima $L^{(\mathrm{max})}_{l'm'}$ (Figure 3.1k) are selected.

Optionally, (1D) SGF filtering of the data set can be performed beforehand to reduce the influence of high-frequency noise in determination of the maximum position. Then, a mask is generated from the selected maxima via application of a dilation operation in X-direction (Figure 3.1l) and multiplied with the correlation map $L_{l'm'n'}$ thereby yielding the selection $L^{(\mathrm{dil}')}_{l'm'n'}$ (Figure 3.1l). To undo the line-wise jitter along X the offset of the line-wise maximum from the line-center position in $L^{(\mathrm{win})}_{l'm'n'}$ is determined via 1D Center of Mass (CoM) (Figure 3.1n) calculation $L^{(\mathrm{1D})}_{m'n'}$ along X-direction according to:

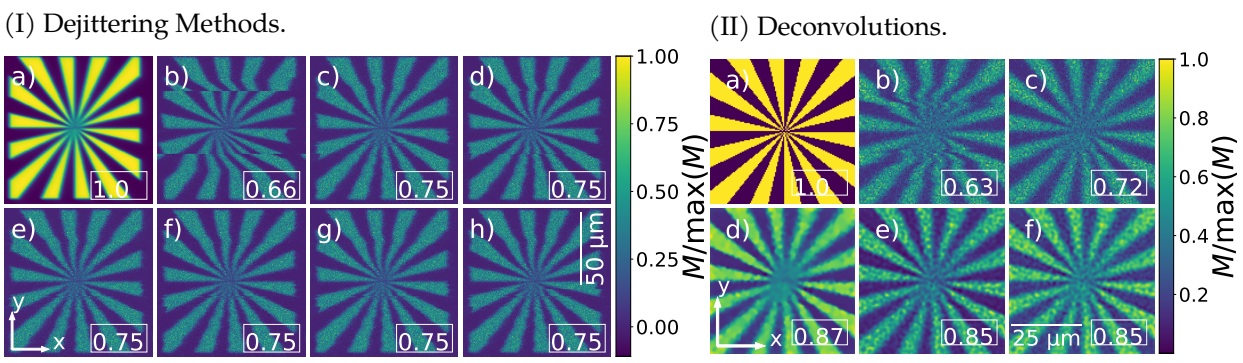


FIGURE 3.3: **Dejittering and Deconvolution Comparison.** (I) Dejittering (displayed at z_{-4}) of linear lateral drifted data and (II) of regularized single-view DEConvolution (DEC) (Table 1.2) in case of linear lateral drift. (I)a) non-drifted, non-jittered noise-free raw-image, b) jittered and drifted image, c-h) dejittered images $M^{(\mathrm{dej})}$ after application of c) CoM of mutual line-wise correlation without window, d) CoM using Gaussian window ($\sigma^{(\mathrm{G})} = 30\,pix$), e) CoM using rect windowed (windowsize $\Delta x = 30\,pix$), f) SGF just using no-window but direct maxima, g) SGF using a rect-window with CoM or h) SGF using Gaussian-window with CoM to b). (II)a) ground-truth S, b) image $M^{(\mathrm{raw})}$, c) dejittered and drift-corrected image $M^{(\mathrm{dej})}$, d-f) DEC with regularizers d) TV, e) GR, f) GS. NCC results are displayed in lower right box in each panel.

$$L^{(\mathrm{1D})}_{m'n'} = \frac{\sum_x x \cdot L^{(\mathrm{dil'})}_{l'm'n'}}{\sum_x L^{(\mathrm{dil'})}_{l'm'n'}} \tag{3.5}$$

A comparison of some possible dejittering configurations is demonstrated in Figure 3.3I. Here, a) is a Z-slice at $z = z^{(-4)}$ of the linear drifted (Equation 3.3) ideal noise-free and non-jittered 3D image and b) shows a) after application of noise and jitter. c-h) show individually processed results of b) where different filters and search routines where combined for comparison. c-h) show dejittered images $M^{(\mathrm{dej})}$ using c) CoM of mutual line-wise correlation without window, d) CoM using Gaussian window with $\sigma^{(\mathrm{G})} = 30\,\mathrm{pix}$), e) CoM using rect window of size $\Delta x = 30$ pix, f) SGF just using no-window but direct maxima, g) SGF using a rect-window with CoM and h) SGF using Gaussian-window with CoM.

The line shifts in Figure 3.3Ib are better reduced in this case by a completely open window (Figure 3.3Ic) or only unattenuated by rect function bounded window (Figure 3.3Ie) than by a Gaussian-attenuated window (Figure 3.3Id), since this better incorporates the tails of the correlation function. Since this is not considered a robust implementation for pathological configurations, as described above, a combination of SGF, Gaussian-windowing, and CoM computation (Figure 3.3Ih) is proposed. The results Figure 3.3If-h, i. e. the combination of SGF without window and with rect- or Gaussian-windowing, all provide reliable and comparable results and the combination Figure 3.3Ih can be recommended for general processing.

The determined shifts for line-wise dejittering are then applied to the Z-stack (Figure 3.10) and the results are deconvolved on a laptop (for specifications see Section 1.8) with DEC using Total Variation (TV) regularization with $\gamma^{(\mathrm{TV})} = 10^{-2.5}$ where convergence is reached around $N^{(\mathrm{iter})} = 180$ (Figure 3.1p-r).

The parameters used for the DEC were determind via a simple parameter-search as follows. The tested regularizers are TV (Equation 1.56), Good's Roughness (GR) (Equation 1.57) and Gradient Square (GS) (Equation 1.58), see Figure 3.3II. Here, a) shows the *spokes2d* target used as ground-truth S for the parameter-search, b) the image $M^{(\mathrm{raw})}$, c) the dejittered and drift-corrected image $M^{(\mathrm{dej})}$ and d-f) the 2D DEC reconstructions $\hat{S}_{mn}$ with regularizers d) TV with $\gamma^{(\mathrm{TV})} = 10^{-2.5}$, e) GR with $\gamma^{(\mathrm{GR})} = 10^{-2.5}$, f) GS with $\gamma^{(\mathrm{GS})} = 10^{-3.5}$. The ground-truth a) is an edged target which is smooth along its struts. Of the three priors compared, the TV is the most likely to have a local smoothing but edge-texturing property.

On example of TV, the dependence of the DEC result on the choice of the regularization parameter is shown in Figure 3.4. Here, a) is the ground-truth *spokes2d* target, b) is the noisy jittered and c) the noisy dejittered image. d-m) show 2D DEC results for different $\gamma^{(\mathrm{TV})} = 10^{-m}$ for $m \in [-1, -0.5, \dots, 3.5]$ where convergence was reached around $N^{(\mathrm{iter})} \leq 200$ iterations. The NCC (Equation 1.38) best agreement of the reconstruction $\hat{S}$ with the ground-truth S (Figure 3.4a) for $\gamma^{(\mathrm{TV})} = 0.01$ (Figure 3.4j) and thus $C = 0.89$. For comparison, the same parameter search in case of using GS regularization is given in Section C.3.

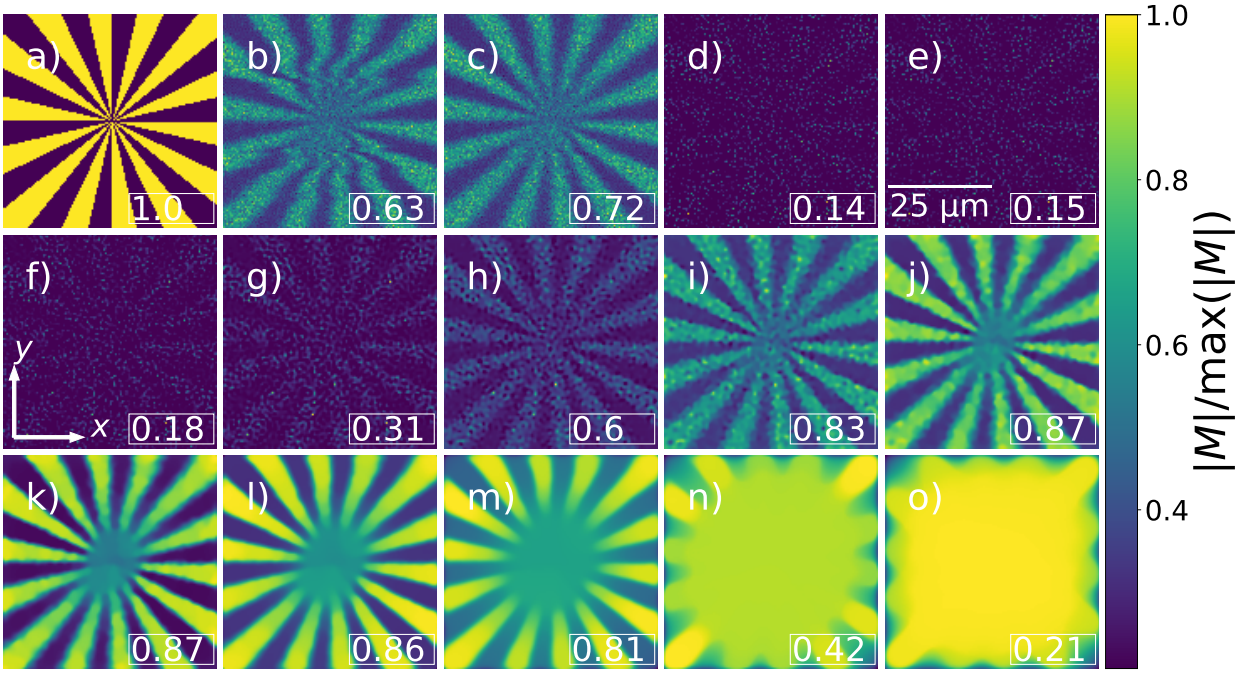


FIGURE 3.4: **TV-Deconvolution Parameters.** Results $\hat{S}_{l^{(o)}mn}$ of processing and tiled deconvolution of a) 3D obj, b) jittered image, c) dejittered image d-m) using TV regularizer with $\gamma^{(\mathrm{TV})} \in 10^{1}, 10^{0.5}, \ldots, 10^{-3.5}$. NCC results are displayed in lower right box in each panel.

The search for suitable parameters depends on many factors like the used update procedure (here: Limited-memory Broyden-Fletcher-Goldfarb-Shanno algorithm (L-BFGS-B)), regularizers or the features mainly expected in the image. The parameter hyperspace search is its own big field of research.

In Figure 3.5 a brute-force two-dimensional search is demonstrated, where the parameters were varied in the ranges: iteration number $N^{(\mathrm{iter})} \in \{1, 4, \ldots, 94\}$ and the regularization parameter $\gamma = 10^{-m}$, $m \in \{-3, -2.5, \ldots, 1\}$. To speed up the search, tiled-processing was omitted here and only a subimage of size $N = [128, 128, 16]$ around the center pixel was used. This is the area that is significantly degraded by the imaging process compared to the raw data, yielding a slightly lower correlation result as compared to the results of the complete stack, see Figure 3.4. All three sub-images of Figure 3.5 show a surface-plot where on the X-axis the iteration number $N^{(\mathrm{iter})}$, on the Y-axis the $\log_{10}$ of the parameter γ and on the Z-axis the NCC result is plotted. The NCC result is further color coded from darkblue ($C = 0.2$) to red ($C = 0.8$). The resulting 2D surface in case of TV shows that

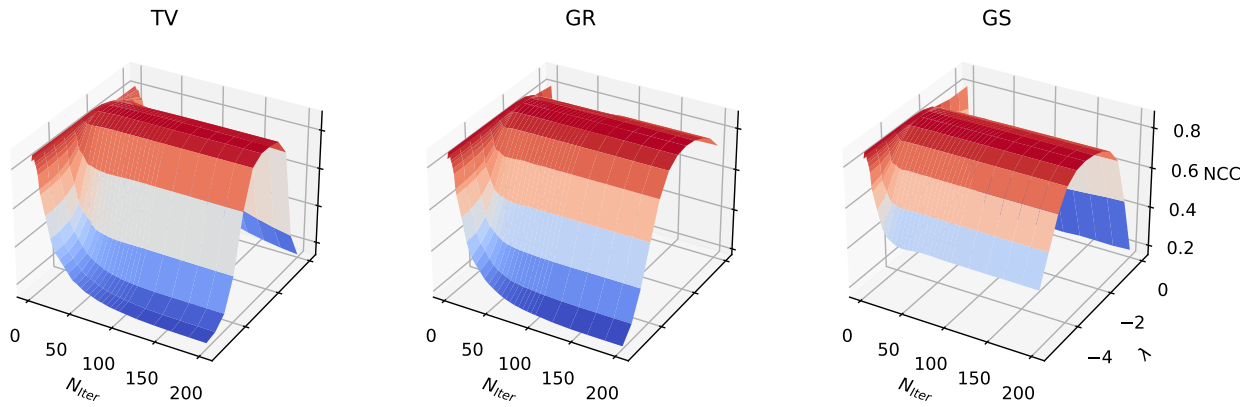


FIGURE 3.5: **2D-Search for Optimal Parameters.** NCC results of 2D-Parameter search for deconvolution of dejittered *spokes2d* stack using 3 regularizers TV,GR,GS with $N^{(\text{iter})} \in \{1, 2, \dots, 10, 15, \dots, 100, 120, \dots, 200\}$, $\gamma = 10^{-m}$ with $m \in \{-5.5, -5.0, \dots, 0.0\}$. The NCC is color coded from darkblue ($C = 0.2$) to red ($C = 0.8$).

only a few iterations $N^{(\text{iter})}$ suffice for convergence towards the maximum NCC, shown in dark red. The NCC looks like a Gaussian centered on the optimal $\gamma^{(\text{TV})} = \gamma^{(\max)}$, i. e. it decreases fast for $\gamma^{(\text{TV})} \neq \gamma^{(\max)}$.

A maximum search using the determined NCC values gives e. g. for TV the optimal configuration: $\gamma^{(\text{TV,max})} = 10^{-2.5}$ and $N^{(\text{iter,TV,max})} = 43$ thereby achieving a correlation of $C^{(\max)} \approx 0.77$. The minimum $C^{(\min)} \approx 0.271$ for this simulation is found at $\gamma^{(\text{TV,min})} = 10$ and $N^{(\text{iter,TV,min})} = 94$. The difference is remarkable. Table 3.1 lists all found minimum and maximum parameters for the three analyzed metrics. The NCC similarity of all simulated and processed images with respect to the ground-truth on example of the central Z-slice $z^{(o)}$ are directly written into the panel images, see Figure 3.1b,e,h,o,q. While the dejittering algorithm provides the basis for a proper deconvolution it only increases NCC similarity by 2.5%. Subsequent deconvolution DEC achieves denoising by creation of a target estimate

	$NCC^{(\max)}$	$N_{Iter}^{(\max)}$	$\gamma^{(\max)}$	$NCC^{(\min)}$	$N_{Iter}^{(\min)}$	$\gamma^{(\min)}$
TV	8.82e-01	1.80e+02	3.16e-03	1.47e-01	2.00e+02	1.00e+00
GR	8.67e-01	2.00e+02	1.00e-02	1.85e-01	1.80e+02	3.16e-06
GS	8.67e-01	2.00e+02	1.00e-03	1.65e-01	8.00e+01	1.00e+00

TABLE 3.1: **Extrema of 2D-Search Parameter Search.** Iteration number $N^{(\text{iter})}$, regularization parameter $\gamma^{(\text{TV})}$ for their respective minumum and maximum NCC w. r. t. to parameter search Figure 3.5 are listed.

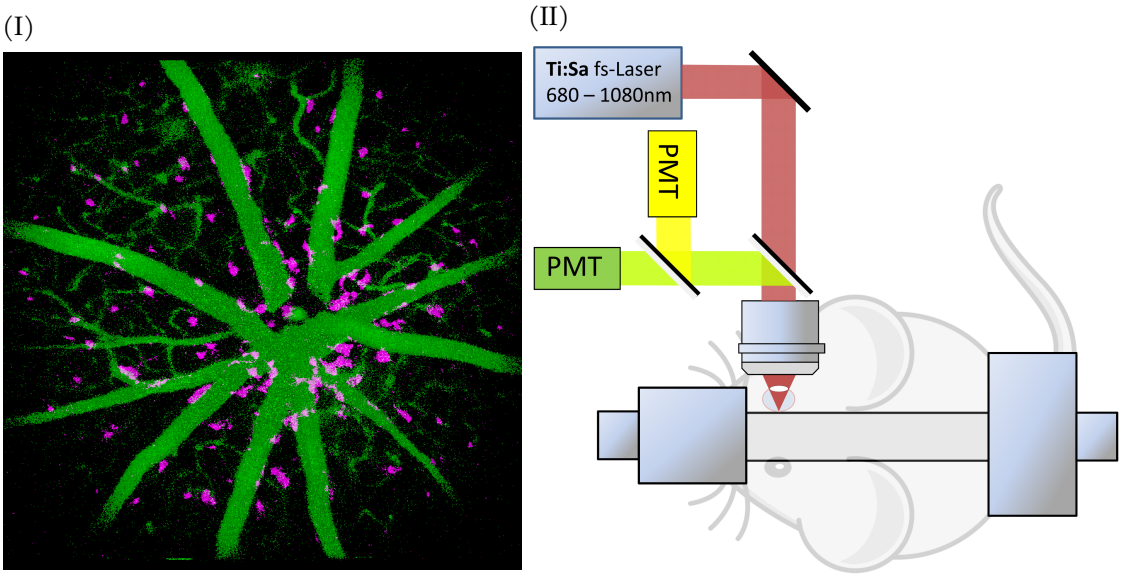


FIGURE 3.6: **Mouse Experiment Setup.** (I) Basic sketch of the imaging setup. The image shows a FoV of 1.5 mm ×1.5 mm. (II) Overlay of iso-rendering of completely processed channels, blood-vessels in green, microglia in magenta.

$\hat{S}$ while emphasizing smooth surfaces and clear edges with the chosen TV prior thereby increasing NCC similarity further by 13.4% and hence reconstructing an image of high quality (see Section A.11).

3.4 CORRECTION OF *IN-VIVO* MOUSE RETINA DATA

Using 2-photon fluorescence-based mouse retinal imaging, the proposed algorithm, already tested on *in-silico* data, is evaluated. The goal is to reconstruct a more reliable image w. r. t. to shape and position of mouse-eye invading microglia, see magenta structures in the reconstructed two color overlay result Figure 3.6I. Therefore, two color channels are employed, one for measuring large supportive and axially slowly changing structures and one for the microglia, which are sparsely distributed small structures.

The data were recorded by Daniel Bremer at the German Rheumatism Research Center (DRFZ) in Berlin on a 2-photon microscope setup built by Ruth Leben [61]. The setup is shown in Figure 3.6II. A Ti:Sa femtosecond laser is used to excite at $\lambda^{(\mathrm{ex})} = 900\,\mathrm{nm}$ and a set of two Photo-Multiplying Tube (PMT) detectors is used to measure in the regions $\lambda = [525 \pm 25, 593 \pm 20]\,\mathrm{nm}$. The scan unit, scan lens as well as the objective are project

specific developments of LaVision Biotec. The objective has 4× magnification at $0.28NA$ with water gel immersion and 6 mm working distance. The objective is well matched to the aperture of the $0.32NA$ mouse eye opened with dilated pupils. By using water immersion, a significant portion of aberrations generated by the mouse eye could be reduced. The theoretically correct Nyquist sampling is $d^{(\mathrm{Ny,lat})} \approx 400\,\mathrm{nm}$ and $d^{(\mathrm{Ny,ax})} \approx 7.5\,\mu\mathrm{m}$. To fit a sufficient Region of Interest (ROI) of the retina into the limiting FoV of $994\,pix \times 994\,pix$ the measurement was sampled at lateral pixel pitch $d^{(\mathrm{SPL,lat})} = 1.5\,\mu\mathrm{m}$, i. e. 3.8× the lateral Nyquist limit, thereby achieving a lateral FoV of 1.5 mm×1.5 mm. Axial sampling had to be reduced to $d^{(\mathrm{SPL,ax})} = 30\,\mu\mathrm{m}$, i. e. 4× of the axial Nyquist limit, due to technical limitations to ensure a sufficiently high scanning speed. The pixel dwell time was $1.6\,\mu s$. In the mouse eye, microglia were labeled with eGFP, measured at $\lambda^{(\mathrm{em1})} = 525\,\mathrm{nm}$, and blood pathways were labeled with sulforhodamine 101, measured at $\lambda^{(\mathrm{em2})} = 593\,\mathrm{nm}$. During measurements, the mouse was fixed via a rotatable and tiltable mount and anesthetized by isoflurane inhalation. The degree of anesthesia could be regulated with the absolute amount of inhalate. A large amount of inhalation leads to a strong anesthesia which noticeably reduces the amount of jitter. In this condition, however, irregular, strong changes in the angle of view of the mouse eye (spontaneous muscle contractions) occurred which led to a spontaneous shift of the FoV about the size of the FoV. If the inhalation intensity is too low, the mouse is sufficiently awake such that spontaneous muscle contractions did not occur and the FoV remains the same over the entire measurement period. In this case, however, the pulse as well as the respiration of the mouse is significantly higher and thus more jitter is present in the image.

The presented dejitter and processing algorithm is based on finding a reference and aligning the stack to it. Accordingly, dominant structures in the data are necessary to use them as dejitter scaffolds. In the case of the mouse eye, the large blood vessels can serve as basic structures. The algorithm is first applied entirely to the yellow ($\lambda^{(\mathrm{em2})} = 593\,\mathrm{nm}$) channel data stack, i. e. blood vessels. The processing of the raw data is shown using the pink-labeled ROI in Figure 3.7I on example of three different Z-slices $z^{(-5)}$, $z^{(0)}$ and $z^{(3)}$. The raw data (Figure 3.7IIa-c) are first drift-corrected (Figure 3.7IId-f), then dejittered (Figure 3.7IIg-i), and finally deconvolved (Figure 3.7IIj-l). The reference was calculated by median selection after SGF application to the data-set. For dejittering, a Gaussian window with $\sigma^{(\mathrm{G})} = 80\,pix$, dilation $D = 80\,pix$, no SGF prefiltering and the proposed CoM approach was used. For deconvolution, DEC with a GR prior with $N^{(\mathrm{iter})} \leq 200$

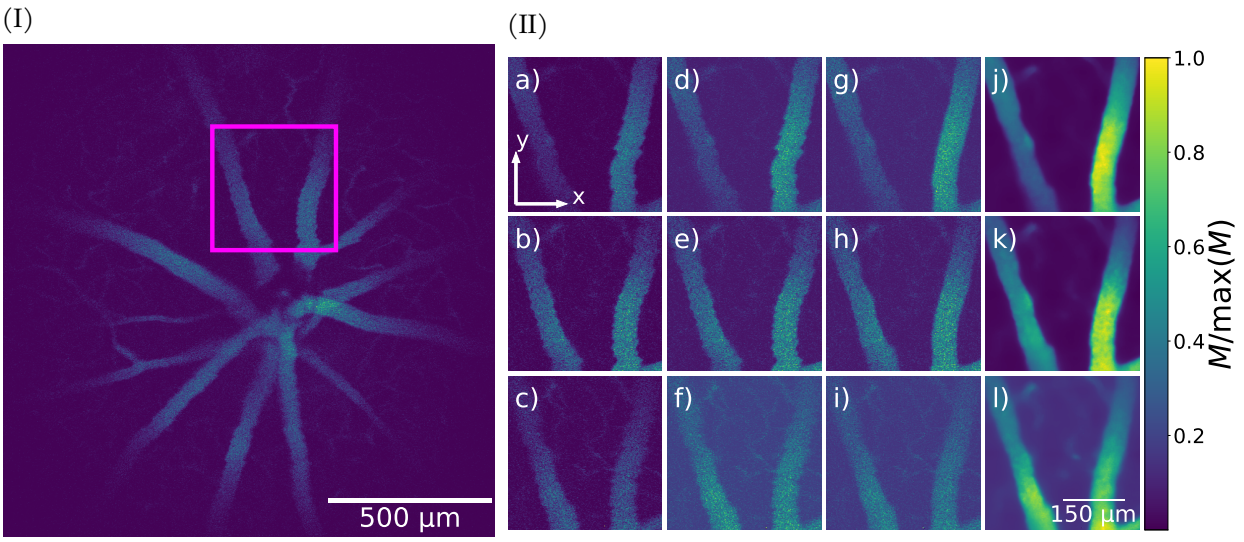


FIGURE 3.7: **Processing Retina Stack.** (I) raw data of Channel 1 (blood vessel) at $z^{(o)}$ with magenta marked ROI. (II) a-c) Raw-data of ROI shown at at Z-slices z_{-5}, z_0, z_3. The next columns always are processing results of the previous columns. d-f) axial-drift corrected, g-i) dejittered, j-l) DEC processed with GR prior.

and $\gamma^{(\mathrm{GR})} = 10^{-2.5}$ and DEC (Table 1.2) was used. The deconvolution parameters where directly transferred from the results of the parameter-search for the *spokes2d* target, see Figure 3.5. Note, that due to the coarse sampling deconvolution was implemented for denoising and contrast enhancement rather then resolution gain.

The measured raw mouse-retina Z-stack $M^{(\mathrm{mouse})}$ was minimum subtracted (up to an $\epsilon \in \mathbb{R}^+$), normalized by its standard deviation $\sigma^{(\mathrm{mouse})}$ and multiplied by the standard deviation $\sigma^{(\mathrm{spokes})}$ of the *spokes2d* target. Thus, the image was normalized according to:

$$M'^{(\mathrm{mouse})} = \frac{M^{(\mathrm{mouse})} - \min\left(M^{(\mathrm{mouse})}\right) + \epsilon}{\sigma^{(\mathrm{mouse})}} \sigma^{(\mathrm{spokes})} \tag{3.6}$$

which allowed to use the same parameters for DEC.

The effect of the drift correction can be visually perceived by comparison of the distance of the blood vessel structure to the reference-line, see Figure 3.7IIc+f). Lateral drift correction values are in the range of $s^{(\mathrm{drift})} \leq 8.6\,pix$, which is less than 1% of e. g. pixels along X-direction. Lateral jitter shifts were slightly greater $L^{(\mathrm{1D})}_{m'n'} \leq 20.1\,pix$, but were $1.34\,pix$ at the median and $1.85\,pix$ at the arithmetic mean.

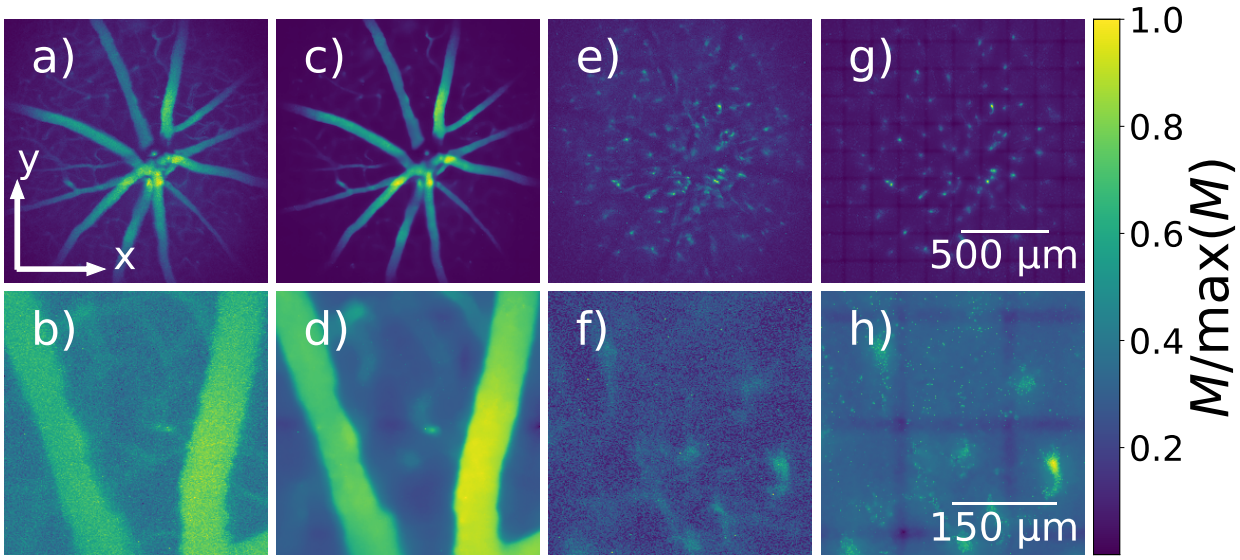


FIGURE 3.8: **Raw-Data and Deconvolution Result Comparison.** Max-Projection of (upper row) full-lateral FoV and (lower row) same ROI as Figure 3.7. Stacks: a-b) Raw and c-d) processed blood vessels. e-f) raw and g-h) processed microglia.

Shift values determined using blood vessels were applied directly to microglia. A color-overlay of a 3D iso-surface representation of the two processed channels is shown in Figure 3.6I where the viewing angle is set to 90 deg w. r. t. the first Z-slice (thus it looks slightly like a maximum projection). Here, the spatial distribution of microglia entering the intraocular space from the optic nerve (center of image) can be observed. In Figure 3.8 a max-projection of the raw and processed data-stacks is given. The upper row shows the full images before (a+c) and after processing (c+g). The lower row shows the ROI (see Figure 3.7I) of the upper row. The lower row is normalized according to $[M/\max M]^{0.3}$ to demonstrate the effects of the reconstructions especially in the background.

Comparing b) with d) demonstrates the effect of dejittering. Now, the blood-vessels have a more smooth surface which seems more reasonable and closer to the ground-truth. Thus, reconstructions based on the dejittered data can be trusted more like e. g. the shape of the lines in the background of d). The deconvolution further led to better distinguishability between fore- and background due to its denoising, optical sectioning and and surface leveling effect. Comparing f) with h) demonstrates clearly the gain in distinguishability between fore- and background. Particularly the microglia cell in the lower right is more defined. The grid pattern in the back of g+h) is a remnant of Hann-windowed recombination of the image after tiled processing with 20% overlap and needs to be improved further in future versions.

3.5 CONCLUSION AND OUTLOOK

Axial drifting, jitter and low-pass filtering can complicate the analysis and interpretation of measured image data and thus generate misinterpretations. Semantic data analysis using axial projections can produce double images due to drifting and jitter, which can lead to interpretation artifacts. In addition, the low-pass filtering inherent in imaging reduces the intensity of high spatial frequencies and thus the clear distinction between signal and background becomes more difficult the higher the frequencies.

To propose and demonstrate a suitable strategy to address these effects for a given real-world measurement problem, the *in-silico* object *spokes2d* was used as data-representitive for blood vessels and a forward simulation including the jitter process was first created. Suitable parameters for strength and type of lateral drift for each individual Z-slice as well as number, strength and abruptness of jitter along X-direction were presented and remedied by a proposed processing algorithm. Calculation of the reference image using 2D SGF filtering and median projection proved to be successful. For dejittering the combination of SGF prefiltering with line-wise CoM (along X-direction) calculation on a Gaussian-window for shift determination proved to be particularly suitable. For subsequent deconvolution, DEC together with priors TV, GR, and GS were tested, with TV being found to be a more suitable prior in the case of the *spokes2d* target due to its tendency to have sharp edges.

The proposed correction algorithm was finally applied to an *in-vivo* 2-photon fluorescence CLSM two-color mouse retina Z-stack. The analysis and correction were first performed for the blood vessels and thus lateral drifts were determined. By normalizing the Z-stack via minimum-subtraction and standard deviation based rescaling to the standard deviation of the *in-silico* data set, the deconvolution parameters from the simulation could be adopted. All determined shifts and used parameters of the blood vessels could be applied directly to the microglia Z-stack, allowing them to be processed appropriately, despite very low signal levels.

The advantage of dejittering during post-processing is difficult to see when comparing f) with h) due to the very detached and noisy structures, but can be estimated indirectly via the change from b) to d). In such a reference-less highly noisy case as f) greater confidence in the dataset can be established by the presented regularization.

For future developments making the toolbox more robust against axial jumps, including stretches or jitter along the Y-axis (i. e. slow scan direction) and improving the flow to work completely on GPU (not only the deconvolution) would be adviseable.

4

THICK SLICE UNMIXING (TU)

Children of infinity, always gazing to our past. We are dust of the stars.

— Ville Friman

4.1 MOTIVATION

In Image Scanning Microscopy (ISM), multiple viewing angles from the same object are recorded in parallel during the scanning process using a detector array (Section 1.4), each detector pixel yielding a seperate view of the sample. By lacking a detection-side pinhole, the photon loss occurring in the detection process can be reduced. Detector-side spatially resolved photon detection enables further processing capabilities.

In this section, the fundamental property of ISM as compared to the Confocal Laser Scanning Microscopy (CLSM) scheme of having more information about the investigated sample available for further analysis processes by using a pixelated camera instead of a single-pixel detector will be exploited. Existing widely used ISM processing methods such as Pixel Reassignment (PiRe) (Equation 1.26) process the available data linearly, allowing higher lateral resolution to be achieved. Nevertheless, the lack of optical sectioning cannot be remedied. In this section, a novel linear approach named Thick slice Unmixing (TU) will be presented to not only allow for optical sectioning by post-processing, but to even reconstruct a 3D sample-distribution from 2D multi-view measurements. Thus, 1) the basic idea of ISM will be explained further with a simulation and afterwards from a slightly changed forward-model the TU will be derived. 2) optimal multi-view DEConvolutions (mDECs) and TU parameters will be determined. 3) properties of TU w. r. t. Out of Focus (ooF) rejection and 3D reconstruction will be analyzed. 4) The processing strategy LEAve-out z-Planes (zLEAP) will be presented.

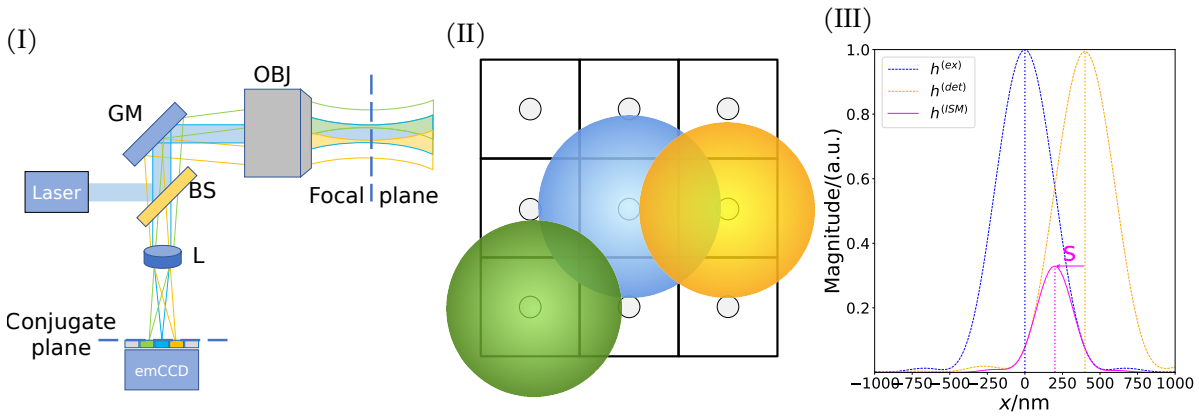


FIGURE 4.1: **Basic ISM-Setup.** (I) ISM schematic. Excitation beam $h^{(\mathrm{ex})}$ is colored blue and emission $h^{(\mathrm{det})}$ from two different lateral positions are colored orange and green. (II) Uni-rectangular detector shape of 3×3 pixels. Center of $h^{(\mathrm{ex})}$ is aligned with central pixel of detector. $h^{(\mathrm{det})}$ from (I) are displayed accordingly. (III) X-cut through $y = y^{(\mathrm{o})}$ and $z = z^{(\mathrm{o})}$ for 3D Simulation of the ISM Point Spread Function (PSF) according to (II). The shift is marked with s.

4.2 THEORY AND METHODS OF TU

4.2.1 *Derivation of TU*

The schematic structure of a scanning ISM is shown in Figure 4.1I. The illumination emitted by the laser is reflected at a Beam Splitting Element (BS), scanned by a Galvanometric Mirror pair (GM) positioned at a plane conjugate to the Back Focal Plane (BFP) of the objective lens, and the illumination is thus sequentially focused at different lateral scan positions. In Figure 4.1I one scan position is demonstrated and the the excitation light is colored blue. Some of the fluorescence emitted from the excited volume is collected back by the objective lens, descanned by the GM, propagates through the BS, and is focused on the camera (here electron-multipliying Charge-Coupled Device (emCCD)) using a focusing lens (L). Note that a tube lens and a scanning lens between GM and objective lens were omitted for simplicity of the drawing. Figure 4.1II shows the descanned excitation PSF $h^{(\mathrm{ex})}$ on the central pixel (blue) and two exemplary detection PSFs $h^{(\mathrm{det})}$ (orange and green) in the case of a rectangular pixel distribution unit cell for one scan-position

[1]. Observe, that the corresponding sample volume of most probable detection for the two detection pixels are shown with orange and green in the focal plane of Figure 4.1I. Figure 4.1III shows the 1D slice at $z = z^{(o)}$ and $y = y^{(o)}$ of a 3D spatial simulation of the ISM PSF model (Section 1.4) for one scan position.

The effect of the PiRe shift factor (Equation 1.26) is demonstrated using the central pixel and its forth neighbor along the horizontal view-direction, for simplicity: X-direction. The selected parameters are: excitation wavelength $\lambda^{(ex)} = 563$ nm, emission wavelength $\lambda^{(em)} = 580$ nm, numerical aperture of the objective $NA = 0.7$, isotropic lateral sampling $d^{(SPL,lat)} = 20$ nm, axial sampling $d^{(SPL,ax)} = 60$ nm, and isotropic detector pixel pitch $d^{(Cpitch)} = 100$ nm in scan coordinates. Both $h^{(ex)}$ and $h^{(det)}$ are max-normalized for visualization. $h^{(ex)}$ is shown dashed in blue and its center coincides with the center of the central pinhole from Figure 4.1II, thus its maximum along X-direction is at $\hat{x}^{(ex)} = 0$ nm. The detection PSF $h^{(det)}$ of the selected pinhole is shown in dashed orange[2]. It has its maximum at $\hat{x}^{(det)} = 400$ nm. The pinholes are simulated as δ-peaks, giving $h^{(ISM)}$ directly as the product of excitation and detection PSF (Equation 1.24). $h^{(ISM)}$ has its maximum at almost $\hat{x}^{(ISM)} = 200$ nm, rounded to full pixels in the simulation, and is drawn with magenta. Being the product of $h^{(ex)}$ with $h^{(det)}$ it is much smaller than either of the excitation or detection PSF, since a molecule has to both be excited and detected to appear in the recorded data. The distance between the excitation and emission maxima is given by $\Delta x = \hat{x}^{(det)} - \hat{x}^{(ex)}$.

The most likely position of the detected fluorophore emission in this example is at $\hat{x}^{(ISM)}$, but is stored at position $\hat{x}^{(ex)} + \hat{x}^{(det)}$. To correct for this effect, the distance $s^{(ISM)} = \hat{x}^{(ISM)} - \hat{x}^{(ex)}$ is determined in case of the PiRe processing and the shift factor $m^{(ISM)} = 1 - s^{(ISM)}/\Delta x$ is calculated above (Equation 1.26)[3]. With $m^{(ISM)}$ at hand, all detector pixel positions can be shifted back to the most probable emission position[4] (in scan coordinates) and

1 Due to the setup geometry the illumination is first scanned by the GM and afterwards descanned. If the optical axis of excitation and detection beam-path match and are centered on the central camera pixel $a^{(o)}$, then $a^{(o)}$ is always the conjugated position w. r. t. to the scan position of the illumination beam.

2 The detection probability of all (ideal) camera pixels is equally high and its sum over all pixels is normalized to 1.

3 For calculation of $m^{(ISM)}$ the distance between e. g. the central and neighboring pixels in scan coordinates must be known.

4 The extension of the calculation of $m^{(ISM)}$ in the case of non-isotropic detector pixel unit cells is straightforward.

afterwards summed up. In this thesis, this shift-back and sum-up processing is called Pixel Reassignment (PiRe). In the given example, the shift distance[5] is $s = m^{(\mathrm{ISM})} \cdot \hat{x}^{(\mathrm{det})}$ and so $s = 0.5 \cdot 400\ \mathrm{nm} = 200\ \mathrm{nm}$. In the case of ISM, an image consists of many scan positions where for each one the above explanation applies.

Since for PiRe all camera pixels are shifted back toward the central pixel $a^{(0)}$ and the total PSF of a laterally shifted detector pixel is asymmetric, it is important to accurately determine the central camera pixel $a^{(0)}$, otherwise asymmetric ISM PSFs may result and thus the reconstruction may suffer qualitative degradation. A brief simulation of the change of the lateral shape of the in-focus ISM PSF as a function of the camera pixel used is shown in Section D.1.

For determination of the lateral shift factor, mainly 3 different methods have been implemented in this work. Method 1) takes a (theoretically) calculated shift factor and applies it to all detector pixels according to given unit vectors of the camera pixel unit cell in the detector plane. This method is robust to any noise, but requires precise knowledge of the detector geometry and an accurate shift factor. Method 2) is called *nearest* and determines the respective distance of the two closest neighboring detector pinholes along the individual basis vectors of the detector unit cell to the central detector pinhole $a^{(0)}$ by determining the maximum of the correlation of these neighboring pinholes with the central detector pinhole. The unit vectors determined in this way are then applied to all remaining detector pinholes. Method 3) is called *complete* and also uses the determination of the maximum of the correlation of different pinholes to the central pinhole, but calculates it for all detector pinholes individually. While *nearest* is statistically less susceptible to noise compared to *complete*, *complete* does not require any further assumptions about the detector geometry. Method 1 requires less system resources by means of Central Processing Unit (CPU) and Random-Access Memory (RAM) than *nearest*, which again requires less than *complete*. For the use cases in this work, especially due to the independence from the theoretical shift factor determination, *nearest* was used.

In Section 1.7 different schemes for reconstructing an object estimate $\hat{S}$ were presented. While direct methods like 1/OTF (Section A.18) or Wiener Deconvolution (WD) (Section A.19) are directly applicable, iterative solutions like Poisson-likelihood maximizing single-view DEConvolution (DEC) using Limited-memory Broyden-Fletcher-Goldfarb-

5 Due to differences in the lateral extent between $h^{(\mathrm{ex})}$ and $h^{(\mathrm{det})}$ normally $m^{(\mathrm{ISM})} \neq 0.5$, but here only a pixel-exact simulation is demonstrated leading to $m^{(\mathrm{ISM})} = 0.5$.

Shanno algorithm (L-BFGS-B) (Section 1.7) naturally require significantly more processing time and ressources. The methods above have been presented only for single-view datasets. For datasets with multiple views, Signal-to-Noise Ratio (SNR)-optimal recombination in case of Weighted Averaging in Fourier Space (WAVG) (Section 1.7) has already been discussed. The aforementioned iterative DEC routine can be extended to multi-view data by extending the forward model with an indvidual PSF for each view. In contrast, the WD directly sums up all views before deconvolution due to the linearity of WD.

For the imaging process a 3D forward model with a 2D in-focus selection (i. e. axial position of sample w. r. t. the nominal focus position of the objective lens) at $z = z^{(0)}$ (Equation A.58) is assumed. The goal is to reconstruct an estimate $\hat{S}$ of a 3D sample distribution from a 2D multi-view ISM single Z-slice measurement of the object. Using the Fourier slice theorem [15], the discretized forward model (Equation 1.42) for measurement of an in-focus slice $z = z^{(0)}$ can also be expressed as:

$$M_{alm} = \mathcal{F}^{-1}_{(2\mathcal{D})} \left\{ \sum_{\tilde{n}} \left[\tilde{h}_{a\tilde{l}\tilde{m}\tilde{n}} \tilde{S}_{\tilde{l}\tilde{m}\tilde{n}} + \widetilde{\mathcal{N}}_{a\tilde{l}\tilde{m}\tilde{n}} \right] \right\} = \mathcal{F}^{-1}_{(2\mathcal{D})} \left\{ \widetilde{M}_{a\tilde{l}\tilde{m}} \right\}_{lm} \tag{4.1}$$

where the summation along the axial spatial frequency axis, indexed by $\tilde{n}$, equals the selection of the slice $z = z^{(0)}$ in real space. $\tilde{l}, \tilde{m}$ are the indices of the discrete lateral Fourier spatial frequencies, l, m are the indices of the lateral real space positions and a is the index of the detector pixel represented as a flattened array to account for arbitrary (2D) detector geometries.[6] For better readability the lateral frequency indices will be dropped, thus e. g. $\widetilde{M}_{a\tilde{l}\tilde{m}} \longrightarrow \widetilde{M}_a$ and the equations have to be evaluated individually for each individual selection of $\tilde{l}$ and $\tilde{m}$.

While $\widetilde{M}_a$ in Equation 4.1 is a 2D spatial frequency distribution, $\tilde{S}_{\tilde{n}}$ represents the complete 3D object, again for all pairs $\tilde{l}$, $\tilde{m}$ individually. In Equation 1.41 the forward-model uses a point-wise multiplication of the Optical Transfer Function (OTF) $\tilde{h}$ with the Fourier spectrum $\tilde{S}$ of the object in case of a complete 3D dataset or integration along the $\tilde{n}$ axial frequency if a 2D Z-slice is measured. In Equation 4.1 on the other hand a matrix product between $\tilde{h}$ and $\tilde{S}$, i. e. a summation along $\tilde{n}$, is assumed for the forward model.

6 Note that not for every l, m there exists an assigned $\tilde{l}, \tilde{m}$, but rather the indices are meant to visualize the axes they belong to, so in case of m or $\tilde{m}$ the Y- or $\tilde{m}$ axis respectively.

The goal is to determine an inversion-matrix $\tilde{h}^{\dagger}$ to reconstruct an estimate of the sample distribution $\hat{\tilde{S}}$ via inversion of the forward-model Equation 4.1, analog to the approach demonstrated in Section A.18, according to:

$$\hat{\tilde{S}}_{\tilde{n}} = \sum_{a} \tilde{h}^{\dagger}_{a\tilde{n}} \widetilde{M}_{a} \tag{4.2}$$

In general, the problem is ill-posed due to the band limitation of $\tilde{h}$ along $k^{(z)}$, the noise $\widetilde{\mathcal{N}}$ and the shape of the matrices, but can become well-posed and thus solvable by introducing a regularization. Assuming that the matrix $\tilde{h}$ has bounded support $supp\left(\tilde{h}\right)$, the reconstruction is performed only within this support, and $SNR\left(supp\left(\tilde{h}\right)\right) \gg 1$ in this domain, the noise term $\widetilde{\mathcal{N}}$ in Equation 4.1 can be neglected. Thus the regularized minimization problem can formulated as [62]:

$$\mathcal{L}_{a}\left(\hat{\tilde{S}}; \widetilde{M}_{a}\right) = \left\| \sum_{\tilde{n}} \tilde{h}_{a\tilde{n}} \hat{\tilde{S}}_{\tilde{n}} - \widetilde{M}_{a} \right\|_{2}^{2} + \gamma^{(\mathrm{TU})} \left\| \hat{\tilde{S}} \right\|_{2}^{2} \tag{4.3}$$

with $\gamma^{(\mathrm{TU})}$ as the Tikhonov regularization parameter. With the choice $\gamma^{(\mathrm{TU})} > 0$, Equation 4.3 is well-posed [62]. The problem was made convex, thus a global minimum exists and a $\gamma^{(\mathrm{TU})}$ can be found which minimizes $\mathcal{L}_{a}$. The regularizing pseudo-inverse $\tilde{h}^{\dagger}$ that solves Equation 4.3 can be computed efficiently using Singular Value Decomposition (SVD) [63]. In the SVD formalism, the mapping matrix $\tilde{h}$ is computed as a linear combination of two unitary transformation matrices U,V (rotations on $\mathbb{R}$) and a diagonal forward matrix Ξ, which holds the Singular Values (SV)s, i. e. square roots of the eigenvalues of $\tilde{h}^{*}_{a\tilde{n}} \tilde{h}_{a\tilde{n}}$ [62], according to:

$$\tilde{h}^{\dagger}_{a\tilde{n}} = \sum_{b} U_{ab} \sum_{\tilde{p}} \left[\Xi_{b\tilde{p}} V^{*}_{\tilde{p}\tilde{n}} \right] \tag{4.4}$$

Together with Equation 4.3 the diagonal entries of the forward matrix Ξ are:

$$\Xi_{nn} = \frac{\xi_{n}}{|\xi_{n}|^{2} + \gamma^{(\mathrm{TU})}} \tag{4.5}$$

with ξ_{n} the SVs of Ξ, which are sorted in descending order, hence $\xi^{(\max)} = \xi^{(0)}$. Using Equation 4.5 it is easy to see that with the choice of the regularization paramater $\gamma^{(\mathrm{TU})}$ a lower bound to the effectively influencing SVs can be set. Thus $\gamma^{(\mathrm{TU})}$ directly influences the conditioning number Θ (Equation 4.6) of the problem, which can be upper bounded by $|\xi^{(\min)}| \ll \sqrt{\gamma^{(\mathrm{TU})}}$, since mostly $0 < \gamma^{(\mathrm{TU})} \ll 1$ holds.

$$\Theta(\tilde{h}^{\dagger}) = \frac{|\xi^{(\max)}|}{|\xi^{(\min)}|} > \frac{|\xi^{(\max)}|}{\sqrt{\gamma^{(\mathrm{TU})}}} \tag{4.6}$$

In this thesis, reconstructions whose regularization parameter $\gamma^{(\mathrm{TU})}$ lies on the order of $O\left(1/\Theta\left(\tilde{h}\right)\right)$ (Equation 4.6) are considered to be unregularized. In this case, only numerical instabilities as well as inaccuracies in the determination of the OTF support are corrected. Parameter choices of $\gamma^{(\mathrm{TU})} \gg 1/\Theta\left(\tilde{h}\right)$ are called regularized. For a brief discussion on the rank of the OTF matrix see Section D.2.

By means of $\epsilon^{(\mathrm{mask})}$ the relative minimum value w. r. t. $\left|\tilde{h}\left(\vec{k}=0\right)\right|$ frequency of the OTF and thus the support cutoff spatial frequency $k^{(\mathrm{C})}$ to be used for the reconstructions is set. While $N^{(\xi\,\mathrm{lim})}$ limits the absolute maximum number of SVs (Equation 4.5) to be used for inversion, $\Delta\xi^{(\mathrm{rel})}$ gives an alternative to allow for all SVs ξ which satisfy the condition $\xi \geq \Delta\xi^{(\mathrm{rel})} \cdot \xi^{(\mathrm{o})}$. Since the pseudo-inverse is computed according to Equation 4.4 for all lateral spatial frequencies (indexed by $\tilde{l}, \tilde{m}$) individually, the respective magnitudes of the individual SVs ξ_n as well as the size of the matrices $\tilde{h}_{a\tilde{n}}$ vary, since a different amount of Z-slices lies within the OTF support $supp\left(\tilde{h}\right)$ at different lateral frequencies.

The reconstruction is called Thick slice Unmixing (TU), because $\hat{\tilde{S}}$ is reconstructed with an axial (3D) extent from a 2D Z-slice.

4.2.2 *Deconvolution Parameter Search*

For comparison of the TU results with according deconvolution results using mDEC and multi-view DEConvolution with 2D to 3D reconstruction (2d3dDEC) in Section 4.2.3, first the simulation parameters and the effects of the different reconstruction schemes on the parameter choices will be demonstrated in this section.

The effective ISM PSF is calculated according to Equation 1.24 with $NA = 1.42$, $n^{(\mathrm{im})} = 1.52$, $\lambda^{(\mathrm{ex})} = 561\,nm$, and $\lambda^{(\mathrm{em})} = 580\,nm$. Parameters are chosen to roughly match the experimental parameters in Section 4.3. Hence the necessary Nyquist sampling upper bound is $d^{(\mathrm{Ny,ax})} = 145\,nm$, $d^{(\mathrm{Ny,lat})} = 50\,nm$ and one lateral airy unit of the detection PSF has the size of $d^{(\mathrm{1AU,lat})} = 498\,nm$. Theoretically, it is sufficient to choose the sampling distances for the scan matching the Nyquist limit of the scanning PSF $h^{(\mathrm{ex})}$ and for the detection matching the Nyquist limit of $h^{(\mathrm{det})}$. In this simulation, however, the much more restrictive Nyquist limit of the effective PSF is chosen and thus the sampling used is $d^{(\mathrm{SPL,lat})} = 42.17\,\mathrm{nm}$ and $d^{(\mathrm{SPL,ax})} = 100\,\mathrm{nm}$. The camera-pixels have a rectangular unit-cell and the pitch in sample coordinates is isotropic $d^{(\mathrm{Cpitch})} = 84\,nm$ or roughly $0.17 \times d^{(\mathrm{1AU,det})}$

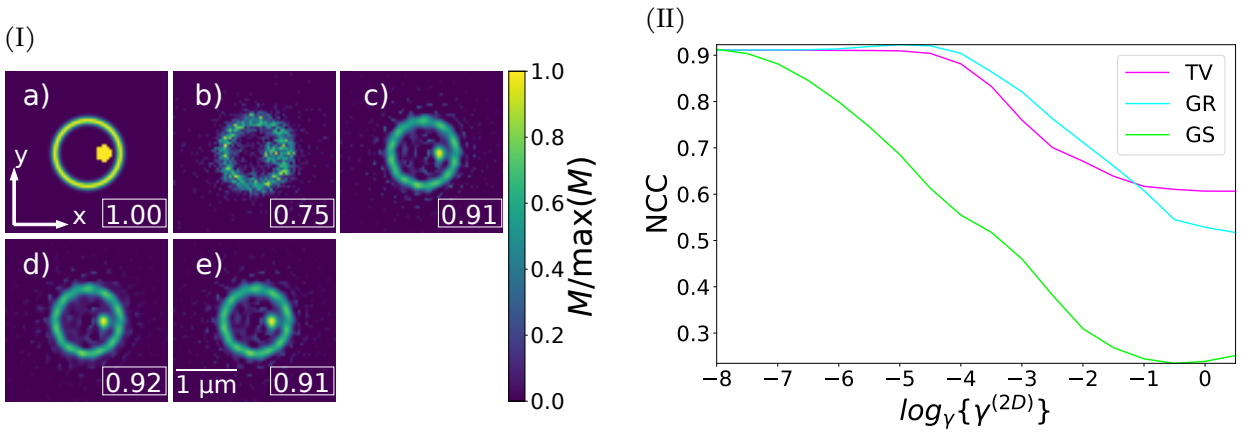


FIGURE 4.2: **2D Deconvolution Parameter Search** (I) for a) $z^{(o)}$ slice of *obj3d* displayed for the b) central detector-pinhole $a^{(o)}$. Results for optimal parameters in case of DEC with priors c) TV, d) GR and e) GS are displayed. NCC similarity is displayed in the lower right corner of each tile. (II) NCC results for whole tested parameter range and regularizers used in (I).

of $1\,AU$ of the detection PSF $h^{(\mathrm{det})}$. Of all camera pixels, an active Region of Interest (ROI) of $5\,pix \times 5\,pix$ or $0.84\,AU \times 0.84\,AU$ of the detection PSF $h^{(\mathrm{det})}$ around the center pixel, defined by the center of the excitation PSF $h^{(\mathrm{ex})}$, is used. Even though a bigger active on-camera pixel-area would allow to reconstruct even deeper Z-slices of the sample[7] and thus a coverage of a lateral region of $1.5 - 2\,AU$ would be preferable, the parameters are chosen to demonstrate the influence of the active on-camera pixel-area on the reconstruction. The object *obj3d* (Figure 2.1 m-p) is scanned three dimensionally leading to a 4D data set of size $N = [N^{(\mathrm{a})}, N^{(\mathrm{x})}, N^{(\mathrm{y})}, N^{(\mathrm{z})}] = [25, 64, 64, 32]$. Furthermore, Poisson noise is applied to the normalized image such that a maximum pixel has $N^{(\mathrm{phot})} = 10$.

Effects of finding the optimal iteration number $N^{(\mathrm{iter})}$ have already been presented in Section 3.3, so in this section the focus is placed on finding the ideal regularization parameters γ for the used regularizers Total Variation (TV), Good's Roughness (GR) and Gradient Square (GS), see Section 1.7. The Normalized Cross Correlation (NCC) Equation 1.39 is again used as a measure for reconstruction quality. All detector pinholes $N^{(\mathrm{a})}$ are utilized.

Reconstructions from a 2D (e. g. $z = z^{(o)}$) to a full 3D stack are particularly relevant for comparability to 3D reconstructions using the TU algorithm (Section 4.2.1).

7 Because even more information can be acquired from the more displaced views.

In order to use comparable regularization parameters for all further reconstructions, all images M used for reconstruction are reduced to the pixel sum of the reference stack $\sum_{almn}^{N} M^{(\mathrm{ref})}$ as well as normalized by the number of Z-slices to be reconstructed $N^{(\mathrm{z,recon})}$. For the 2D mDEC reconstruction only the $z^{(0)}$-slice of the simulated multi-view Z-stack is used. The results for this case are shown in Figure 4.2I, with all images normalized to their individual maximum. Here, the lateral cuts at $z^{(0)}$ of a) the ground-truth S, b) the noisy non-aberrated image M_0 and the 2D mDEC reconstructions in case of c) TV, d) GR and e) GS are shown. The regularization factors γ are found via an NCC-based maximum search, shown in Figure 4.2II. The NCC maximum for TV and GR precedes the plateau ($\gamma^{(2\mathrm{D})} \in [10^{-8}, \dots, 10^{-5}]$) but differs only slightly from it. Thus, in most cases of this reconstruction example, regularization is hardly advantageous to reconstruction. The optimal $\gamma^{(2\mathrm{D})}$s found are $\vec{\gamma}^{(2\mathrm{D})} = [\gamma^{(\mathrm{TV})}, \gamma^{(\mathrm{GR})}, \gamma^{(\mathrm{GS})}] = [10^{-8}, 10^{-5}, 10^{-8}]$ with which the NCC maxima $\vec{C}^{(2\mathrm{D})} = [C^{(\mathrm{TV})}, C^{(\mathrm{GR})}, C^{(\mathrm{GS})}] = [0.91, 0.92, 0.91]$ and thus a high-quality reconstruction is achieved. The maximum number of iterations was set to $N^{(\mathrm{iter})} = 200$ for all deconvolutions and convergence was reached within this bound.

The same analysis procedure for 2d3dDEC as well as 3D mDEC reconstruction is shown in Figure 4.3. In Figure 4.3I the upper row shows a lateral-cut at the $z = z^{(0)}$ and the lower row an axial cut at $y = y^{(0)}$. a+b) shows the ground-truth and c+d) the noisy but aberration-free image of the central detector pixel M_0. In e+f,i+j,m+n) the results of 2d3dDEC reconstruction, i. e. reconstruction of a 3D-sample distribution from only the in-focus $z = z^{(0)}$ slice, for all three regularizers are displayed. g+h,k+l,o+p) display the results of 3D mDEC using the complete Z-stack M. In case of 2d3dDEC, Figure 4.3II-4.3III and for 3D, mDEC Figure 4.3IV the NCC similarity of the reconstructed thick slice to the ground truth as a function of the regularization parameter γ is displayed. While Figure 4.3II uses the entire Z-stack for NCC calculation, Figure 4.3III uses a reduced axial range of $\Delta z = \pm 4$ Z-slices.

Due to the mirror symmetry property of the PSF at $z^{(0)}$, the iterative 2d3dDEC reconstruction (Figure 4.3 f,j,n) is equally mirror symmetric with respect to $z^{(0)}$. The limited axial extent of the PSF further leads to that relevant information from more distant Z-slices (e. g. $z^{(16)}$) do not contribute to the measured slice, or only as a nearly constant background. Thus, the reconstructed stack seems limited to a small range $\Delta z = \pm 4$ around $z^{(0)}$ even though the reconstruction accounted for the full 32 Z-slices. The NCC results are given in case of 1) utilizing the full reconstructed Z-range (Figure 4.3II) and 2) only using

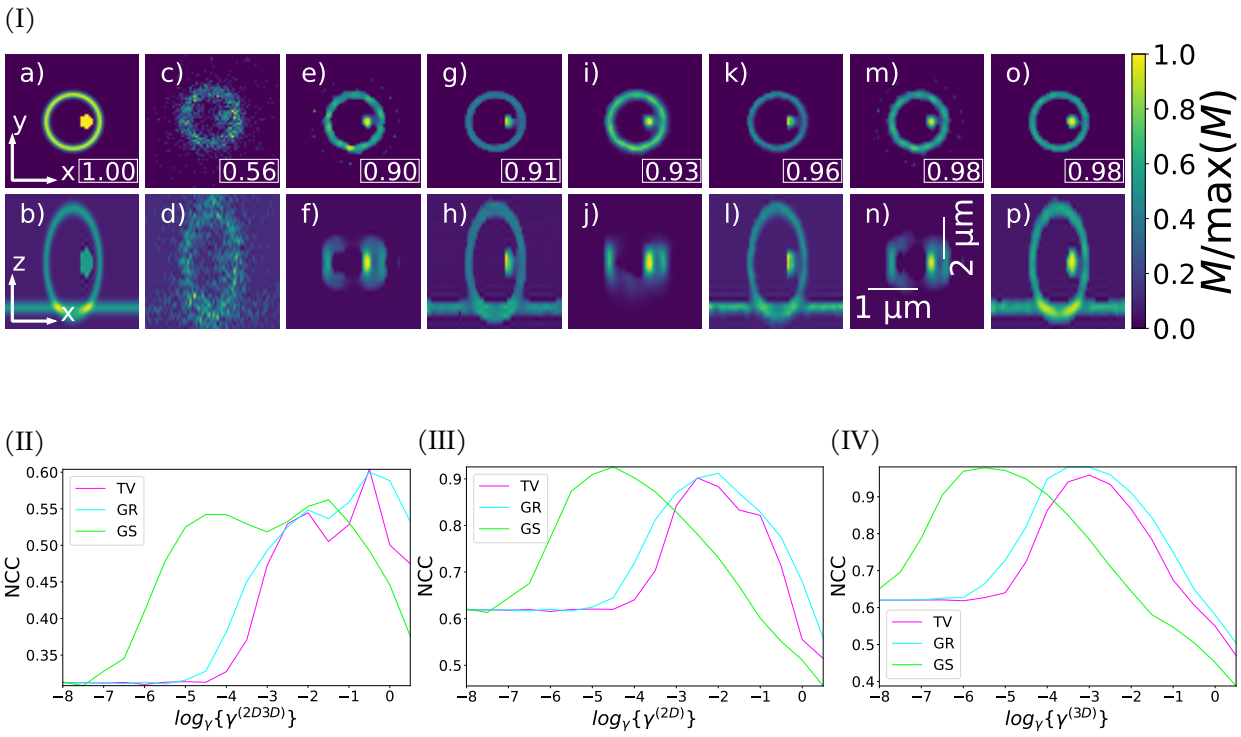


FIGURE 4.3: **3D Deconvolution Parameter Search** on an example of the *obj3d* target. (I) Upper row lateral slice (at $z^{(o)}$) lower row axial slice (at y_0). a+b) RAW object c+d) image of central pixel $a^{(o)}$. e-p) Results of mDEC with e-h) TV, i-l) GR and m-p) GS prior where left column (eg e+f) of each 4-image block (eg e-h) uses 2d3dDEC reconstruction and right column (eg g-h) 3D mDEC reconstruction. 3D NCC similarity is displayed in lower right corner of upper row tiles. NCC results using complete Z-range in case of (II) 2d3dDEC and (IV) 3D mDEC search. (III) shows NCC results for comparison of 2d3dDEC with ground-truth within a limited Z-range of $\Delta z = \pm 4$ around central slice.

a limited Z-range of $\Delta z = \pm 4$ around $z^{(o)}$ for calculation of the NCC (Figure 4.3III). In case of 1), e. g. two local maxima of the NCC of TV are found with the bigger maximum at $\log_{10}\left(\gamma^{(2D3D,TV)}\right) = -0.5$ while in case of 2) one maximum at $\log_{10}\left(\gamma^{(2D3D,TV)}\right) = -2.5$ is found. Visual comparison of the two results shows that the found maximum of 1) leads to an over-regularized reconstruction where the the high axial regularization together with the approximate symmetry of the used *obj3d* leads to more pixels $\neq 0$ that can be positively correlated thereby yielding a comparatively higher NCC result even though

the resolution gain is not optimal.[8] Thus, for the NCC evaluation on the limited Z-range $\Delta z = \pm 4$ the parameters $\vec{\gamma}^{(\mathrm{2D3D})} = [10^{-2.5}, 10^{-2}, 10^{-4.5}]$ with $\vec{C}^{(\mathrm{2D3D})} = [0.90, 0.92, 0.93]$ are found. For pixel normalization, the factor $N^{(\mathrm{z,recon})} = N^{(\mathrm{z})} = 32$ was used.

The complete 3D reconstruction mDEC (g+h,k+l,o+p) can reliably recover the entire object and finds the maxima $\vec{\gamma}^{(\mathrm{3D})} = [10^{-3}, 10^{-3.5}, 10^{-5.5}]$ with $\vec{C}^{(\mathrm{2D3D})} = [0.96, 0.98, 0.98]$.

4.2.3 *Properties of TU*

The simulation parameters and 3D sample used in this section are the same as in Section 4.2.2. Here, a 3D thick slice sample estimate $\hat{S}_{lmn}$ from a measured Z-slice M_{amn} at $z^{(\mathrm{o})}$ is reconstructed. The simulated image dataset M_{amn} consists of $N^{(\mathrm{M})} = [25, 64, 64, 1]$ pixels[9], while the multi-view 3D PSF dataset used for reconstruction consists of $N^{(\mathrm{h})} = [25, 64, 64, 32]$ pixels.

A parameter search for TU, where a 3D thick slice object estimate $\hat{S}^{(\mathrm{TU})}$ is reconstructed from a 2D multi-view measurement M_{amn}, is shown in Figure 4.4. The graphic is meant to be read column-wise. The 1st and 3rd row show a lateral cut at $z^{(\mathrm{o})}$ while the 2nd and 4th show an axial cut at y_0. While in the upper two rows the reconstruction results with minimal NCC, within the given simulation range, in the lower two rows the results with maximum NCC are displayed. The parameters were varied in the range: $\epsilon^{(\mathrm{mask})} = 10^{-0.5-b}$, $\gamma^{(\mathrm{TU})} = 10^{-5-b}$, $N^{(\xi\,\mathrm{lim})} = 1 - 2b$ and $\Delta\xi^{(\mathrm{rel})} = 10^{-1-b}$ for basic steps $b \in [0, , 0.5, \ldots, 5]$. In the first four columns one of the four parameters relevant for the TU algorithm is always varied, while the others remain fixed at the basic settings $\epsilon^{(\mathrm{mask})} = 10^{-5}$, $\gamma^{(\mathrm{TU})} = 10^{-8.5}$, $N^{(\xi\,\mathrm{lim})} = 100$, and $\Delta\xi^{(\mathrm{rel})} = 0$, thus imposing no significant constraint on the SVs. The varied parameters per column are: a-d) $\epsilon^{(\mathrm{mask})}$, e-h) $\gamma^{(\mathrm{TU})}$, i-l) $N^{(\xi\,\mathrm{lim})}$ and m-p) $\Delta\xi^{(\mathrm{rel})}$. NCC was calculated for the axial thick slice of $[z^{(\text{-}3)}, z^{(\text{-}2)}, \ldots, z^{(3)}]$ to anylize how close the reconstruction is to the original 3D sample distribution around the central Z-slice. Nevertheless, Figure 4.4I displays the entire reconstructed Z-stack. For comparison, multi-view deconvolution in case of q-r) 2d3dDEC and 3D s-t) mDEC using GS regularization

8 Note, that a bigger active on-camera pixel-area could remedy this problem by allowing to reconstruct a bigger 3D volume.

9 Note, that the axis order is $N^{(\mathrm{M})} = [N^{(\mathrm{M,a})}, N^{(\mathrm{M,x})}, N^{(\mathrm{M,y})}, N^{(\mathrm{M,z})}]$.

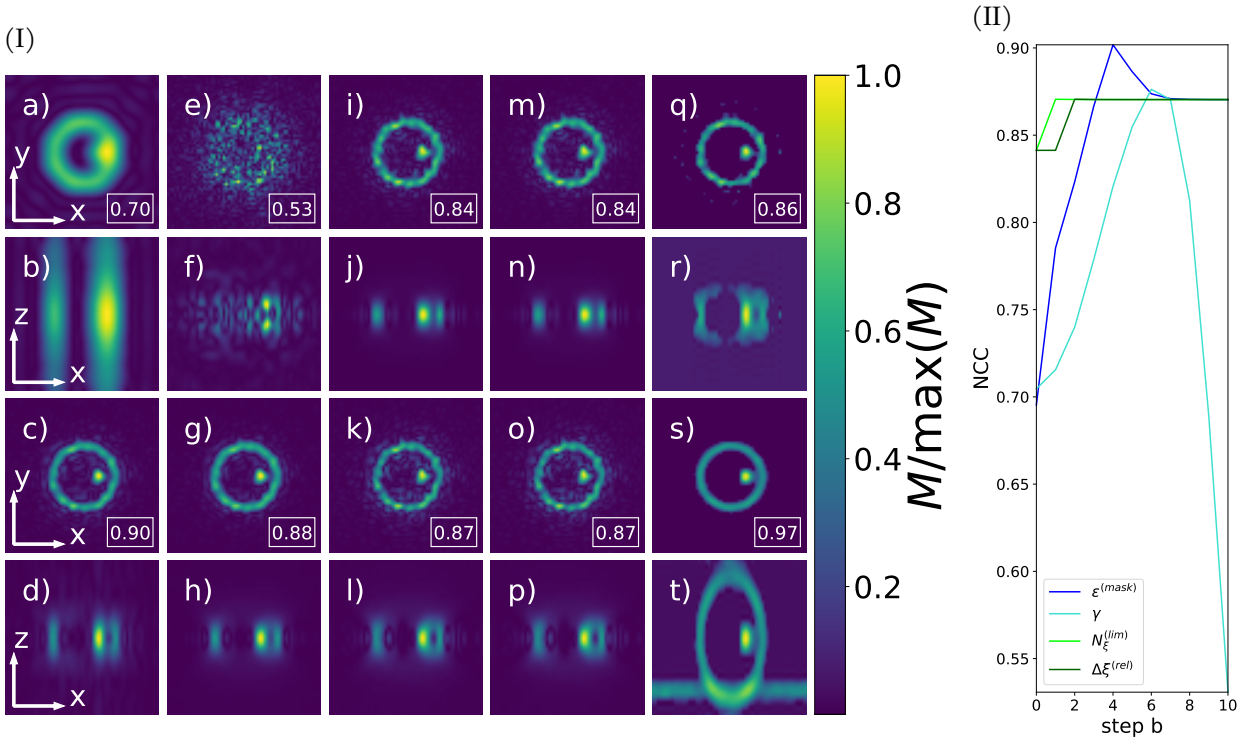


FIGURE 4.4: **Thickslice Parameter Search** on example of obj3d. (I) Results of found parameters yielding (first 4 columns of) row1+2) minimal NCC, row3+4) maximal NCC. Parameters varied are a-d) $\epsilon^{(\text{mask})}$, e-h) $\gamma^{(\text{TU})}$, i-l) $N^{(\xi\,\text{lim})}$ and m-p) $\Delta\xi^{(\text{rel})}$. Multi-view deconvolutions q+r) 2d3dDEC and s-t) 3D mDEC are depicted for visual comparison. NCC similarity is displayed in lower right corner of each tile. (II) NCC results of TU reconstruction for individually varied parameters $\epsilon^{(\text{mask})}$, $\gamma^{(\text{TU})}$, $N^{(\xi\,\text{lim})}$ and $\Delta\xi^{(\text{rel})}$. Individual parameter values are derived from steps $b \in [0,5]$ with step size $\Delta b = 0.5$.

with $\gamma^{(\text{2D3D})} = 10^{-5} = \gamma^{(\text{3D})}$ was calculated. Figure 4.4II shows a plot of the NCC results in case of TU reconstruction for the four varied parameters of Figure 4.4I as a function of the step parameter b.

Optimal parameters found for reconstruction of the sample *obj3d* are $N_\xi^{(\max)} = 25$, $\xi^{(\max)} = 9.5 \cdot 10^{-4}$, $\xi^{(\min)} = 1.12 \cdot 10^{-15}$, thus the conditioning number is $\Theta = 8.5 \cdot 10^{11}$ and the problem is regularized as $\gamma^{(\text{TU})} \gg 1/\Theta$. Here, $\xi^{(\max)}$ and $\xi^{(\min)}$ as well as the conditioning number Θ (Equation 4.6) are calculated for each lateral spatial frequency individually.

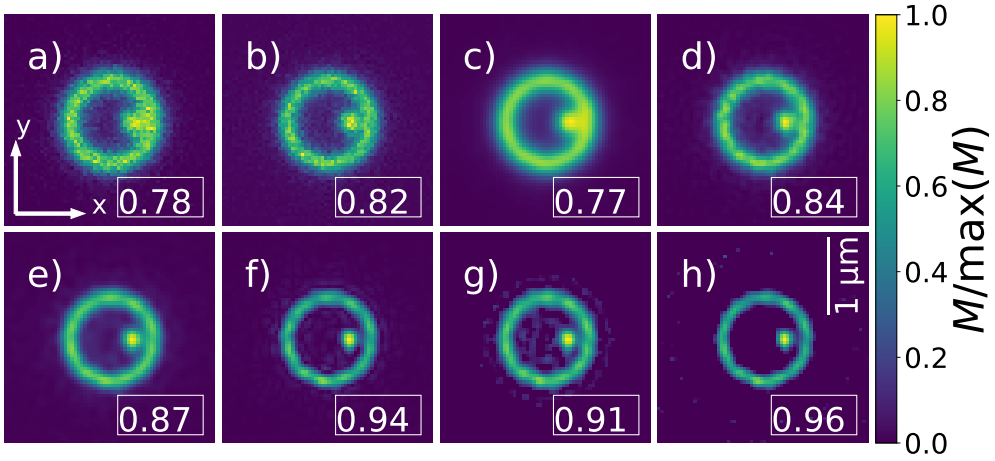


FIGURE 4.5: **ooF Rejection Comparison.** (I) comparison of ooF rejection and image reconstruction quality for different methods using all available views. Processing methods are: a) $M^{(\mathrm{sum})}$, b) $M^{(\mathrm{ISM})}$, c) $M^{(\mathrm{wavg})}$, d) $M^{(\mathrm{nwavg})}$, e) $M^{(\mathrm{wd})}$, f) $M^{(\mathrm{tu})}$, g) 2D $M^{(\mathrm{mDEC})}$, h) 3D $M^{(\mathrm{mDEC})}(z = z^{(\mathrm{o})})$ with selection of central slice. NCC similarity to ground-truth is displayed in lower right corner of each tile.

While $\epsilon^{(\mathrm{mask})}$ and $\gamma^{(\mathrm{TU})}$ have the largest impact on reliable unmixing, the relative impact due to the constraint on SVs is almost imperceptible. The reconstruction matrix is highly rank-reduced, thus SVs ($\equiv$ large Θ) drop rapidly and only the first two SVs (see $N^{(\xi\mathrm{lim})}$ Figure 4.4II) have a significant impact on reconstruction quality and stability (Figure 4.4II). Thus, the reconstruction is dominated by limiting OTF support and Tikhonov regularization.

The reconstructions are symmetric around the slice $M^{(\mathrm{o})}$ measured at $z^{(\mathrm{o})}$ and reconstruct the sample only around a limited range δz. The symmetry is because TU tries to distribute the measured frequency components of $\widetilde{M}$ (at $n^{(\mathrm{o})}$) of a given lateral position over all frequencies $k^{(\mathrm{z})}$ of $\hat{S}$ by clever weighting. Since h is axially symmetric this leads to the fact that the weights are also distributed uniformly and symmetrically around $n^{(\mathrm{o})}$. The axial limitation depends on the number of pinholes used, so that the inversion problem is well defined, the information content per pinhole (i. e. the angle of the view) and the SNR. Views further away from the central detector pixel allow the acquisition of more distant sample volumes and thus a more complete image of the investigated volume.

If only individual slices (or regions) are to be reconstructed rather than an entire stack, the inverse $\tilde{h}^{\dagger}_{a\tilde{n}}$ can be computed according to Equation 4.2 only for a specific $\tilde{n} = \tilde{n}^{(\mathrm{o})}$. Thus, the necessary numerical effort can already be reduced by the axial dimension

and a performance boost can be achieved. By using a full 3D forward model, but only reconstructing a single axial slice $z^{(0)}$, ooF are removed from the reconstructed slice and inherently regularized.

In Figure 4.5 a selection of reconstruction methods w. r. t. their ooF rejection capability is compared. The simulation parameters remain unchanged. The methods used are: a) direct summation of all views $M^{(\mathrm{sum})}$, b) PiRe $M^{(\mathrm{ISM})}$, c) WAVG $M^{(\mathrm{WAVG})}$, d) noise normalized Weighted Averaging in Fourier Space (nWAVG) $M^{(\mathrm{nWAVG})}$, e) WD $\hat{S}^{(\mathrm{WD})}$, f) TU $\hat{S}^{(\mathrm{TU})}$ with limitation to one central slice, g) 2D $\hat{S}^{(\mathrm{mDEC})}$, h) 3D $\hat{S}^{(\mathrm{mDEC})}$ with subsequent selection of $z^{(0)}$. The NCC similarity was determined only for the $z^{(0)}$ slice. While a-e,g) use 2D forward model with 2D multi-view dataset and PSF at $z^{(0)}$, f+h) use a 3D forward model. Here, f) uses only the central spatial frequency $k_0^{(z)}$ of the multi-view 3D OTF while 3D mDEC utilizes the full 3D PSF. The comparison shows that the deconvolutions (e-h) tend to achieve the NCC-better results, with the mDEC and TU processings in particular achieving reconstructions of high quality (see Section A.11).

For a more in-depth analysis in case of 21 different active detector-pixel-configurations, see Section D.3. By means of simulation it was shown that the detector pixel pitch has influence on the reconstructability of a thick slice. Within the same simulation parameters, an optimal pixel pitch of $d^{(\mathrm{Cpitch})} = 3 \cdot d^{(\mathrm{SPL,lat})} \approx 126$ nm could be obtained for the target *obj3d* for TU and of $d^{(\mathrm{Cpitch})} = d^{(\mathrm{SPL,lat})} \approx 42$ nm for 3D mDEC. For details, see Section D.4. In particular, it turns out that not every camera pixel has the same importance in the case of a 3D thick slice reconstruction. Should a limitation of the used detector elements be necessary for faster data readout, data volume limitation or cost limitation when building a dedicated detector array, it was found that especially the central pinhole is of essential importance. In addition, to achieve sufficient thick slice reconstruction quality in a sparse axial region around the measured slice $z^{(0)}$, the direct four neighbors (vertical and horizontal) are of particular importance. For more details see Section D.5.

4.2.4 *zLEAP*

In the previous sections, it was illustrated how a 3D thick slice can be reconstructed from a 2D slice using the available views of the ISM measurement. However, the axial thickness of the reconstructible thick slice is 1) limited to region Δz around the position of the measured

2D slice at $z^{(0)}$ due to the finite axial extent of the PSF. Further, due to the symmetry around the central slice $z^{(0)}$ of the PSF h, it is 2) indistinguishable to the algorithm in the case of an ideal (axially symmetric) PSF whether an image information has come from above or below the slice at $z^{(0)}$. One possibility to achieve the aforementioned algorithmic axial discrimination 2) is to use an aberrated PSF in order to create an axial asymmetry. Among the manifold possibilities, two methods should be mentioned here. First, manipulation of the excitation PSF by using a cylindrical lens has become standard in 3D single molecule localization microscopy (SMLM) microscopy [64] and can easily be adopted to other setup geometries like ISM. Second, manipulation of the detection PSF using phase mask e. g. to generate a single or double helix in REfocusing after SCanning using Helical phase engineering (RESCH, [12]). While the former method can be easily inserted into existing setups at low cost, the latter method is more costly and complex. Since phase masks can be used to generate arbitrary patterns, a extended 3D thick slice is reconstructed from a 2D slice using the example of slight spherical aberrations and strong astigmatism.

In order to 1) reconstruct a larger Δz area without changing the experimental setup and 2) achieve local axial asymmetry the demonstrated TU will now be extended by recording $N^{(\mathrm{p})}$ additional images $M^{(\mathrm{p})}$ at positions $z = z^{(\mathrm{p})}$. Thus, e. g. by measuring an image at every 3rd Z-position of an available 3D PSF of axial extent $N^{(\mathrm{z})} = 32$, if the measurement is centered around $z^{(0)}$, an image would be recorded at $\left\{z^{(-6)}, z^{(-3)}, z^{(0)}, z^{(+3)}, z^{(+6)}\right\}$ and afterwards this extended input Z-stack $M^{(3)}$ is used for reconstruction of a bigger connected region within the available axial extent of PSF.

For zLEAP processing, all recorded image Z-slices are interpreted as additional set of views of the same sample and thus are concatenated along the detector axis A, i. e. a new data set $M'_{a'lm}$ of the form $N = [N^{(a')}, N^{(\mathrm{x})}, N^{(\mathrm{y})}]$ with $N^{(a')} = N^{(a)} \cdot N^{(\mathrm{z})}$ is created. The same is done for the PSFs, thus a new effective multi-view PSF $h'_{a'lmn}$ is created of the form $N = [N^{(a')}, N^{(\mathrm{x})}, N^{(\mathrm{y})}, N^{(\mathrm{z})}]$. Finally, TU is used to reconstruct the estimate $\hat{S}^{(\mathrm{TU})}$ from the dataset M' and h' where the image information of the $M^{(\mathrm{p})}$ is now used to solve the symmetry problem of TU and additionally fill the regions in-between the $M^{(\mathrm{p})}$. This approach coined the name zLEAP to the presented method, as it jumps at measurement time over non-selected Z-slices, but reconstructs them at processing time. Note, that h' scales with $N^{(\mathrm{z})}$ for every new Z-slice that is used for the thick slice reconstruction zLEAP. For comparison of the reconstruction performance a similar approach in case of mDEC is implemented. Here, an empty data volume M'' of same size as the PSF h, i. e. of

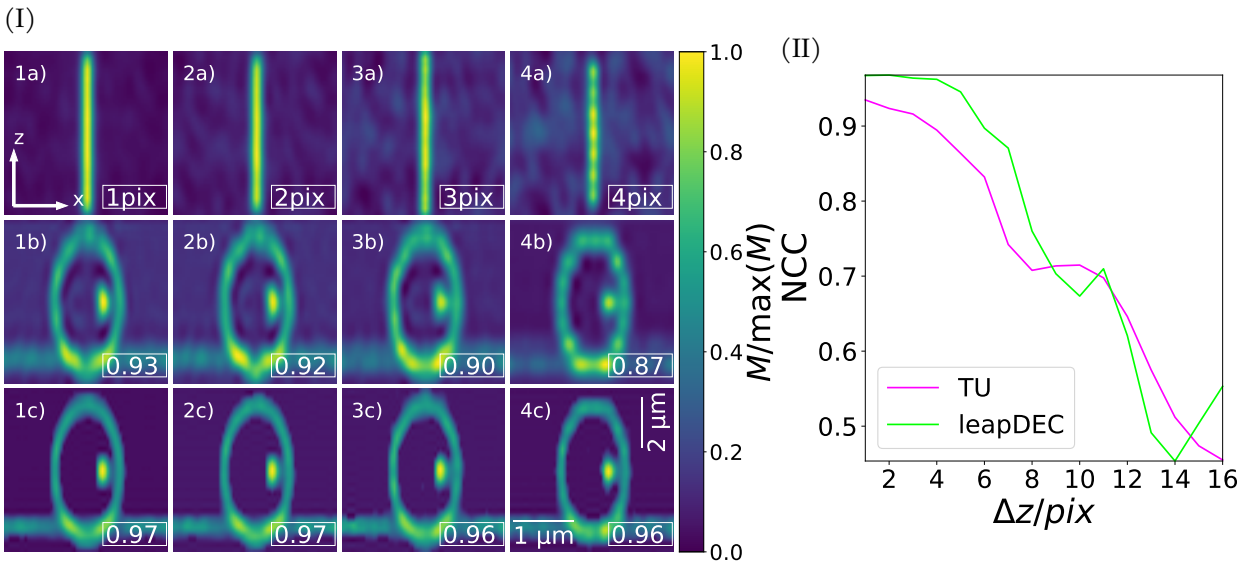


FIGURE 4.6: **zLEAP 3D-Reconstruction from Thickslice** for different axial spacings Δz. (I) 1st row: summed axial PSFs for central Pinhole, 2nd row: TU 3rd row: mDEC, all 3D-Volumes cut at $y = y^{(o)}$. The axial pitch Δz is written in lower right corner of upper row tiles. The NCC results are plotted in the lower right corner of tiles from the lower two rows. (II) Calculated NCC for all tested Δz pitches for TU (magenta) mDEC (lime) reconstruction.

size $N = [N^{(a)}, N^{(x)}, N^{(y)}, N^{(z)}]$, is created and filled with the measured images $M^{(p)}$ at $z = z^{(p)}$. Then, 3D mDEC with a chosen prior and PSF h reconstructs a full 3D thick slice estimate $\hat{S}^{(mDEC)}$. For calculation of the loss an additional mask is fed to the algorithm that limits the loss calculation to the Z-slices $z^{(p)}$ that are filled with image data $M^{(p)}$.

For the simulation the same parameters as before are used and thus the 3D object S and noisy image of the central view M_0 is given in Figure 4.3I. For the given target and simulation parameters the NCC improvement by using more then the central five views is negligeable and thus for this reconstruction only the central five views, i. e. central view and left-right-top-bottom direct neighbors, were used, see Section D.3 for further information. Thus, the stack size is $N^{(a)} = 5$, $N^{(x)} = N^{(y)} = 64$ and $N^{(z)} = 32$.

The dependence of the reconstruction quality of the whole 3D thick slice on the number and axial spacings of the PSFs used is shown in Figure 4.6. For the comparison of 3D mDEC to TU, an optimal set of mDEC as well as TU parameters was determined for the case $\Delta z = 1$ and used for all Δz spacings. The PSFs summarized in $N^{(a')}$ were sum-normalized

to one for each Δz and the corresponding image stacks were subsequently multiplied to the sum value of the full reference 3D Z-stack divided by the number of reconstructed Z-slices $N^{(z)}$. In Figure 4.6I each column has a different number of Z-slices Δz from one axial measurement $z^{(1)}$ to the next axial measurement at $z^{(2)}$ which are denoted in the lower right box of each tile with name a), i. e. the upper row. Hence, e. g. in case of the first column all axial slices are used for reconstruction while in the second column only every 2nd Z-slice symmetrically positioned around the center slice $z^{(0)}$ were used. For the first row, the 3D PSFs in case of the central detector pixel $a^{(0)}$ are shifted to the Z-slices $z^{(p)}$ selected for measurement and are summed up axially to display the spatial extent covered by this imaging scenario. The plots are given as an axial cut at $y^{(0)}$. The second row shows the TU and the third row the mDEC based zLEAP reconstruction results. The NCC similarity to the 3D ground-truth image is shown in the lower right box of each tile. Figure 4.6II shows the NCC as function of the distance Δz to the next measured Z-slice as 1D plot for a range of $\Delta z \in \{1, 2, \dots, 16\}$.

While Figure 4.6I1a-c) demonstrate the best possible reconstructions with the given techniques, the difference between TU and mDEC is particularly noticeable in the bumpiness at the upper and lower end of the big spherical shell. In the displayed summed PSF profiles, which are not used for any calculation but just for visualization, a structuring effect becomes apparent starting from the 3rd column, but can already be seen clearly in the 4th column of Figure 4.6I. The same axial rippling like in 4a) is visible in the TU reconstruction 4b) thereby leading to a noticeable decrease of NCC similarity. The effect is weak in 4c). By comparison of the reconstructions of column 2-4 with column 1 it can be stated by visual impression that for the given simulation only measuring every third Z-slice leads to a sufficiently good reconstruction in case of TU processing. This is in agreement with the in Section A.11 given definition of a high quality reconstruction. In case of mDEC even measuring every 6th Z-slice is still acceptable, see green curve of Figure 4.6II.

The reconstruction is NCC-optimal when every single Z-slice was used and decreases almost continuously for increasing distances and thus a decreasing number of Z-slices used (Figure 4.6II). The case $\Delta z = 16$ corresponds to recording only the central slice $z^{(0)}$. Using only the central five views for reconstruction and together with additional Z-planes that are spaced within $\Delta z \leq 5$ the inversion problem Equation 4.3 is overdetermined or exact while in case of $\Delta z > 5$ it is underdetermined leading to a smaller reconstructable

axial thick slice than in the calculations of $\Delta z \leq 5$ and thus increasingly worse NCC results. This can be countered by using more lateral view $N^{(A)}$, but still does not lead to reconstructions with NCC ≥ 0.9 and is thus not further included for consideration here.

The comparison demonstrates how TU can be used to reconstruct PSF and noise-limited axial slices. Since in practical applications 3D datasets should be recorded, but the phototoxic load on the sample should be kept low, TU and mDEC suggest that recording every 4^{th} Z-slice, or about 2.7× the axial Nyquist-limit of the system PSF h^{ISM}, still allows for a NCC similarity to the ground-truth of $C \geq 0.9$ (Figure 4.6II).[10] This leads further to at least 2× faster stack acquisition and 2× reduction in data size while allowing for the same reconstruction quality.

Finally, a performance comparison between 3D mDEC and TU demonstrates the fast reconstruction performance of TU needing $\approx 2s$ for processing a $[5, 128, 128, 64]$ Z-stack while mDEC needed $\approx 22s$ for the same task. On the other hand, TU needed (peak amount) $\approx$ 34× as much RAM for this processing than mDEC needed on the Graphics Processing Unit (GPU) memory. For an in-depth analysis see Section D.6.

4.3 TU EXPERIMENTS ON THE AIRYSCAN SYSTEM

The processing strategies presented in Section 4.2.1 will be evaluated in this section using data measured by means of the commercially available Zeiss Airyscan confocal scanning system *LSM 880*.

Thus, this section addresses the following issues: 1) determination of experimental PSF and comparison of 3D TU with multi-view deconvolution results using different PSFs, 2) comparison of ooF rejection capability of TU with deconvolution-based methods, and 3) comparison of 3D reconstruction capability of TU with deconvolution-based methods.

4.3.1 *Airyscan Setup*

Introduced by Zeiss in 2014, the LSM 880 with Airyscan extension is a commercial confocal scanning unit with ISM image acquisition capability [65]. The detailed setup is given in [65] and is not reprinted here, because its basic shape does not differ much from

10 Note, that the PSF h has to be sampled at least at the Nyquist limit for this result to hold true.

Figure 4.1I except that a hexagonal detector, called Airyscan detector, is used instead of the depicted emCCD. For its basic detector pixel distribution see Figure 4.7II. Instead, a short description will be given. Two laser lines can be injected simultaneously or sequentially. A GM scans the angles of the BFP of the objective, creating a lateral offset of the focus in the conjugate image plane, i. e. in the sample. The objective is a Zeiss 1.42*NA*, 63× plan apochromat DIC, which was used with $n^{(\mathrm{im})} = 1.52$ immersion oil (applied to the cover-slip). An intermediate pinhole including a focus and defocus lens pair removes coarse ooF light before excitation light is blocked further by a detection filter wheel. The focal spot changes its size depending on the used wavelength as well as lens. In order to normalize the first 1.25 Airy Unit (AU) of the detection wavefront to the diameter of the detector array, mechanically adaptable magnifying optics are built in. The detector array consists of 32 Gallium-Arsenide Phosphate (GaAsP) Photo-Multiplying Tubes (PMTs) that can be read and processed individually. In total, the scanning system can record a field of view (FoV) of 512 × 512 pixels at $13\,fps$, which corresponds to an average pixel dwell time $t^{(\mathrm{pix})} \approx 0.3\,\mu\mathrm{s}$.

4.3.2 *Measuring and Comparison of Airyscan PSF*

For the characterization of an incoherent imaging system, the PSF as well as the associated transfer function is particularly suitable. Once these are known, post-processing steps such as the presented TU or mDEC can be applied. The system PSF of the Airyscan system is now 1) determined directly from a recorded multi-view Z-stack of $100\,nm$ sized beads, 2) a 3D Savitzky-Golay Filter (SGF) fit of 1) and 3) simulated based on the parameters of the (optical) components and manual adjustments after comparing to the experimental PSF. The three PSFs will then be used for zLEAP TU and mDEC processing of a ROI of a measured experimental dataset. Finally, the results are compared w. r. t. reconstruction quality and one PSF will be chosen to be used for further processings within this section.

Alexa568 (Thermofisher, Invitrogen) stained $100\,nm$ beads with oil immersion were used for experimental determination of the PSF. In Figure 4.7, first the selected beads (see Figure 4.7I), second the determined effective detector pixel positions (see Figure 4.7II) and finally a comparison of the zLEAP and mDEC reconstructions w. r. t. to the used PSFs (see Figure 4.7III) are displayed for the central detector-pixel. For the reconstructions

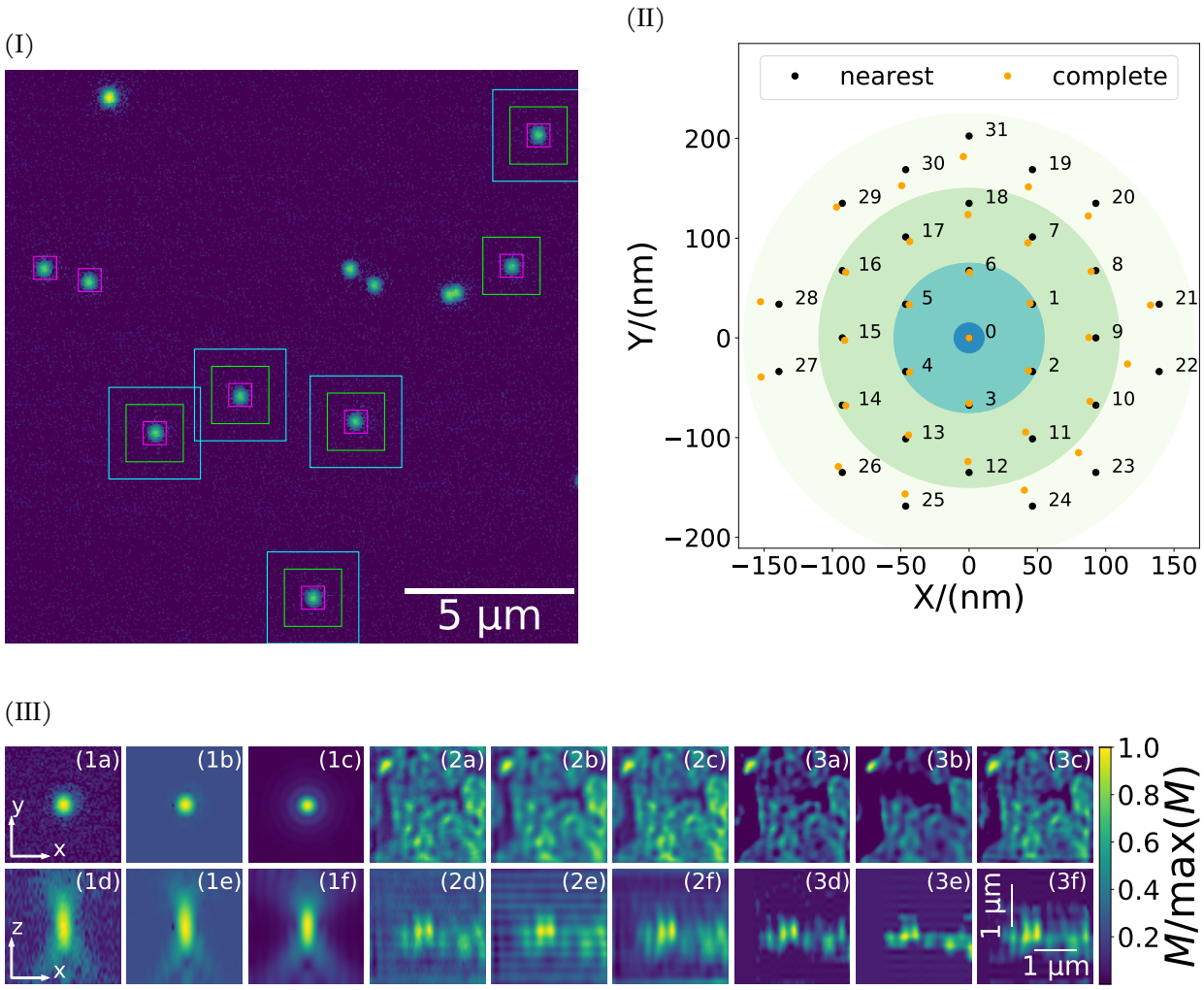


FIGURE 4.7: **Airyscan PSF Comparison.** (I) Selected Beads for calculating the average system-PSF shown at slice $z = 14$. ROI sizes are color marked: small (magenta), middle (green) and large (turquoise). (II) Calculated $d_a^{(\mathrm{ISM})}$ positions of all views (complete) and via calculation using pinholes 1 and 6 (nearest). Positions are displayed in scan coordinates. Nominal rings are color coded. (III) Upper row: X-Y-slice at (1) $z = 12$ and (2-3) $z = 14$. Lower row: X-Z-slice at (1) $y = 32$ and (2-3) $y = 50$. (1) PSF a+d) measured, b+e) Savitzky-Golay Filtered and c+f) simulated. (2) zLEAP TU and (3) zLEAP mDEC reconstructions using (1a-c) PSFs.

every 2nd, i.e. $p = 2$, Z-slice was used. Figure 4.7III is divided into blocks of three columns and 2 rows (a-c,d-f) each. The columns of the blocks are linked by means of e. g. 1a+d) are cuts of the experimental PSF that is used to calculate 2a+d) TU and 3a+d) mDEC reconstruction. The same follows for the second and third columns in each case and is discussed in more detail below.

In the considerations of this work, the PSF is assumed to be translationally invariant (Table A.2). Thus, determination at the top or bottom of the image will result in an identical PSF. To account for these assumptions, compensate for axial offsets as well as local distortions, and obtain a better SNR, different beads were selected distributed over the entire FoV and then averaged, see Figure 4.7I. The beads are extracted with a ROI of $N^{(\mathrm{roi,s})} = [16, 16, 41]$ (magenta), $N^{(\mathrm{roi,m})} = [40, 40, 41]$ (green) and $N^{(\mathrm{roi,l})} = [64, 64, 41]$ (turquoise). Due to the lateral expansion of the PSF during axial defocusing, edge effects of the smaller ROIs become noticeable. Here, the largest ROI achieved the best results, so that further comparisons and processing was performed with this one. The total system PSF of the central detector pixel, determined in the presented way, differs slightly from the actual PSF because it has already been folded with the 100 *nm* beads. Figure 4.7III shows an X-Y cut (1a) as well as X-Z cut (1d). The particular PSF is then smoothed and denoised by using a SGF (see Section 2.2), with window-size $w^{(\mathrm{SGF})} = 5$ pix and $p^{(\mathrm{SGF})} = 2$, see Figure 4.7III (1b+e). Based on the visual impression, the experimental PSF exhibits a slight axial tilt and spherical aberrations. Incorporating these findings into the synthetic PSF calculation leads to a reduced $NA = 1.3$ and application of slight spherical aberrations with $c^{(\mathrm{spher})} = 0.1$ (Equation 2.2), see Figure 4.7III (1c+f). Comparison with the measured PSFs shows a smaller axial extent of the simulated PSF.

In order to determine the simulated PSF not only for the central but all detector pixels (see Figure 4.1III), the individual actual distance of the detector pixels from the central pixel in sample coordinates is needed. First, the methods *complete* and *nearest* are used to determine the maximum position d_a of each view $h_a^{(\mathrm{ISM})}$. Then, assuming a shift factor of $m^{(\mathrm{ISM})} \approx 1/(1 + (\lambda^{(\mathrm{em})}/\lambda^{(\mathrm{ex})})^2)$ according to Roth *et al.* [9] a guess of the original detector position is constructed via $(1 - m^{(\mathrm{ISM})}) \cdot d_a$. Finally, a parameter-search tries to minimize the distance between the determined effective pixel positions $d_a^{(\mathrm{nearest})}$ (see below) and the view positions of the simulated $h^{(\mathrm{ISM})}$.

Two methods are presented. 1) By the maximum and Center of Mass (CoM) (Equation 3.5) preserving property of SGF, the effective pixel distances $d_a^{(\mathrm{complete})}$ (position of the maximum of the non-deformed $h_a^{(\mathrm{ISM})}$) of all views are first determined by correlating each individual detector pixel with the central pixel. The positional deviation of the correlation maximum from the central pixel is then assigned to $d_a^{(\mathrm{complete})}$, see Figure 4.7II (*complete*, orange). 2) *nearest* limits the correlation-based search to two pixels located close to the central pixel, which are expected to still have a high SNR. Using correlation of

pixel 1 and 6 to the central detector pixel 0 (see Figure 4.7II), the basis vectors for the distribution of detector pixels are determined. The distances $d_a^{(\text{nearest})}$ for all other pixels are calculated by using the basis vectors and by knowing the basic hexagonal structure of the detector, i. e. the indexed positions of the pixels w. r. t. the hexagonal unit-cell, see Figure 4.7II (*nearest*, black).

Single detector pixels can be combined according to their nominal position and distance to the central pixel, leading to rings which are color coded in Figure 4.7II. Ring 1 (blue) only contains the central pixel, ring 2 (turquoise) contains pixels 1-6, ring 3 (pastel green) contains pixels 7-18 and ring 4 (pale green) contains pixels 19-31.

A comparison of the calculation results shows a subpixel exact match of $d_a^{(\text{nearest})}$ and $d_a^{(\text{complete})}$ for ring 2, but up to half a sample pixel separation in ring 4. In ring $1-3$ $d_a^{(\text{complete})} \leq d_a^{(\text{nearest})}$ tends to hold, while in ring 4, more local offsets thwart the estimation of a global trend. For further processings the effective pixel distances $d_a^{(\text{nearest})}$ are used.

Using zLEAP (see Figure 4.7III (2a-f)) and zLEAP-like multi-view deconvolution with $p = 2$ (see Figure 4.7III (3a-f)) of the ROI in the sample further investigated in the coming sections (see Figure 4.8I), the usability of the measured (1a+d), filtered (1b+e) as well as calculated PSF (1c+f) for subsequent post-processing methods is visually evaluated. Since the ground truth is unknown, the processing methods are run for a sequence of reconstruction parameters and a subjectively appropriate result is selected. While the reconstructions using measured and filtered experimental PSF are of high quality they show axial striations (see Figure 4.7III2d+e) where each Z-position at which a measured image $M^{(\text{p})}$ was available for reconstruction is brighter than the intervening reconstructed regions. This effect is more severe with smaller ROI-sizes used for PSF determination. The reconstructions using the simulated PSF (1c+f) suffer less from axial striations while achieving a comparable quality to the mDEC reconstruction (3c+f). Among the mDEC reconstructions (3a-f) the results 3c+f) using the simulated PSF look subjectively best. Thus, for further processing the simulated PSF (1c+f) was used.

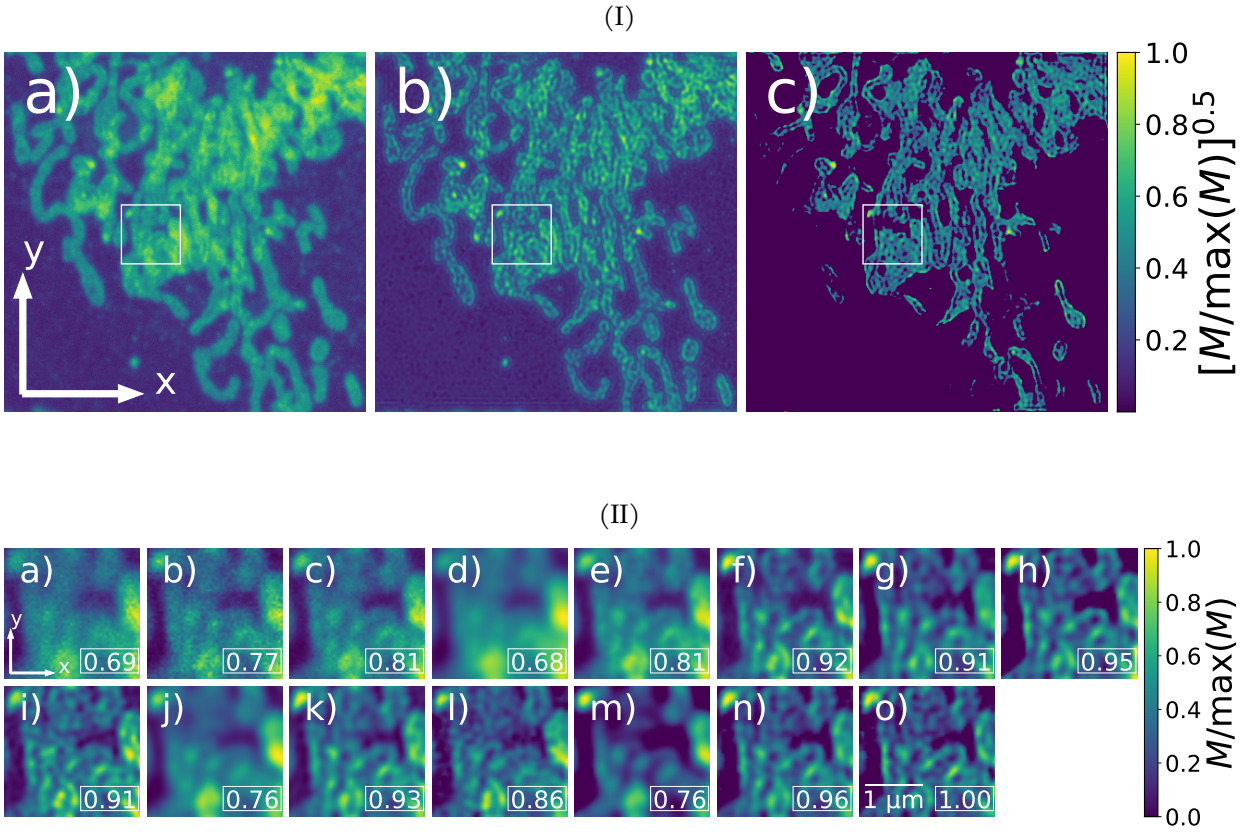


FIGURE 4.8: **2D Reconstructions of TOMM20.** (I) Reconstruction results of TOMM20 at $z = 14$ for a) PiRe, b) WD and c) PiRe deconvolution. White box marks ROI. (II) reconstruction results on ROI of (I) at slice $z = 14$. a) open CLSM at 1.25 AU, b) CLSM at 0.6 AU pinhole size, c) PiRe at full 1.25 AU pinhole size, d) WAVG e) nWAVG, f) generalized WD, g) confocal single-view DEConvolution (cDEC), h) pixel reassigned single-view DEConvolution (prDEC), i) ooF using TU (oofTU) j) 2D to 3D TU k) zLEAP using every 2nd image-Z-slice. l-o) multi-view deconvolutions using: l) $z = 14$ slice and 2D PSF, m) 2D slice and 3D PSF, n) like k), o) full 3D image and PSF. Images are normalized to their individual maximum local density (or amplitude, respectively). NCC results are displayed in the lower right corner of each panel.

4.3.3 *Comparing TU and Deconvolution - Out-of-Focus Rejection of TOMM20*

An Alexa568 labelled import receptor subunit TOMM20 in the outer membrane of mitochondria of Human Bone Osteosarcoma Epithelial Cells (U2OS) cells is used as experimental data. The raw multi-view Z-stack was stored as 16bit uint, consists of $N = [32, 400, 400, 41]$ pixels and was recorded at $5\,fps$ imaging speed. A cut of the processed multi-view 3D

data-stack at $z = z^{(14)}$ is shown in the case of (a) PiRe[11] (Figure 4.8Ia), WD (Figure 4.8Ib), and prDEC (Figure 4.8Ic) processing. The white box marks the ROI used henceforth for comparisons. Because the labelled structure TOMM20 is located in the outer membrane of the mitochondria, the goal of the reconstructions is to obtain the lowest-noise, continuous, thin cell borders possible from the noisy, filled cell structures. In the 3 methods compared, PiRe (Ia) shows almost no such separation of mitochondria body and membrane, WD (Ib) and prDEC (Ic) succeed better. prDEC (Ic) shows a stronger axial sectioning, whereby in comparison to the WD (Ib) the contrast to the background is larger and altogether fewer structures are contained in the same Z-slice. Thus, prDEC (Ic) is chosen as the subjectively closer to reality and thus better reconstruction.

The ooF rejection ability is tested using methods that require knowledge of the PSF (IId-o) and those that do not (IIa-c), see Figure 4.8II which shows the ROI of Figure 4.8I. For NCC based reconstruction parameter determination, the subjectively best reconstruction using full multi-view deconvolution mDEC was chosen and used as ground truth. For input only the in-focus $M^{(14)}$ slice at $z^{(14)}$ was used while for processing, depending on the method, 2D or 3D multi-view PSFs were used. Suitable reconstruction parameters were determined automatically for all methods in a predefined range. For open CLSM, the images of all detector pixels are summed (Figure 4.8IIa), whereas for CLSM only the inner detector pixels in the diameter of $0.6\,AU$ are summed (Figure 4.8IIb). PiRe (Figure 4.8IIc) uses all detector pixels, but shifts them beforehand to their most likely emission position (Equation 1.26). WAVG (Figure 4.8IId), nWAVG (Figure 4.8IIe), WD (Figure 4.8IIf) and mDEC (Figure 4.8IIg+h,l-o) use all detector pixels, TU (Figure 4.8IIi-k) limit the number of pinholes used to the number of Z-slices to be reconstructed, i. e. $N^{(a)} \stackrel{!}{=} N^{(z)}$, so as not to overdetermine the system of equations used to compute the inversion matrix (Equation D.1). In the case of ooF rejection using TU (for comparison named oofTU) and associated mDEC (for comparison named oofDEC), this means that only one Z-slice of the PSF and image was used, i. e. a 2D mDEC.

In subjective visual comparison, mDEC (Figure 4.8IIo) gives the best reconstruction, but differs only slightly from WD, cDEC (Figure 4.8IIg), prDEC (Figure 4.8IIh), 2D mDEC (oofDEC, Figure 4.8IIl), and zLEAP mDEC (zleapDEC, Figure 4.8IIn). The TU reconstructions achieve mutually similar reconstructions whose ooF rejection is comparable

11 For calculation see Equation 1.26. Shifts are determined as described in Section 4.3.2.

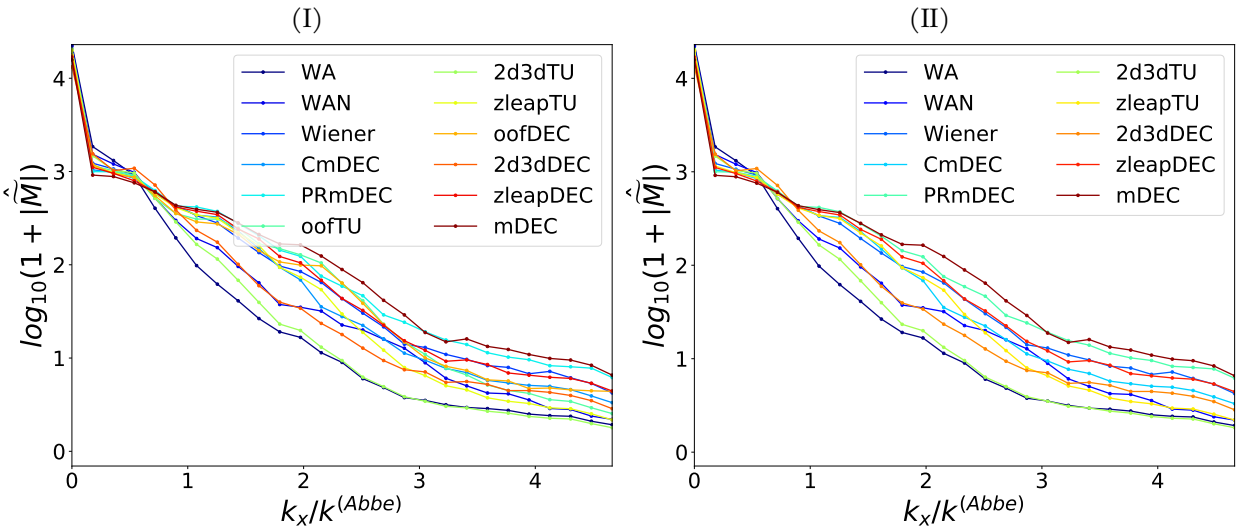


FIGURE 4.9: **Radial Frequency Sums of TOMM20 Reconstructions.** (I) Radial sums of 2D (see Figure 4.8) and (II) 3D (see Figure 4.8) reconstructions. The calculated modulus is displayed log-normalized while all curves are normalized to same 0-frequency value.

to mDEC, but the reconstructed thickness of the membrane structures is larger. The TU reconstructions TU (2d3dTU) and zLEAP are on average as NCC similar to mDEC as cDEC and prDEC. Subjectively, 2D mDEC sharpens the sample slightly more than 3D mDEC with similar ooF rejection, subject to the above criteria, but retains more existing details in the same focal plane, which in mDEC have already been moved to other Z-slices. The 2D TU reconstruction (oofTU), on the other hand, appears to reconstruct about the same number of object features as the other TU reconstructions in the same Z-slice, but tends to sharpen the image less than e. g. zleapTU.

A comparison of the lateral Fourier frequency moduli supports the above findings, see Figure 4.9I. The graphs are normalized to the same total photon number ($\widetilde{M}(|\vec{k}| = 0)$) so that the relative distribution of available photons at each spatial frequency can be compared. For spatial frequencies $k_x \leq 1.5 \cdot k^{(\mathrm{Abbe})}$, the amplitude distributions of the deconvolutions and the TU reconstructions are similar, while for $k_x > 1.5 \cdot k^{(\mathrm{Abbe})}$ the TU tends to assign less weight to the high frequencies and for $k_x > 3 \cdot k^{(\mathrm{Abbe})}$ even approach the amplitude distribution of the WAVG (WA). The distribution of 2D deconvolution results (oofDEC) illustrates the higher degree of sharpness of this reconstruction noted earlier, as its amplitudes for $k_x \geq 1.5 \cdot k^{(\mathrm{Abbe})}$ are above those of all other amplitudes, while the reconstructed image still looks subjectively reasonable (Figure 4.8II l). The amplitude

distribution of prDEC is above the amplitude distribution of cDEC for $k_x \geq 1 \cdot k^{(\mathrm{Abbe})}$, but further away from mDEC for $k_x \leq 1.5 \cdot k^{(\mathrm{Abbe})}$ which results in a lower NCC similarity of prDEC to mDEC than cDEC.

4.3.4 *Comparing TU and Deconvolution - 3D Reconstruction of TOMM20*

Using the complete 3D data set of the sample already used in Section 4.3.3, the reconstruction capability of the TU (2d3d) and zLEAP is now compared in Figure 4.10 with the other methods presented. Here, the graphic consists of 2× two rows, where always the upper one row shows a lateral cut at $z^{(14)}$ and the lower row an axial cut at $y^{(15)}$. Every method of the comparison is presented in two rows of one column each, i. e. with two panel images.

Again, with an appropriate choice of parameters, the mDEC is used as ground truth for the NCC based automated parameter determination of the other methods. TU (2d3dTU) (Io+p) and 2d3dDEC (Is+t) reconstruct a 3D estimate $\hat{S}$ of the 3D object from the single $z^{(14)}$ slice using the 3D PSFs of all detector pixels. Finally, zLEAP, TU and complete mDEC use every other Z-slice, i. e. $p = 2$, of the measured image stack along with the entire 3D PSF for reconstruction of the 3D object distribution.

While the non-deconvolution methods (a-h) are axially very smeared out, all other methods succeed in better sectioning. But again visually the WD as well as variants of mDEC (Ii-n,u-x) appear very similar. However, axial sectioning is best for 3D mDEC (Ix) and zLEAP deconvolution (zleapDEC, Iv), which also makes them quite NCC similar, although zleapDEC uses only half the Z-slices compared to mDEC. prDEC has some axial reconstruction artifacts (e. g. right edge of In), resulting in a lower NCC similarity to mDEC (Figure 4.10) than cDEC, yet edges are more prominent in prDEC and better axial sectioning is achieved. This is also confirmed by the significantly higher modulus values of prDEC for $k_x \geq 1.5 \cdot k^{(\mathrm{Abbe})}$ (Figure 4.9II).

The TU (2d3dTU, Io+p) as well as the 2d3dDEC (Is+t) confirms the theoretical expectations and simulation results (Section 4.2.3). Although the simulated PSF is slightly spherical aberrated it is not asymmetric enough and thus both algorithms distribute axial

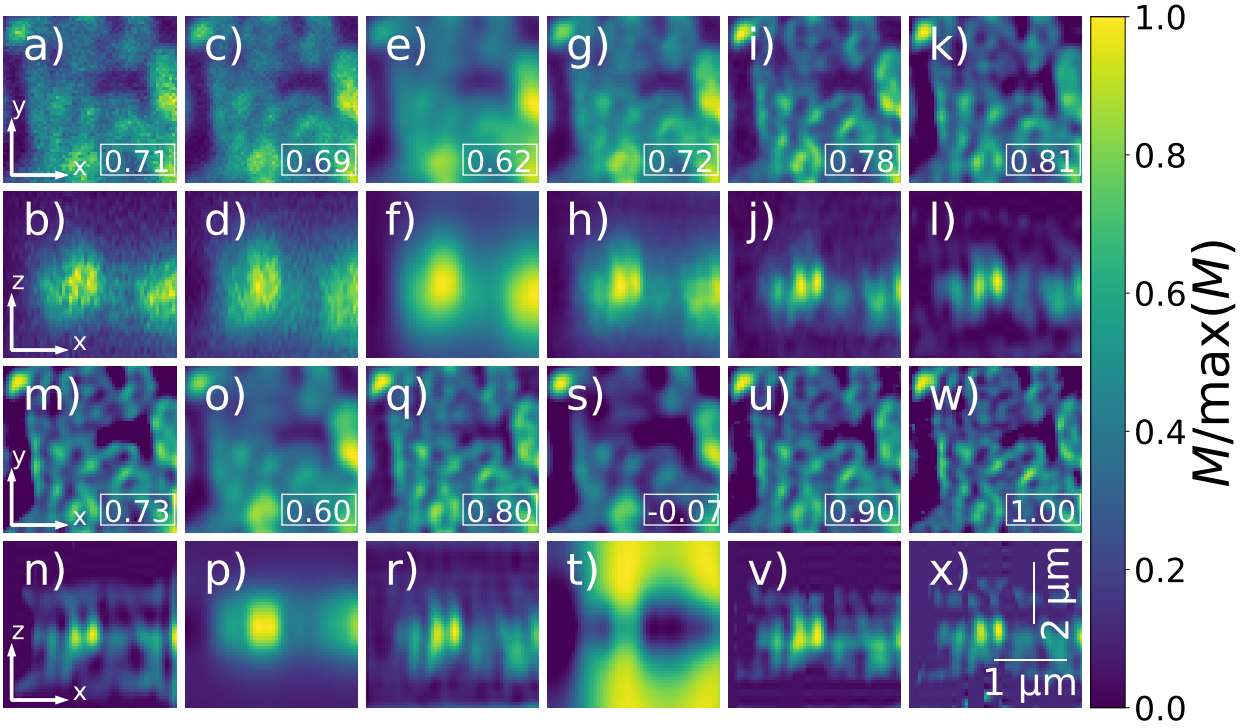


FIGURE 4.10: **3D Reconstructions of TOMM20.** Reconstruction results of ROI (see Figure 4.8I). First and third row show an X-Y-slice at $z^{(14)}$, second and forth show an X-Z-slice at $y^{(15)}$ of the respective reconstruction methods. a+b) CLSM at 0.6 AU pinhole size, c+d) PiRe at full 1.25 AU pinhole size, e+f) WAVG g+h) nWAVG, i+j) generalized WD, k+l) cDEC, m+n) prDEC, o+p) 2D to 3D TU q+r) zLEAP using every 2nd image-Z-slice. s-x) multi-view deconvolutions using: s+t) 2D slice and 3D PSF, u+v) like q+r), w+x) full 3D image and PSF. Images are normalized to their individual maximum local density (or amplitude, respectively). NCC results are displayed in the lower right corner of the 1st and 3rd row.

information approximately symmetrically around the central reconstruction slice. Due to this vague and axially significantly extended reconstruction, both methods achieve an approximately similar NCC to mDEC of $C^{(2d3dTU)} = C^{(2d3dDEC)} \approx 0.6$.

In the case of zLEAP, TU achieves an 0.8 NCC similarity to mDEC while zleapDEC achieves 0.9. In particular the axial sectioning of zleapDEC is better than TU. In the unmixing process of TU, parameter selection must be used to decide between lateral sharpness with good axial sectioning versus a smooth axial amplitude transition. The selected parameters achieve the maximum possible NCC similarity to mDEC, but a weak axial step pattern is discernible which hints where input Z-slices where placed.

It has been shown that TU-based zLEAP can accelerate the measurement process and reduce the data set to be recorded without significant loss in image quality while maintaining fast processing and low implementation overhead.

4.4 CONCLUSION AND OUTLOOK

For the presented method TU makes use of the additional views of the sample which are encoded into the different detector pixels. The additional sample information is used for ooF rejection, reconstruction of an axial thick slice from 2D multi-view data, and data reduction by measuring each p^{th} slice using zLEAP were tested with a set of simulations. zLEAP thereby lowers the phototoxic load on the sample, increases the data acquisition speed by 2×, and reduce the stack size by 2× in case $p = 2$. All investigated reconstruction methods achieved the NCC best result when all detector pixels were used. Since already by using the five central pixels about 98% in case of deconv3D of the achieved NCC maximum was achieved, this suggests a reduction of the amount of required pinholes used for reconstruction in favor of reduced memory requirements.

The additional information from the different views can also be used to achieve a better SNR via PiRe allowing to recover higher nonlinear orders of fluorescence in saturated excitation imaging as compared to CLSM schemes. For an extensive analysis, see Section E.1.

The findings of the simulations were confirmed experimentally using data of fluorescent beads and stained U2OS cells acquired with a Zeiss Airyscan LSM880 setup.

Overall, it is shown that the TU reconstructions do not provide NCC optimal reconstructions with respect to mDEC, which was subjectively chosen on the example of the experimental data presented, but do provide a remarkable reconstruction quality with consistently fast processing.

In contrast to excitation and emission PSF manipulation-based methods, TU does not require further modifications of the measurement setup and can be used directly. However, due to the axial symmetry of the PSF used, TU cannot distinguish, in the case of only one measured Z-slice, whether ooF influences originate from above or below the Z-slice under consideration. In this case, TU can be used for ooF-rejection as the precise assignment of the ooF to other Z-slices is irrelevant. By measuring further Z-slices and

using them for reconstruction, the necessary asymmetry can be generated indirectly here, thus enabling the distinguishability of the axial position of the object information. Due to the direct usability of TU without PSF-manipulation, the *NA* of the measurement setup is not degraded, can be fully used and the system transfer efficiency is not reduced. Although the reconstruction by deconvolution used for comparison yields overall more stable and NCC better results, TU can be used performantly on the CPU and without complicated setup steps. Also the parameter search for TU turns out to be more compact than for the used deconvolution. For further analysis the influence of the detector pixel shape factor on the detector mask choice is interesting. The presented method zLEAP can also be applied laterally, which allows further time and measurement savings.

TU was implemented in Python using the *C*-optimized Numpy interface. TU is currently not fully but only partially processed in parallel despite its excellent suitability for parallel processing. The implementation of deconvolution on the GPU, on the other hand, uses all CUDA units of the entire card simultaneously. Due to the GPU RAM limitation, memory-consumption-optimized deconvolution methods have been applied, while TU currently does not yet optimally use memory resources. Optimization of memory usage and parallelization of the unmixing process of TU is conceivable by strict use of Cython as well as adaptation to *C* standards or direct implementation in *C* and is left as an outlook for future engineering. To overcome the fast increasing necessary RAM for calculation of the unmixing matrix, using Hierarchical Data Format Version 5 (HDF5) [66] containers and therewith directly calculating on the system storage might further increase the methods versatility.

5

CONCLUSION

Something unique, something imperishable formed in my head. An artful structure of words and sentences, which materialized like an extraterrestrial creature of strange beauty in my thinking - and spoke to me, in immaculate verses. It was [...] *a gift from the stars.*

— Walter Moers

The microscopy landscape is in a constant state of change. New labs are being built and equipped with the latest imaging methods, old labs are being upgraded, imaging centers are offering their services across labs and systems can now even be rented out. Existing microscopes are not immediately discarded, but still serve as workhorses and continue to be used as a reference or gold standard. Widefield Microscopy (WF) and Confocal Laser Scanning Microscopy (CLSM) type systems still make up the majority of the currently existing and used microscopy systems. Also, in terms of price, these are currently still more affordable than state-of-the-art high-end microscope systems.

Many new microscopy methods are based on fundamentally the same geometries as in the case of WF or CLSM setups. The understanding, analysis, and improvement of these basic methods can also be extended to novel methods where appropriate, making them of particular interest.

The goal of this work was to consider microscopy systems holistically and across methods. This was done by starting with the fundamental problem of focus quality and loss and continuing with a comparatively shallower problem of data reliability due to jitter effects. Finally, after identifying ways to increase data reliability in the case of WF and CLSM, a method was presented to get more out of existing Image Scanning Microscopy (ISM) 2D data without changing the setup or purchasing extra large computational systems.

In chapter 2, the presented software autofocus represents a possibility to keep the actual focus position by evaluation and subsequent classification of the recorded data w. r. t. the proposed quality measure Q. The presented metrics were stress tested in terms of required

computational power and time as well as susceptibility to system aberrations. The results can be easily applied to existing setups by processing the recorded images live on the computer and, according to the obtained value Q, the focus position can be evaluated and adjusted. However, this process may be slow, but this can be remedied by system extension using Field Programmable Gate Array (FPGA). The presented method was tested for 2D and 3D object distributions. Since the captured images are reduced to a single value Q, in the case of a 3D stack all acquired images must first be completely recorded and each slice evaluated. The advantage of this approach is that the similarity of different slices and therefore the system drift can be determined by comparing the slicewise Q values and not by correlating the entire 3D stacks. This can significantly save system resources. However, in case of a rapid change of the sample between two consecutive temporally spaced measurement points, e. g. due to cell genesis or strong contrast change, the method has its limitations and hardware based methods have to be resorted to. In the future, the method may be directly transferred to Neural Networks (NN) to e. g. address the problem of rapid image content change and provide augmented feedback directly in a live view of the measurement using one-shot applications.

In chapter 3, the task of removing imaging artifacts due to random (micro) movements of a mouse eye was accomplished using a novel dejitter algorithm. The algorithm was adapted to the specific problem and uses a correlation-based measure with respect to an existing coarse continuous structure for finding global drift as well as local, line-by-line correction. Random jitter occurs not only on the living organism under consideration, but in any measurement situation, since measurement devices are always connected to their environment and even the best vibration-free mounting cannot completely dampen strong, spontaneous oscillations. Thus, the presented dejitter algorithm has a further and more general field of application. In the future, it is conceivable to use some output parameters of the dejitter algorithm as a measure for assessing the reliability of the recorded data both live and in post-processing.

In chapter 4, the deconvolution method of Thick slice Unmixing (TU) is presented. It is based on a 3D forward model together with ISM imaging, where the different detector pixels measure different projections of a 3D volume of the observed sample (views). Thus, by knowing or simulating the 3D system Point Spread Function (PSF), a 3D distribution of the sample can be reconstructed based on the different views. Three application scenarios for TU were presented. In scenario one, TU is used for Out of Focus (ooF) rejection.

Due to a compact formulation of the TU equations on only one slice to be reconstructed based on a single-slice multi-view dataset of a three-dimensional sample, only a few computational steps are needed here, which allows the method to be used on hardware with low computational power. In particular, computational power can be saved by limiting the imaging model in the reconstruction to include only a small axial region around the considered Z-slice. In Scenario two, a 3D volume is reconstructed from a 2D multiview dataset. For comparison, existing deconvolution methods were extended to include the ability to compute a 3D stack from a 2D multi-view dataset. Although the Maximum A Posteriori likelihood (MAP)-based deconvolution results have on average a higher Normalized Cross Correlation (NCC) similarity to the used *in-silico* ground-truth the differences to the the volume-reconstructing TU results are small.

LEAve-out z-Planes (zLEAP) was presented as scenario three, in which only every p^{th} slice is recorded and intervening slices are reconstructed as best as possible by knowing the multi-view 3D system PSF and including the full Z-scan multi-view information in the model. The results compared to the original dataset as well as deconvolution are of surprisingly good quality, as both computational approaches, model complexity and implementation complexity are significantly lower than in the case of MAP-based deconvolution. zLEAP is particularly intended for the use case of Nyquist correct imaging with every second slice omitted. Here, every intermediate slice can be reconstructed and thus the Nyquist correctly sampled image is gained while the phototoxic sample load is significantly reduced. While the MAP-based deconvolution scales with $N \log(N)$[1] with the number of Z-slices the inversion matrix of TU scales proportionally to the number of views used per Z-slice, which could lead to a quadratic dependence on the number of Z-slices. Thus, the presented implementation of TU is comparatively performant only up until a Z-stack size or amount of views threshold. Unlike other implemented deconvolution methods, TU does not require Graphics Processing Unit (GPU) to run in a time-performant manner, does not require cumbersome (driver) setup as well as costly computer components, and can be used straightforward in the form of matrix multiplication. The strength of TU lies in its reusability use in case of stable imaging parameters and PSFs for different measurement data, since here the inversion matrix has for each spatial frequency to be computed only exactly once thereby saving computational time for all

1 Due to properties of Fast Fourier Transform (FFT) calculation.

consecutive applications. The easy manageability of TU could also be seen as another user advantage, since only a few parameters have to be set here for a robust reconstruction. For future developments, the parallelization and optimization of the view usage in the inversion matrix determination is of particular importance, since this can once again significantly save system resources and accelerate the processing. Thus TU can be still more competitive compared to existing MAP-based deconvolution methods.

For the presented processing methods a complete own Python toolbox *MicroPy* [44] was created and further toolboxes (*InverseModelling* [67], *NanoImagingPack* [38], *tiler* [42], UC2-Software-GIT [68]) were actively advanced and supported in collaborative development.

The presented post-processing based enhancements to the entire optical system are directly applicable and have a direct impact on image acquisition quality (autofocus), reliability of measurement data (dejitter) and quality of the 3D reconstruction (TU). In particular, the zLEAP use case of the TU method can noticeably reduce scan duration and recording volume. All methods/reconstructions were implemented on a laptop using the CPU or its dedicated small GPU, allowing any researcher in e. g. biomedical imaging not only to extend their existing methods without changes to the ISM setup, but also to process the acquired data directly on their laptop. The present work demonstrates how a holistic improvement of the measured data can be achieved by a set of well chosen imaging optimizations.

A

APPENDIX: GENERAL INTRODUCTION

A.1 GREEN'S IDENTITY

Here Green's second identity is invoked to connects the bound volume-integral of volume V of a set of differentialable functions with their surface-integral A [69]:

$$\iiint_V d^3\vec{x}\ \left(E(\vec{x})\nabla^2 G(\vec{x}) + G(\vec{x})\nabla^2 E(\vec{x})\right) = \iint_A d^2\vec{\chi}\ \left(E(\vec{x})\partial_{\vec{n}} G(\vec{x}) + G(\vec{x})\partial_{\vec{n}} E(\vec{x})\right) \quad \text{(A.1)}$$

where $\vec{n}$ marks the surface-normal on A that points away from the volume V.

A.2 HELMHOLTZ-EQUATION

To describe light and matter interaction the Maxwell-Equations (Equation A.2 - A.5) ([70]):

$$\sum_{mn} \epsilon_{lmn}\partial_m H_n = \partial_t D_l + j_l \quad \text{, with:} l, m, n \in \{x, y, z\} \quad \text{(A.2)}$$

$$\sum_{mn} \epsilon_{lmn}\partial_m E_n = -\partial_t B_l \quad \text{(A.3)}$$

$$\sum_{l} \nabla_l D_l = \rho \quad \text{(A.4)}$$

$$\sum_{l} \nabla_l B_l = 0 \quad \text{(A.5)}$$

with $E_m \equiv E_m(\vec{x}, t)$ electric vector, $H_m \equiv H_m(\vec{x}, t)$ magnetic vector, $D_m \equiv D_m(\vec{x}, t)$ electric displacement, $B_m \equiv B_m(\vec{x}, t)$ magnetic induction and j_l electric current density, will be used. ∂_m are the partial derivatives with respect to the 3 spatial dimensions $\sum_m \partial_m = \partial_x + \partial_y + \partial_z$ and ∂_t is the time-derivative.

Assuming *time-harmonic (=sinusoidal) waves, very slowly - with respect to the period of interaction with fields - moving bodies and isotropic material*, the following material equations are obtained:

$$j_m = \zeta^{(\text{cond})} E_m \quad \text{(A.6)}$$

$$D_m = \zeta^{(\text{elec})} E_m \quad \text{(A.7)}$$

$$B_m = \zeta^{(\text{magn})} H_m \quad \text{(A.8)}$$

where $\zeta^{(\text{cond})}$ is the n, $\zeta^{(\text{elec})}$ the electric permittivity and $\zeta^{(\text{magn})}$ the magnetic permeability. Typically these are tensors, but it suffices to assume them to be *constants* within the framework of this thesis. Taking the *curl* $= \nabla\times = \sum_o \epsilon_{pol}\partial_o$ of Equation A.3 yields for the left side of the equation

$$\begin{aligned}\epsilon_{pol}\partial_o\left(\epsilon_{lmn}\partial_m E_n\right) &= \sum_{olmn}\epsilon_{pol}\epsilon_{lmn}\partial_o\partial_m E_n \\ &= \sum_{olmn}\epsilon_{lpo}\epsilon_{lmn}\partial_o\partial_m E^n \\ &= \sum_{omn}\left(\delta_{pm}\delta_{on}-\delta_{pn}\delta_{om}\right)\partial_o\partial_m E_n \\ &= \sum_{o}\left(\partial_o\partial_p E_o+\partial_o\partial_o E_p\right) \end{aligned} \tag{A.9}$$

where the assumption that the observed region is source-free ($rho = 0$) and $\zeta^{(\text{elec})} \neq 0$ non-vanishing and hence from A.7 the relation:

$$\begin{aligned}\sum_l \nabla_l D_l &= \sum_l \nabla_l(\zeta^{(\text{elec})}E_l) = \sum_l\left(E_l\nabla_l\zeta^{(\text{elec})}+\zeta^{(\text{elec})}\nabla_l E_l\right) \\ \sum_l \nabla_l E_l &= \sum_l E_l\frac{\nabla_l\zeta^{(\text{elec})}}{\zeta^{(\text{elec})}} = \sum_l E_l\nabla_l\ln\left[\zeta^{(\text{elec})}\right]\end{aligned} \tag{A.10}$$

was used.

Under the assumption that the magnetic permeability is spatially homogeneous ($\zeta^{(\text{magn})} = 0$) the right side of Equation A.3 yields:

$$\begin{aligned}\sum_{ol}\epsilon_{pol}\partial_o\sum_{mn}\epsilon_{lmn}\partial_m E_n &= \sum_{lo}\partial_t\left[H_l\epsilon_{pol}\left(\partial_o\zeta^{(\text{magn})}\right)+\zeta^{(\text{magn})}\epsilon_{pol}\partial_o H_l\right] \\ &= \sum_{ol}\zeta^{(\text{magn})}\epsilon_{pol}\partial_o H_l\end{aligned} \tag{A.11}$$

Assuming *charge- and current-free field-regions* and combining the A.9 and A.11 yields the wave-equation for e. g. the E_m field [16]:

$$\begin{aligned}\textstyle\sum_m \Delta_{lm}E_m &= \frac{\zeta^{(\text{elec})}\zeta^{(\text{magn})}}{c}\partial_t^2 E_l - \textstyle\sum_{mnop}\epsilon_{lmn}\left(\partial_m\ln\zeta^{(\text{magn})}\right)\left(\epsilon_{nop}\partial_o E_p\right) \\ &\quad - \textstyle\sum_m \nabla_l\left[E_m\left(\partial_m\ln\zeta^{(\text{magn})}\right)\right]\end{aligned} \tag{A.12}$$

Assuming a *monochromatic*, *time-harmonic* external field with time-dependence $\exp\{-i\omega t\}$ and $\omega = k\cdot c$, the electric field E_m (and respective H_m) can be described as:

$$E_m^{(real)}(\vec{x},\omega) = \mathcal{R}e\left\{E_m(\vec{x},\omega)\cdot e^{-i\omega t}\right\} \tag{A.13}$$

and $\mathcal{R}e$ being the real-part selector of a complex quantity. Assuming that $\zeta^{(\text{magn})}(\vec{x})$, the magnetic permeability, *is approximately constant over the spatial extend of* 1λ of the exterior field its spatial derivative $\partial_m\ln\zeta^{(\text{elec})} \propto 0$ and thus the compontens of the differential

equation Equation A.12 decouple. It suffices to do a scalar derivation for one vectorial component, because the extension for all components is similar and straightforward. While omitting the ω-dependency Equation A.12 reduces to the Helmholtz-Equation:

$$\Delta E(\vec{x}) = -k^2 n^2(\vec{x}) E(\vec{x}) \tag{A.14}$$

where $\Delta = \partial_x^2 + \partial_y^2 + \partial_z^2$ is the 2nd derivative or Laplace operator and e. g. $\partial_x^2 = \partial^2/\partial^2 x$ the Einstein notation. By splitting the scalar field $E(\vec{x}) = E^{(\mathrm{i})}(\vec{x}) + E^{(\mathrm{s})}(\vec{x})$ into the illuminating field $E^{(\mathrm{i})}(\vec{x})$ and the scattered field $E^{(\mathrm{s})}(\vec{x})$ and introducing the scalar scattering potential $F^{(\mathrm{s})}(\vec{x}) = -\frac{1}{4\pi n_0^2} k^2 \left(n^2(\vec{x}) - n_0^2\right)$ Equation A.14 can be split into a homogeneous and an inhomogeneous part:

$$\left(\Delta + k^2\right) E^{(\mathrm{i})}(\vec{x}) = 0 \tag{A.15}$$

$$\left(\Delta + k^2\right) E^{(\mathrm{s})}(\vec{x}) = -4\pi F^{(\mathrm{s})}(\vec{x}) E(\vec{x}) \tag{A.16}$$

where n_0 is the homogeneos refractive index of the surrounding medium and n the refractive index-distribution of the sample. Equation A.15 can directly be solved via the Ansatz $E^{(\mathrm{i})}(\vec{x}) = e^{i\vec{k}\vec{x}}$ and thus the Fourier-relation:

$$E^{(\mathrm{i})}(\vec{x}) = \mathcal{F}^{-1}\left\{\tilde{E}(\vec{k}')\right\}(\vec{x}) \tag{A.17}$$

$$= \int_{\mathbb{R}^3} d\vec{k}' \tilde{E}(\vec{k}') e^{i\vec{k}'\vec{x}} \tag{A.18}$$

allows to represent the analysed field via a superposition of a plane-wave spectrum which, in case of the plane-wave solution $E^{(\mathrm{i})}$, reduces to a delta-peak $\tilde{E}^{(\mathrm{i})}(\vec{k} - \vec{k}') = \delta(\vec{k} - \vec{k}')$. Only vectors $\left|\vec{k}\right| = \frac{2\pi n}{\lambda}$ that fulfill the requirement:

$$\left|\vec{k}\right| = k_x^2 + k_y^2 + k_z^2 \tag{A.19}$$

are valid solutions (wave-vectors) to Equation A.15. Within a bounded region $\vec{k}$ can be complex and thereby frequencies bigger $\left|\vec{k}\right|$ are valid, but these so called *evanescent waves* are spatially fast decaying. Hence, only waves with real-valued solutions to Equation A.19 can propagate to far distance and are thus called *propagating waves*.

The inhomogeneous Equation A.16 still contains the complete $E(\vec{x})$-field on its right side which makes an analytic solution - without any further assumptions impossible. A standard way to study the response of an analysed set of differential equations is by hitting it gently with a small hammer thereby exciting all possible harmonics within the

equations' response range. The same can achieved in case of differential equations by applying the Green's operator $G(\vec{x})$ which changes Equation A.16 to:

$$\left(\Delta + k^2\right) G(\vec{x} - \vec{x}') = -4\pi\delta(\vec{x} - \vec{x}') \tag{A.20}$$

Combining Equation A.16 and Equation A.20, applying Green's Second Identity Equation A.1 and making the particular choice Equation A.21 for the Green's function leads to Equation A.23:

$$G(\vec{x} - \vec{x}') = \frac{e^{ik|\vec{x}-\vec{x}'|}}{|\vec{x} - \vec{x}'|} \tag{A.21}$$

$$E^{(s)}(\vec{x}) = \iiint_V d^3\vec{x}'\, F(\vec{x}')E(\vec{x}')\frac{e^{ik|\vec{x}-\vec{x}'|}}{|\vec{x} - \vec{x}'|} \tag{A.22}$$

$$= [G \otimes S]\,(\vec{x}) \tag{A.23}$$

Note that $k = \frac{2\pi n}{\lambda}$ in Equation A.21 is the environment (n) dependend spatial-frequency belonging to the wavelength λ and should not be confused in general with the spatial frequency-vector $\vec{k}$, yet. In case of propagating(=non-evanescent) monochromatic fields, with $\vec{\kappa} = (\kappa_x, \kappa_y)$, the condition $|\kappa| < k$ is fulfilled and then $k = |\vec{k}|$ is correct. Even though Equation A.22 looks like a solution it is still not tractable, because the complete field $E(\vec{x}')$ is still contained on the right side of the equation. By interpreting Equation A.22 as a convolution $(\otimes)$, Equation A.23 is found and allows to understand the Green's function as a propagator of the new scattered field $S(\vec{x}) \equiv F(\vec{x})E(\vec{x})$. Instead of finding a model for the field-scattering interaction let us first take a break and analyse the propagation of fields through a given optical system.

A.3 RAYLEIGH-SOMMERFELD DIFFRACTION

Deriving the field distribution at a far point $P_1 = x^{(1)} = (\vec{\chi}^{(1)}, z^{(1)})$, with $\vec{\chi}^{(1)} = (x^{(1)}, y^{(1)})$, from a source at $P_0 = (\vec{\chi}^{(0)}, z^{(0)})$ can be achieved by applying Huygen's wavelet theory where *every wave-front* can be assumed to be the starting point of a set of new spherical waves that can mutually interfere. For a source-free region between P_0 (in plane 0) and P_1

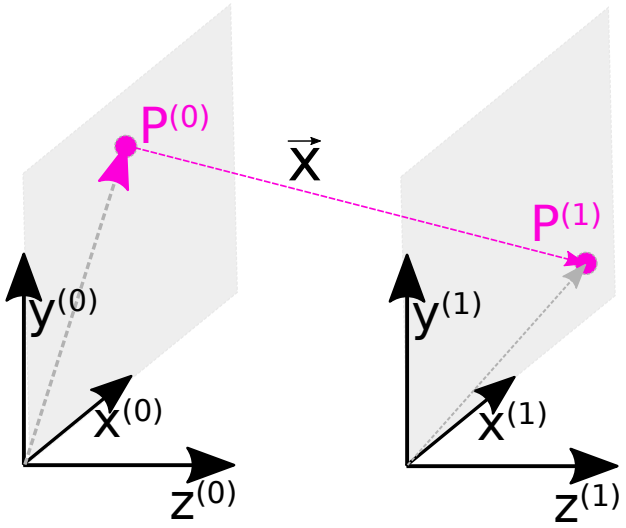


FIGURE A.1: **Coordinate Systems** used to describe the field propagation between a plane (0) with a point $P^{(0)}$ and a plane (1) with a point $P^{(1)}$.

(in plane 1) (see Figure A.1) it can be shown that the field-propagation can be described by:

$$E^{(1)}\left(\vec{x}^{(1)}\right) = E^{(1)}\left(\vec{\chi}^{(1)}, z^{(1)}\right) = \iint d^2\vec{\chi}^{(0)} D\left(\vec{\chi}^{(1)} - \vec{\chi}^{(0)}, z^{(1)} - z^{(0)}\right) E^{(0)}\left(\vec{x}^{(0)}\right) \tag{A.24}$$

$$E^{(1)}\left(\vec{\chi}^{(1)}, z^{(1)}\right) = \left[D \otimes E^{(0)}\right]\left(\vec{x}^{(1)}\right) \tag{A.25}$$

$$D\left(\vec{\chi}, z\right) = -\partial_{z^{(0)}}\left[G\left(\vec{\chi}, z\right) - G\left(\vec{\chi}, z - 2z^{(1)}\right)\right]_{z^{(0)}=0} \tag{A.26}$$

with $\vec{\chi} \equiv \vec{\chi}^{(1)} - \vec{\chi}^{(0)}$, $z \equiv z^{(1)} - z^{(0)}$ and $z^{(0)} = 0$ evaluated after the derivation. The assumptions of a *finitely bounded* field $E^{(0)}$, which is *not influenced by the bounding-edges, a homogeneous medium and a forward-propagating wave* - meaning $z^{(1)} > z^{(0)}$ were made. Coining this situation as *free-space* the Green's function Equation A.21 can be used yielding:

$$D(\vec{\chi}, z) = \left(\frac{1}{k\left|\vec{x}^{(1)} - \vec{x}^{(0)}\right|} - i\right) \frac{kz}{\pi\left|\vec{x}^{(1)} - \vec{x}^{(0)}\right|} \frac{e^{ik\left|\vec{x}^{(1)} - \vec{x}^{(0)}\right|}}{\left|\vec{x}^{(1)} - \vec{x}^{(0)}\right|} \tag{A.27}$$

$$= \left(\frac{1}{k|\vec{x}|} - i\right) \frac{kz}{\pi|\vec{x}|} \frac{e^{ik|\vec{x}|}}{|\vec{x}|} \tag{A.28}$$

with $\vec{x} = \vec{x}^{(1)} - \vec{x}^{(0)}$. In the bounding case of $k|\vec{x}| \gg (2\pi)^{-1}$ it becomes apparent that the propagator shifts the phase by $i = e^{\frac{\pi}{2}}$ which is the crucial difference of secondary sources (=fields) to primary sources (=scatterers, absorbers, emitters).

Assuming only radiating fields (=being far away from the sample), allowing only small lateral devitions from the optical axis ($|\vec{x}| \to z$) and applying the Taylor-expansion around $\vec{\chi}_0 = 0$:

$$|\vec{x}| = \sqrt{z^2 + \vec{\chi}^2} = z\sqrt{1 + \left(\frac{\vec{\chi}}{z}\right)^2} \propto z\left(1 + \frac{\vec{\chi}^2}{2z^2}\right) \tag{A.29}$$

for the $|\vec{x}|$ in the exponential leads to the Fresnel-Propagator $D^{(\mathrm{F})}$:

$$D^{(\mathrm{F})}(\vec{\chi}, z) = -i\frac{k}{2\pi z}e^{ikz}e^{i\frac{k}{2z}\vec{\chi}^2} \tag{A.30}$$

Inserting Equation A.30 in Equation A.24 finally yields the Fresnel-approximation:

$$E^{(1)}\left(\vec{\chi}^{(1)}, z^{(1)}\right) = D^{(F')}\left(\vec{\chi}^{(1)}, z\right)\mathcal{F}_{\vec{\chi}}\left\{E^{(0)'}\left(\frac{z\vec{\chi}^{(0)'}}{k}, z^{(0)}\right)\right\}\left(\vec{\chi}^{(1)}, z^{(0)}\right) \tag{A.31}$$

$$D^{(F')}\left(\vec{\chi}^{(1)}, z\right) \equiv -ie^{ikz}e^{i\frac{k}{2z}\vec{\chi}^{(1)2}} \tag{A.32}$$

$$E^{(0)'}\left(\frac{z\vec{\chi}^{(0)'}}{k}, z^{(0)}\right) \equiv E^{(0)}\left(\frac{z\vec{\chi}^{(0)'}}{k}, z^{(0)}\right)e^{-i\frac{k}{2z}\vec{\chi}^{(0)2}} \tag{A.33}$$

$$\tag{A.34}$$

where the coordinate-rescaling $\vec{\chi}^{(0)'} \equiv \frac{\vec{\chi}^{(0)} - \vec{\chi}^{(1)}}{z}$ under the integral was used for reduction. Further $\mathcal{F}_{\vec{\chi}}$ is $2D$ Fourier-Transform with respect to χ. Interestingly, instead of the typical Fourier-component representation of the field in terms of k-frequencies the field at $\vec{\chi}^{(1)}$ rather represents the field composed of spatially distributed plane waves. Instead of having to invoke two Fourier-transforms as would be the case for the free-space propagator Equation A.35:

$$D^{(\mathrm{free})}(z) = e^{ik_z z} = \begin{cases} e^{iz\sqrt{k^2-\kappa^2}} & , |\kappa| < k \\ 0 & , \text{else} \end{cases} \tag{A.35}$$

the Fresnel-Propagator only needs a single transformation.

The Fraunhofer approximation $D^{(\mathrm{FH})}$ Equation A.36 can be derived from Equation A.24 by limiting $E^{(0)}$ to a maximum radius $\chi^{(0,max)}$ and assuming very far distances such that $|\vec{x}| \gg k|\vec{\chi}^{(0)}|^2$, hence yielding:

$$E^{(1)}\left(\vec{\chi}^{(1)}, z^{(1)}\right) = D^{(\mathrm{FH})}\left(\vec{\chi}^{(1)}, z\right)\mathcal{F}_{\vec{\chi}}\left\{E^{(0)'}\left(\frac{z\vec{\chi}^{(0)'}}{k}, z^{(0)}\right)\right\}\left(\vec{\chi}^{(1)}, z^{(0)}\right) \tag{A.36}$$

$$D^{(FH)}\left(\vec{\chi}^{(1)}, z\right) \equiv -\frac{iz}{x^{(1)}}e^{ikz} \tag{A.37}$$

A.4 LENS, 2F & 4F

With these tools in hand the field propagation through an optical system can be calculated as follows - for ease of explanation assuming 2D-fields only:

1. propagate field (plane) from $z^{(0)}$ a distance $s^{(0)} = z^{(1)} - z^{(2)}$ to the entrance pupil at $z^{(2)}$ of the optical system
2. have the field interact with the optical system
3. propagate field from the exit-pupil $z^{(2)'}$ a distance $s^{(1)} = z^{(1)} - z^{(2)'}$ to the plane of measurement at $z^{(1)}$

The simplest configuration in case of free-space propagation (where there is no system in-between) was demonstrated in the former section and is coherent with this approach. An object of particular interest is the propagation through a lens. For a thin lens, the lens-curvature is small compared to the lens diameter and the field's lateral entry and exit-coordinates do not change $\vec{\chi}^{(1)} \approx \vec{\chi}^{(1)'}$. Hence it can be described with a complex-valued pupil function $P(\vec{\chi})$, f the focal length and $\Phi = k\sqrt{f^2 + \chi^{(1)2}}$ the locally introduced phase change dependend on the lens-shape and material. The transmission through a thin lens is then given in Equation A.38 and by Taylor-expanding Φ only until the 1st order Equation A.39 can be derived:

$$E^{(1)'}\left(\vec{x}^{(1)}\right) = P\left(\vec{\chi}^{(1)}\right) e^{-i\Phi\left(\vec{\chi}^{(1)}\right)} E^{(1)}\left(\vec{x}^{(1)}\right) \tag{A.38}$$

$$= \mathcal{F}_{2D}^{-1}\left\{\tilde{P}\right\}\left(\vec{\chi}^{(1)'}\right) e^{-i\frac{k}{2f}\vec{\chi}^{(1)}} E^{(1)}\left(\vec{x}^{(1)}\right) \tag{A.39}$$

with $P\left(\vec{\chi}^{(1)}\right) e^{-i\Phi\left(\vec{\chi}^{(1)}\right)} e^{-i\frac{k}{2f}} = \mathcal{F}_{2D}^{-1}\left\{\tilde{P}\right\}\left(\vec{\chi}^{(1)'}\right)$. To calculate the propagation of a field from $-z^{(0)}$ before the lens until a distance $z^{(1)}$, hence propagating the field by a total distance of $z = z^{(1)} + z^{(0)}$ the above mentioned algorithm is applied and evaluated in case of Fresnel-approximation thereby yielding:

$$E^{(1)}\left(\vec{\chi}^{(1)}, z^{(1)}\right) = D^{(\text{lens})}\left(\vec{\chi}^{(1)}, s^{(1)}, s^{(0)}\right) \int d^2\vec{\chi}^{(0)} \tilde{P}\left(\frac{k^{(1)}}{s^{(1)}}\vec{\chi}^{(1)} + \frac{k^{(0)}}{s^{(0)}}\vec{\chi}^{(0)}\right) E^{(0)'}\left(\vec{\chi}^{(0)}, 0\right) \tag{A.40}$$

Here, different immersion media in front $n^{(0)}$ and after the lens $n^{(1)}$ were taken into account which led to different $k^{(0)}$ and $k^{(1)}$. The replacements

$$D^{(\text{lens})}(\vec{\chi}^{(1)}, s^{(1)}, s^{(0)}) = -\frac{k^{(0)}}{s^{(0)}} e^{i\left(k^{(1)}s^{(1)} + k^{(0)}s^{(0)}\right)} e^{i\frac{k^{(1)}}{s^{(1)}}\chi^{(1)2}} \tag{A.41}$$

$$E^{(0)'}(\vec{\chi}^{(0)}, 0) = e^{i\frac{k^{(0)}}{2s^{(0)}}\chi^{(0)2}} E^{(0)}(\vec{\chi}^{(0)}, 0) \tag{A.42}$$

were used for simplification and to introduce the lens-propagator $D^{(\text{lens})}$ to propagate from one plane at $s^{(0)}$ to a plane at $s^{(1)}$. Adding a second lens with focal length $f^{(1)}$ ($f^{(0)}$ is the focal length of the first lens), placing the two lenses such that their focal-planes coincide, and placing the input field at the front-focal plane of the first lens, a 4f-configuration is constructed. Hence a second propagator over the distance $s^{(1)} = 2f^{(1)}$ is added to the first lens with the propagation distance $s^{(0)} = 2f^{(1)}$. Hence, applying Equation A.40 again for the second lens leads to:

$$E^{(1)}(\vec{\chi}^{(1)}, 2f^{(1)} + 2f^{(0)}) = D^{(4f)}(f^{(0)}, f^{(1)}) \int d^2\vec{\chi}^{(0)} \tilde{P}\left(\frac{k^{(1)}}{f^{(1)}}\vec{\chi}^{(1)} + \frac{k^{(0)}}{f^{(0)}}\vec{\chi}^{(0)}\right) E^{(0)}(\vec{\chi}^{(0)}, 0) \tag{A.43}$$

with the 4f propagator $D^{(4f)}$ as:

$$D^{(4f)}(f^{(0)}, f^{(1)}) = -\frac{k^{(0)}k^{(1)}}{4\pi f^{(0)}f^{(1)}} e^{i\left(k^{(0)}f^{(0)} + k^{(2)}f^{(0)'} + k^{(2)}f^{(1)'}k^{(1)}f^{(2)}\right)} \tag{A.44}$$

Equation A.43 has the typical shape of a homogeneous Fredholm-integral equation of 1st kind with the kernel $\tilde{P}$, the unknown function $E^{(0)}$ und the measurable/known function $E^{(1)}$ [71]. It describes the system-response (kernel) to an incoming signal. By rescaling the pupil-function $\tilde{P}$ according to Equation A.45 the integral can be simplified into a convolution Equation A.46:

$$a(\vec{\chi}) = \left(\frac{k^{(0)}}{2\pi f^{(0)}}\right)^2 \tilde{P}\left(-\frac{k^{(0)}}{f^{(0)}}\vec{\chi}\right) \tag{A.45}$$

$$E^{(1)}\left(\vec{\chi}^{(1)}\right) = \frac{1}{M^2} D^{(4f)}(f^{(0)}, f^{(1)}) \left[a \otimes E^{(0)}\right]\left(\frac{\vec{\chi}^{(1)}}{M}\right) \tag{A.46}$$

where $M = -(n^{(0)}f^{(1)})/(n^{(1)}f^{(0)})$ is the system magnification and a the 2D Amplitude Point Spread Function (APSF) of the analysed 4f-system. This equation can be simplified by assuming that all three different zones consist of the same surrounding medium n and that the front and back-focal planes of each lens are identical. Further, if the Abbe-Sine condition is fulfilled (=constant magnification over whole FOV) *invariance under translation*

of the system-response function can be approximately assumed. Note that the derivation discusses plane-to-plane trasport and hence 2D in-plane solutions are derived. The full 3D-solutions around the planes of interest/focus can be calculated from the in-plane solutions via different propagation approaches and will be introduced in Section A.6.

Hence we saw that a lens is not only a fast Fourier-transformator (Equation A.31), but that the field at a displaced position $z^{(1)}$ can be described as a convolution of the incoming field at $z^{(0)}$ with the APSF a (Equation A.46).

A.5 COHERENT AND INCOHERENT IMAGING

In the derivation of the Helmholtz-Equation (Equation A.14) the time-harmonic $e^{i\omega t}$ term was omitted and shall be viewed again. All system (e. g. 4f-configuration) responses are assumed to be *static* while the incoming and outgoing fields are allowed to change with time. A intensity measurement of the incoming light-field means to temporal average it for the measurement duration $\Delta t = T$ due to the limitations of detectors. By adding time t as variable to Equation A.46 it follows:

$$I^{(1)}(\vec{\chi}) = \left\langle \left| E^{(1)}(\vec{\chi}, t) \right|^2 \right\rangle = \frac{1}{T} \int_t^{t+T} dt' E^{(1)}(\vec{\chi}, t' - \tau') E^{1*}(\vec{\chi}, t' - \tau'') \tag{A.47}$$

$$= \iint d^2\vec{\chi}' d^2\vec{\chi}'' a\left(\vec{\chi} - \vec{\chi}'\right) a^*\left(\vec{\chi} - \vec{\chi}''\right) \cdot \underbrace{\frac{1}{T} \int_t^{t+T} dt' E^{(0)}(\vec{\chi}', t' - \tau') E^{(0)*}(\vec{\chi}'', t' - \tau'')}_{\equiv J(\vec{\chi}', \vec{\chi}'') \ldots \text{mutual intensity}} \tag{A.48}$$

Note that $t - \tau$ marks the time delay necessary to propagate through the optical system characterized by $a(\vec{\chi})$. The introduced mutual coherence is a measure of spatial coherence of the propagated field. The behaviour of J can further be described by limiting it to the two cases that will be used within this thesis, namely (1) quasi-monochromatic ($\Delta\lambda \ll \lambda$, but not $\Delta\lambda = 0$) and (2) incoherent fields E. In case of (1) the field is completely self-correlated and distant points within the field still have a phase-relation with each other and hence still can interfer. In case (2) on the other hand no two points of the field have a continuous phase-relation with each other and hence only self-interference of a point is possible thereby yielding:

$$(1) \quad J^{(coh)}(\vec{\chi}', \vec{\chi}'') = E^{(0)}(\vec{\chi}') E^{(0)*}(\vec{\chi}'') \tag{A.49}$$

$$(2) \quad J^{(inc)}(\vec{\chi}', \vec{\chi}'') = I^{(0)}(\vec{\chi}') \delta^2(\vec{\chi}' - \vec{\chi}'') \tag{A.50}$$

Inserting these mutual intensity functions into Equation A.47 the relations

$$(1) \quad I^{(1)}(\vec{\chi}) = \left|\left[a \otimes E^{(0)}\right](\vec{\chi})\right|^2 \tag{A.51}$$

$$(2) \quad I^{(1)}(\vec{\chi}) = \left[|a|^2 \otimes I^{(0)}\right](\vec{\chi}) \tag{A.52}$$

can be found. Remarkably, a coherent field can be propagated with a coherent propagator, namely the a or APSF, while an intensity distribution (no phase information left) is propagated by an intensity-propagator h, the so called Point Spread Function (PSF). By Fourier-transforming these operators the Amplitude Transfer Function (ATF) $\tilde{a}$ (Equation A.54) and the Optical Transfer Function (OTF) $\tilde{h}$ can be found.

$$h(\vec{\chi}) = \left|a(\vec{\chi})\right|^2 \tag{A.53}$$

$$\tilde{a}(\vec{\kappa}) = \mathcal{F}_{(2D)}\{a(\vec{\chi})\}(\kappa) \tag{A.54}$$

$$\tilde{h}(\vec{\kappa}) = \mathcal{F}_{(2D)}\{h(\vec{\chi})\}(\kappa) \tag{A.55}$$

Already at this state it can be seen that retrieving the PSF in case of incoherent imaging

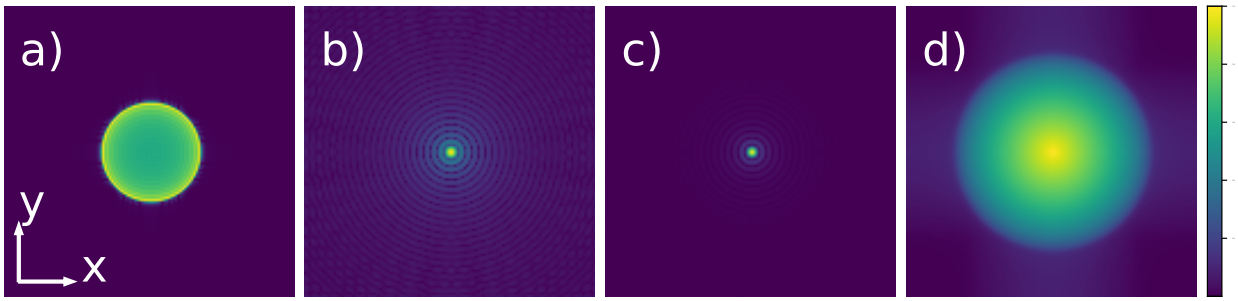


FIGURE A.2: **Fourier-Relations in 2D.** a) ATF, b) APSF, c) PSF, d) OTF.

(Equation A.52) can be achieved by e. g. imaging a sparse set of infinitesimal small points (beads, quantum dots, gold nanoparticles) that can be described as δ-combs for $I^{(0)}$ and averaging over the resulting shifted PSFs according to Equation A.52. In case of the APSF (Equation A.51) regaining the necessary phase-information needs many assumptions about the nature of the system and the incoming field as well. Yet it can be reconstructed by using an iterative phase-retrieval approach, if possible at all.

A.6 3D-IMAGING

Intuitively, a simple but reasonable strategy for extending a given 2D-model to 3D is to assume that in case of the 2D-model the *in-focus sample-plane* (typically $z^{(0)} = 0$) is imaged onto the detector plane. Then by adding a free-space propagation either to the sample or detector in $\pm z$ direction the *in-focus* solution can be extended to Z-planes $z \neq z^{(0)}$.

According to the presented strategy for the transition from 2D to 3D a defocus-operation in terms of convolution of the 2D APSF with the Rayleigh Sommerfeld Field Propagator D (Equation A.28) directly leads to the 3D APSF $a\left(\vec{\chi}, z\right) \equiv a\left(\vec{x}\right)$ Equation A.56, where $\vec{\chi}$ is the lateral 2D and $\vec{x}$ the 3D coordinate vector. In case of Fourier space representation the ATF $\tilde{a}\left(\vec{\kappa}\right)$ is multiplied by the Fourier transformed Rayleigh Sommerfeld Field Propagator $\widetilde{D}$ thereby yielding the 3D ATF $\tilde{a}\left(\vec{k}\right)$ Equation A.57, where $\vec{\kappa}$ is the lateral 2D and $\vec{k}$ the 3D frequency vector.

$$a\left(\vec{x}\right) = \left[D\left(\vec{\chi}, z\right) \otimes_{2D} a\left(\vec{\chi}, 0\right)\right] \tag{A.56}$$

$$\begin{aligned}\tilde{a}\left(\vec{k}\right) &= \mathcal{F}_{\vec{\chi}}\left\{a\right\} = \widetilde{D}\left(\vec{\kappa}, z\right)\tilde{a}\left(\vec{\kappa}, 0\right) \\ &= e^{iz\sqrt{k^2-\kappa^2}}\tilde{a}\left(\vec{\kappa}, 0\right)\end{aligned} \tag{A.57}$$

In Equation A.57 the particular choice $D = D^{(\text{free})}$ (Equation A.35) was used in Equation A.57. The ATF $\tilde{a}\left(\vec{x}\right)$ can be interpreted as a 3D imaging pupil. It is crucial to bear in mind that according to the presented strategy the defocus was solely modeled into the APSF a which in turn is assumed to *plane-wise* interact with the sample. *mutual independence of different planes of the sample* by means of coherence and absorption/emission is assumed. In case of coherent imaging this assumption is typically heavily violated as will be discussed further in Section A.8, but can be used for description of incoherent processes with great success. For example, the incoherent 3D-image of a 3D-sample in the detector-plane can be described via:

$$I^{(1)}\left(\vec{\chi}^{(1)}, 0\right) = h\left(\vec{\chi}^{(1)}, +z^{(1)}\right) \otimes_{2D} I^{(0)}\left(\vec{\chi}^{(0)}, 0\right) \equiv \left[h \otimes_{2D} I^{(0)}\right]\left(\vec{\chi}^{(1)}, +z^{(1)}\right) \tag{A.58}$$

where $I^{(0)}$ is a 2D intensity distribution at plane 0 at $z^{(0)} = 0$ and accordingly $I^{(1)}$ is a 2D intensity distribution at plane 1 at $z^{(1)} = z$, see Figure A.1. The points $P^{(0)}$ and $P^{(1)}$ in Figure A.1 demonstrate how a point is not only free-space propagated but also displaced

Topic	Approximation
Helmholtz-Equation	monochromatic, time-harmonic (=sinusoidal) external field/waves; very slowly moving bodies as compared to the time-window of interaction with surrounding fields; isotropic material; specific conductivity, electric permittivity and magnetic permeability are constants; charge- and current-free field-regions
in-focus	finitely bounded; no influence by the bounding-edges; homogeneous medium over at least 1λ spatial extent; forward-propagating waves; APSF invariance under translation; system responses are static over time, while in-/out-going fields can change; APSF plane-wise interacts with sample; mutual independence of different sample-planes
measurement	far-away distance; finite temporal resolution of detector

TABLE A.2: **Approximations.** Brief overview of all applied approximations for the derivation of the 3D-APSF, ATF, PSF and OTF.

when going from one plane to another through the optical system. The identity ($\equiv$) displays an alternative way of writing the functional dependencies of the first equality of Equation A.58 in a more loose but easier to read manner.

The result is remarkable, as it states that if the sample can be described as a 3D-stack of individual axially-displaced planes (slices) and such a stack can be acquired by (e. g. successive) 3D-measurements. Equation A.58 generalizes to Equation 1.9.

A.7 SCALAR HIGH-*NA* IMAGING

In the last sections the basic foundation for light-propagation and optical systems was set. The concept of a system-transfer function was introduced by using sequential application of the Rayleigh-Sommerfeld integral or its approximations (mainly: Fresnel). Especially a (semi-) convex lens in 2f [1][2] or 4f [3] configuration generates the very far away Fraunhofer-pattern, which can be described as a Fourier-Transform of the incoming front-focal plane

1 i. e. 2f before and 2f after one lens

2 Only directly true if Abbe-sine condition is fullfilled.

3 i. e. 2 subsequent lenses with matching focal planes and imaging from first to last focal plane

field-distribution (Fraunhofer-approximation of the Rayleigh-Sommerfeld diffraction integral), into its back-focal (Fresnel-Approximation) plane (Equation A.31). Note that this is only true for imaging of the intensities in both lens configurations, but does not hold for the phases in case of the 2f configuration[4]. While various approximations have been made to derive how a field propagates through a given system (see Table A.2), the question remains: How to properly model such a system without the particular assumptions of *paraxiality*[5] or *measuring at a far away distance*? In Equation A.24 the incoming 3D-field distribution $E^{(0)}(\vec{x})$ is projected onto a superposition of 2D-wavefronts that reaches the entrance pupil of the optical system at $z^{(0)}$ and is then propagated to a distant plane at $z^{(1)}$. Here the assumption of a *finite sized pupil* and effects of the pupil onto the incoming field (as introduced by Equation A.38) can directly be included into the pupil function P:

$$E^{(1)}\left(\vec{x}^{(1)}\right) = \iint d^2\vec{\chi}^{(0)} P(\vec{\chi}^{(0)}) D\left(\vec{\chi}^{(1)} - \vec{\chi}^{(0)}, z^{(1)} - z^{(0)}\right) E^{(0)}\left(\vec{x}^{(0)}\right) \tag{A.59}$$

Equation A.59 is still exact within the framework of the Sommerfeld-Rayleigh integral for a (locally) *homogeneous surrounding medium*. Assuming a *uniform circular pupil* in *paraxial approximation* yields:

$$P\left(\vec{\chi}^{(0)}\right) = \begin{cases} 1 & ,\chi^2 < NA \\ 0 & , \text{else} \end{cases} \tag{A.60}$$

and a uniform incoming light distribution ($E^{(0)}(\vec{\chi}^{(0)}) = 1$, e. g. a plane wave-front $z \gg \lambda$ away from a δ-like emitter) Equation A.59 can be solved to the *jinc* function [70]:

$$E^{(1)}\left(\vec{\chi}^{(1)}, z^{(1)} = 0\right) = \frac{J^{[1]}\left(\vec{\chi}\,\omega\right)}{\vec{\chi}\,\omega} \equiv \text{jinc}\left(\vec{\chi}\,\omega\right) \tag{A.61}$$

with $J^{[1]}$ being the Bessel-Function of 1st kind and $\omega \equiv kNA$. Note that $J^{[1]}$ is a periodic function whose period is $\omega^{(-1)}$. Equation A.60 is the in-focus solution (at position $z^{(1)} = 0$) for an optical system whose entrance- and exit-pupil can altogether be described with a non-apodized, hard edged circular pupil. The full 3D-field distribution $E^{(1)}\left(x^{(1)}\right)$ around the focus can be calculated by applying the free-space propagator Equation A.35.

When stepping away from the paraxial approximation, hence allowing for bigger lateral angles of the illumination/detection beam and thus $\vec{\chi}^{(0)} \ll z$ is no longer valid, polarization effects on the resulting 3D-vector field in-focus distribution become noticeable. Within

4 e. g. imagine the image of a parallel incoming wave in 2f configuration.

5 *Paraxiality* is typically a wording used in geometrically optics where the light propagation is characterized by using propagating rays. In case of *paraxiality* only small angles between the rays and the optical axis are allowed, thus allowing for the approximation $\sin\alpha \approx \alpha$.

the general framework of super-resolution microscopy many systems rely on high-*NA* objectives with 1.4*NA* and e. g. oil-immersion $n^{(\mathrm{oil})} = 1.518$. Here the half aperture angle of the objective lens is $\alpha = \arcsin\left(\frac{NA}{n^{(\mathrm{oil})}}\right) \approx 67°$. For example, there will be two opposing spatial frequencies[6] in the objective lens' Back Focal Plane (BFP) at $\pm 0.42 \cdot d^{(\mathrm{BFP})}$, where $d^{(\mathrm{BFP})}$ is the objective lens' BFP diameter, for which the focus of an incoming plane-wave meets under an angle of $\theta^{(\mathrm{crit})} = 90°$. These two beams cannot interfere due to a missing spatial overlap. Therefore, the degree and spatial extend of interference of the interfering beams - coming from opposing sides of the optical axis - depends on their lateral position and thus the focal spot gets deformed.

Even though this was only a qualitative explanation it displays the potentially severe discrepancy of using high-*NA* objectives together with the paraxial approximation. The linearity of the Helmholtz equation (Equation A.14) allows to treat the Laplace operator as separable and hence only limit the analysis to 1 field-direction at a time. In this thesis the high-*NA* scalar case is considered, but transition to vectorial treatment is straightforward and could be added at anytime if necessary.

A.8 BORN APPROXIMATION

In Section A.6 field propagation, based on the homogeneous Helmholtz equation Equation A.15 led to the derivation of Fourier-imaging theory for arbitrary imaging systems. The inhomogeneous Helmholtz equation Equation A.16 in case of interaction with coherently scattering/absorbing matter will briefly be analysed.

Since the scattered field $E^{(\mathrm{s})}$ depends directly on the complete field E in Equation A.16 an analytical solution leads to an infinite (self-recursive) series of E and thereby needs additional assumptions for tractability. Without much further restriction and for arbitrary matter distributions, direct Maxwell solvers can compute the matter-field interactions in a small volume of space enclosing the matter distribution. Typical Finite Difference Time Domain (FDTD) algorithms discretize space and time and thus can solve Equation A.16 on a discrete grid for a large frequency range for virtually arbitrary matter-field interactions (e. g. nonlinear media) simultaneously, but require large computational capacities and time [72]. Since the field is computed on all points in space, wrap-around artifacts can

6 This even holds true for a whole ring around the zero frequency.

arise due to the finite domain size (computer memory limited), which can be reduced by introducing a perfectly absorbing boundary surface. The Discrete Dipole Approximation (DDA), on the other hand, assumes a distribution composed of individual dipoles in a constant medium instead of an arbitrary sample distribution and solves Equation A.16 in the frequency domain for 1 frequency at a time. The field distribution calculated in the medium composed of dipoles can then be propagated into the external space and the near and far field can be derived from it. This method is already significantly more resource-efficient, but also entails significantly greater restrictions with respect to material properties and structure [72]. Note that both FDTD and DDA can achieve the same performance depending on the calculation parameters [73]. Further developments for the efficient calculation of the field distribution around the scattering medium could be achieved by the suitable discretization of the Green's function [74].

Another possibility is to divide the medium into individual layers, calculate the field distribution per layer $E^{(\mathrm{l})}$ with the result of the propagated field distribution from the previous layer as an incoming wave $E^{(\mathrm{l})} = D(dz) \otimes E^{(\mathrm{l\text{-}1})}$. The method is susceptible to strong refractive index changes at interfaces, backscattering, and an appropriate choice of propagation distance dz between layers.

Assuming that the matter distribution is only *weakly scattering* that is, $n(x) - n_0 \approx \delta n$ holds for very small δn, $E \approx E^{(\mathrm{i})}$ can be assumed for the right-hand side of Equation A.23 [16] and it follows:

$$E^{(\mathrm{s})}(\vec{x}) = -4\pi \left[G \otimes \left(F^{(\mathrm{s})} E^{(\mathrm{i})} \right) \right] (\vec{x}) \tag{A.62}$$

with the scalar scattering potential $F^{(\mathrm{s})}(\vec{x}) = -\frac{1}{4\pi n_0^2} k^2 \left(n^2(\vec{x}) - n_0^2 \right)$ (see Equation A.16) and $\otimes$ being the 3D-convolution. The total field $E(\vec{x}) = E^{(\mathrm{i})}(\vec{x}) + E^{(\mathrm{s})}(\vec{x})$ is the sum of the ballistic field $E^{(\mathrm{i})}$ and the scattered field $E^{(\mathrm{s})}$ and can be calculated directly. With application of the *far-field approximation* the interaction with the medium can be directly incorporated into the established Fourier imaging framework Equation 1.9, since the weak scattering property allows individual thin slices to be considered as mutually independent.

A.9 THE JABLONSKI DIAGRAM

The possible excitation states of an atom can only be described exactly in the case of hydrogen by means of the harmonic oscillator approach. Atoms with higher atomic numbers can only be described approximately. For molecules, the harmonic oscillator approach is not realistic for several reasons. On the one hand, bonds cannot break despite arbitrarily large energies, and on the other hand, two atomic nuclei can come too close to each other under very small energetic expenses [28]. The introduction of the anharmonic oscillator solves this problem. Restricting the solutions of the anharmonic oscillator to zero and first order, we obtain the Morse potential [28]. For a diatomic molecule, the two first electronic states (S_0, S_1) have a different center of charge, i. e. the central point of symmetry of their spatial charge distribution differs between S_0 and S_1. The lower level S_0 is called Highest Occupied Molecular Orbital (HOMO) and the upper level S_1 Lowest Unoccupied Orbital (LUMO). If the total spin of the excited and the paired electron remaining unexcited in the lower level 0, i.e. they are oriented in opposite directions, a singlet state exists. In this state, the spatial overlap of the electron wave functions is particularly large, which means that more energy must be expended for excitation than in configurations with a smaller overlap, such as the triplet state. Here, the two electrons have the same spin orientation, which separates their wave function spatially (fermion property). The triplet state is more stable and due to the selection rule of emission generally longer-lived than the singlet state [75]. Due to the approximately 1000x greater weight of the nucleus compared to the electrons, the Born-Oppenheimer approximation can be used and thus the wave function of the nucleus can be described separately from the electrons. Thus, the change of the electronic state can be considered instantaneous with respect to the motion of the nucleus (depicted as a straight arrow in Figure 1.2). According to the Franck-Condon principle, transitions with large overlap of their vibronic wavefunctions are more likely than those with smaller overlap [27].

A set of time scales for transitions within a molecule in case of photon absorption and emission are summarized in Table A.3.

Transition	**t** order
Decoherence-Time (a)	fs
Vibrational Relaxation (VR)	ps
Internal Conversion (IC)	ps
Fluorescence (f)	ns
Intersystem Crossing (ISC)	ns
Phosphorescence (p)	ms

TABLE A.3: Time-scale overview for relevant transitions displayed in Figure 1.2 [27].

A.10 DETECTORS

In this work, three main detector technologies were used, and their operation and general parameters are briefly described below.

PMT In a Photo-Multiplying Tube (PMT) incoming photons are converted into Electron (e^-) at the (e. g. Gallium-Arsenide Phosphate (GaAsP)-) entrance-window and then accelerated by the surrounding static E-field towards a set of concave dynodes. By hitting the dynodes each time more e^- are emitted by secondary emission thereby leading to an overall amplification of up to 10^7 for each incoming photon. The created current can than be measured or converted into voltage, depending on the Analog-Digital-Conversion Unit (ADC). While multiple-parameters influence the tube notably, especially the entry window material, the amount of dynodes and the potential between them are of particular interest. Changing the latter can be used to optimize the amplification while minimizing noise-gain, i. e. minimizing the chance of amplifying electron-emission events that where not induced by incoming photons. On the other hand, changing the window-material allows to optimize for response-time or wavelength-dependend sensitivity. Further undesired effects are split into additive noise, where the Signal-to-Noise Ratio (SNR) changes proportional to the photon-influx, or multiplicative noise, where SNR is independend of a change in photon-influx. Dark-current, an additive noise-source, happens due to e. g. after-pulsing (e. g. light emission by dynodes), field-emission (due to high-fields between dynodes), thermal-emission (spontaneous emission of a valence-electron due to temperature depending potential energy) or radioactive decay (e. g. of the glass material). While the average effect of this noise-source can be diminished by simple subtraction

the uncertainty of the measurement is increased. Avalanche noise, a multiplicative noise-source on the other hand, occurs during carrier and lattice interaction at the dynodes where, if energy of the hitting carrier is high enough, more second carriers than intended (according to dynode design) are generated thereby leading to a bigger amplification and hence variance in the measurement. Due to geometry and material, the relative gain by electron multiplication is subject to variations which increase the relative inaccuracy (= variance) of the measurement current prediction per photon measurement. Finally, before converting into digital units by using an ADC, the e^- are converted to voltage. Even though the additive conversion variance (Johnson noise) scales anti-proportionally with the used impedance, too high impedances should not be used as they might influence the device response time together with the system inherent capacitances. The typical quantum efficiency of a GaAsP-PMT around $\lambda = 540\,\text{nm}$ is $q^{(\text{PMT})} \approx 40\%$ where measurements at 20 ns (electron transit time) pixel dwell-time are possible in the case of linear-focusing design and continuous (non-counting) mode [76].

(em)CCD A Charge-Coupled Device (CCD) consists of multiple detectors arranged on a rectangular grid. There are photon conversation pixels and read-out registers. These can vary in number and position depending on the type of device (e. g. full-frame, frame-transfer, interline). Incoming photons are converted into e^- per pixel and shifted to the read-out register after the end of the exposure time. Depending on the type, the pixels can be exposed further during the transfer, which can result in smearing. The accumulated charge per pixel is converted pixel by pixel into voltage via the readout register, amplified and converted into discrete units by ADC. Overexposure, i. e. utilization of more than the available dynamic range of the pixels, can lead to e^- jumping over to adjacent pixels' wells which after readout can be seen as blooming, i. e. brighter regions with lower contrast [77]. In the case of an electron-multipliying Charge-Coupled Device (emCCD), electron multiplication due to avalanche ionization is added before the voltage conversion. emCCD cameras are ideal for high contrast (= ==low or no background) imaging at moderate imaging speeds, e. g. 30 frames per second (fps), due to their read-noise of up to $< 1\,e^-/pix$ as Root-Mean-Square (RMS) of the read-noise distribution . For slow imaging and existing background, long integration times are necessary, causing dark-noise to accumulate. Due to the low dark-noise level of CCDs compared to emCCDs or Scientific-

Metaloxide Semiconductor (sCMOS), they are preferred here [14]. In back-illuminated configuration, (em)CCDs achieve a quantum efficiency over 90%. Common emCCDs have 1024 × 1024 pixels with a pixel-pitch of 13 μm ×13 μm [78].

sCMOS In a sCMOS an incoming e^- creates an electron-hole pair in the depletion region of p-n-transistor which get amplified and accumulated within capacitors directly on the individual pixels. The pixel values are thus pre-amplified, which can significantly reduce the relative read noise arising at the horizontal read-out registers (row). sCMOS achieve down to RMS 1.6 e^- (at 100 fps) at 0.006 e^-/pix at −30 °C cooling [79]. The presence of a microlens array (=1 lens per pixel) can increase quantum efficiency allowing QEs of up to $\geq$ 80% to be achieved. Hence, sCMOS are ideal for fast measurements with background present and weak ($\geq$ 4 $photons/pixel$) signals. Typical detectors have about 2048 × 2048 pixels at 6.5 μm ×6.5 μm pixel pitch [80].

A.11 SHOT NOISE

The quantum nature of light leads to fluctuations in the discrete conversion of photon to electron during a detection process. The measured photons during an detection interval T are thus a sequence of discrete-time events $\delta^{(\mathrm{tl})} \equiv \delta(t - t_l)$, whose temporal resolution depends on the temporal response of the detector $R^{(\mathrm{T})}(t)$. A random variable K, which equals the amount of photons detected in a time interval T:

$$K = \int_T dt \sum_l \left[R^{(\mathrm{T})} \otimes \delta^{(\mathrm{tl})} \right](t) \tag{A.63}$$

can be described by its statistical (=time-averaged) mean $\mu_K \equiv \mathbb{E}\{K\}$ and variance $\sigma_K^2 \equiv \mathbb{V}\{K\}$ according to [14]:

$$K = \mu_K + \sigma_K^2 \tag{A.64}$$

The measurement process can also be formulated as asking the question: „What is the probability that k photons were measured in the time interval $\Delta t = T$?“. The Poisson statistic $\mathcal{P}$ answers this question:

$$\mathcal{P}\{\mu_K\} \equiv P_K(K = k|\mu_K) = \frac{(\mu_K)^k}{k!} e^{-\mu_K} \tag{A.65}$$

with $K = k$ realizations, i. e. measured photons, of the Poisson distributed random variable K. The random variable K is discrete (i. e. realizations $k \in \mathbb{N}^+$) and thus the continuous mean $\mu_K \in \mathbb{R}^+$ can be obtained by using its discrete probability distribution $P_K(K = k|\mu_K)$:

$$\mu_K = \sum_k k P_K(K = k|\mu_K) \tag{A.66}$$

where the formulation $P_K(K = k|\mu_K)$ means: the probability of measuring $K = k > 0$ photons given the mean $\mu > 0$ of the random process takes the value μ_K. The events measurable at the detector depend on both its temporal response $R^{(\mathrm{T})}$ (Equation A.63) and its physical properties, such as e. g. the absorption cross section $\eta^{(\mathrm{D})}$. Thus, the photon-flux $\Phi(t)$ of the incoming Electro-Magnetic (EM) waves measurable by the detector can be described by $\Phi^{(\mathrm{D})}(t) \propto \eta^{(\mathrm{D})}\Phi(t)$ and hence what can be measured by the detector within T is:

$$\Phi^{(\mathrm{T})}(t) = \left[R^{(\mathrm{T})} \otimes \Phi^{(\mathrm{D})}\right](t) \tag{A.67}$$

Most sources (e. g. laser) are subject to temporal fluctuations that last longer than the time scale T relevant for the measurement process at the detector. The counts K are not independent of the measurable flux $\Phi^{(\mathrm{D})}$ at the detector whereby for the description of K the conditional probability $P_K\left(K \cap \Phi^{(\mathrm{T})}\right)$ and thus

$$P(k) = \int_0^\infty d\phi P_K\left(K = k|\mu_K = \Phi^{(\mathrm{T})}\right) p_{\Phi^{(\mathrm{T})}}\left(\Phi^{(\mathrm{T})} = \phi\right) \tag{A.68}$$

must be used. Here $P(k)$ is still a discrete probability distribution while $p_{\Phi^{(\mathrm{T})}}$ describes a continuous probability density. To both the normalization:

$$\sum_k P_K(K = k) = 1 = \int_\phi d\phi\, p_{\Phi^{(\mathrm{T})}}\left(\Phi^{(\mathrm{T})} = \phi\right) \tag{A.69}$$

applies. Note: Equation A.65 follows from Equation A.68 assuming $p_{\Phi^{(\mathrm{T})}}\left(\Phi^{(\mathrm{T})} = \mu_K\right) \approx 1$.

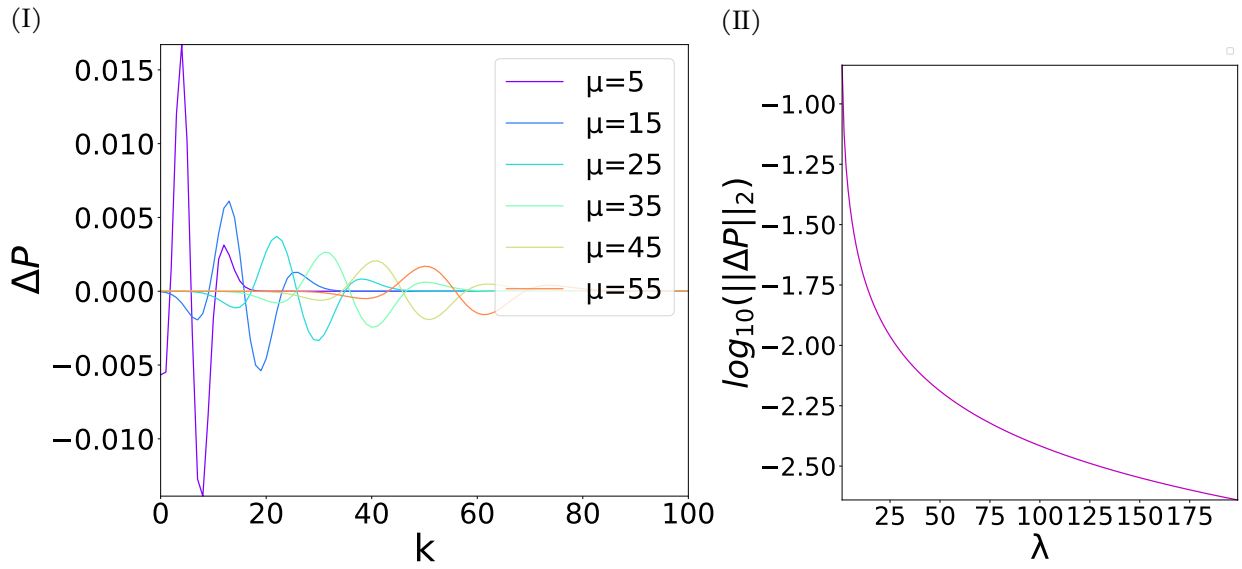


FIGURE A.3: **Distance Poisson and Gauss-Distribution** (a) Difference ΔP of Poisson $\mathcal{P}$ and Gaussian $\mathcal{G}$ distribution, i. e. $\Delta P = \mathcal{P} - \mathcal{G}$, are plotted at discrete integer-points k for different means and variances μ_K. (b) The $\log_{10}$ of l2-norm (Equation 1.8) of the distance of the Poisson- and Gauss distribution for different μ_K within the sampling-range of $[0, 200]$.

A.12 GAUSSIAN NOISE

Interestingly, as displayed in Figure A.3, the relative l_2-distance $\|\Delta P\|_2^2$ (Equation 1.8) with $\Delta P = \mathcal{P} - \mathcal{G}$ between Gaussian $\mathcal{G}$ (Equation 1.32) and Poisson $\mathcal{P}$ (Equation A.65) distribution[7] is already smaller than 10^{-2} for $\mu_K \geq 30$ and thus the two can be used interchangeably if thereby simplification in further analysis can be achieved.

A.13 ALIASING AND UNDERSAMPLING

For example, if the function is a standard cosine as

$$f_n(T) = a(T) \cdot \cos\left(2\pi \frac{n}{T}\right)$$

7 Note, that the Gaussian distribution is continuous in its random variable while the Possoin distribution is discrete. Hence, the comparison of the two is evaluated on an integer domain only.

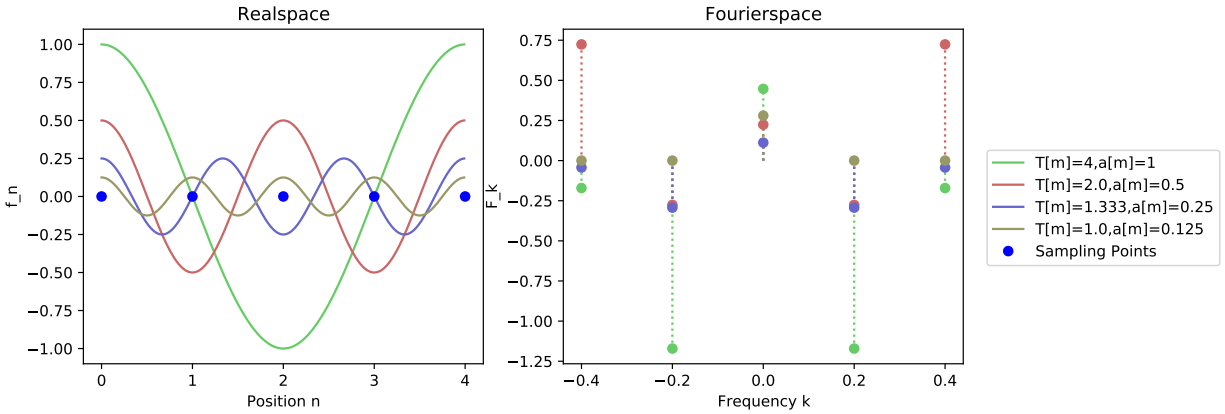


FIGURE A.4: **Aliasing in 1D-case.** Left (Real space): set of Cosine-Functions with different periods. Sampling points are marked with blue. Right (Fourier space): Fourier-Transform of the same functions, but reduced to the given sampling points. Hence: Fourier-Space consists of only 5 pixels.

with T period and $a(T)$ amplitude. If this function is eqidistal sampled with $d_{sample} = 1$, taking $N = 5$ sampling points and $a = 1$ then it results for different values of T:

$$\begin{aligned}
\text{Sampling Points:} \quad & n \in [0,1,2,3,4] \\
f_n(T=4) = cos(0.5\pi n) &= [1,0,-1,0,1] \\
\rightarrow \tilde{f_k} &= [-0.2,-1.2,0.4,-1.2,-0.2] \\
f_n(T=2) = cos(\pi n) &= [1,-1,1,-1,1] \\
\rightarrow \tilde{f_k} &= [1.4,-0.6,0.4,-0.6,1.4] \\
f_n(T=1.3\bar{3}) = cos(1.5\pi n) &= [1,0,-1,0,1] \\
\rightarrow \tilde{f_k} &= [-0.2,-1.2,0.4,-1.2,-0.2] \\
f_n(T=1) = cos(2\pi n) &= [1,1,1,1,1] \\
\rightarrow \tilde{f_k} &= [0,0,2.2,0,0]
\end{aligned}$$

showing that $f_n(T = 1.3\bar{3})$ and $f_n(T = 4)$ have the same values at the sampling points and are thus assigned to the same Fourier-frequencies. Note that while $n \in [0,1,...,4]$ it is $k \in [-2/N,-3/N,...,0,...,2/N-1]$. For better visualization amplitudes of $a = [1,0.5,0.25,0.125]$ for various periods T were used in Figure A.4.

Typical microscopic measurements will be represented as a list of 2D images. An example will be given for an imaged (low-pass filtered) *spokes2d* in Figure A.5. Imaging parameters are according to Section A.11 where $N^{(\text{pix})} = [256,256]$ and for a) the pixel-

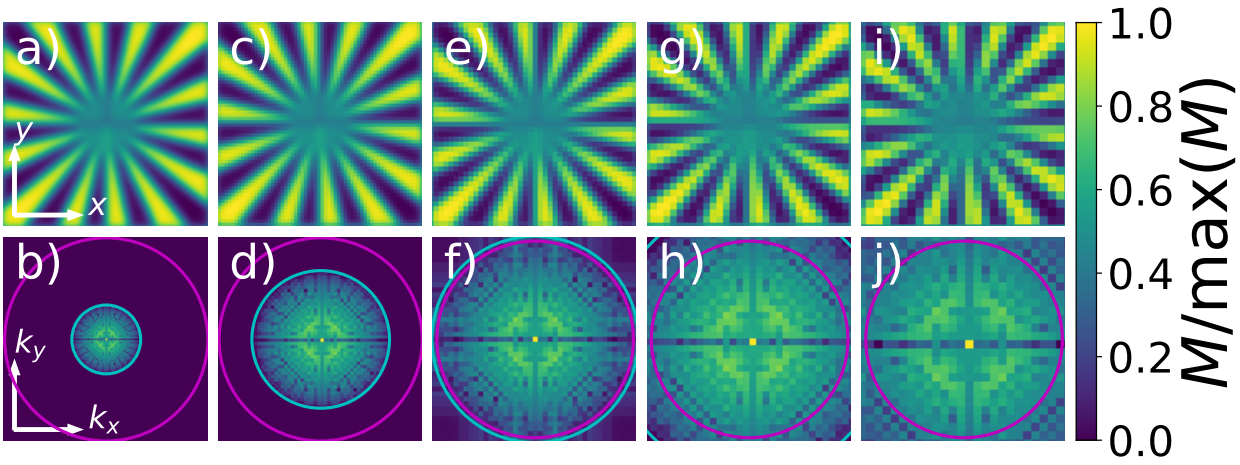


FIGURE A.5: **Undersampling** Ra-Re) Real-space images normalized to individual maximum, Fa-Fe) modulo of Fourier-Space images with normalization $\log_{10}(1 + lrVM)$. While Ra and Fa represent the original image with $d^{(\mathrm{SPL,lat})} = 30\,\mathrm{nm}$ pixelpitch, Rb-Re) and Fb-Fe) represent sub-sampled images with pixel-pitches of $d^{(\mathrm{SPL,lat})} = [2, 3, 4, 5] \cdot 30\,\mathrm{nm}$ and therewith simulate a set of detectors with different pixel-pitches and sizes. The cyan circle marks the OTF-support limit, the magenta circle the sampling frequency of the image.

pitch is $d^{(\mathrm{SPL,lat})} = 30\,\mathrm{nm}$. The further images b-e) are generated by rectangular sub-sampling where the image is locally summed with a mask of size b) 2×2, c) 2×3, d) 4×4 and e) 5×5 thereby reducing the resulting image size. The OTF cutoff frequency $k^{(\mathrm{C,WF})}$ is displayed with a cyan circle while the magenta-circle displays the sampling limit of the image $k^{(\mathrm{SPL,lat})}$. For Ra-c) $k^{(\mathrm{C,WF})} < k^{(\mathrm{SPL,lat})}$ and the images are properly sampled while for Rd-e) $k^{(\mathrm{C,WF})} > k^{(\mathrm{SPL,lat})}$ and the images are under-sampled. A unique reconstruction under the latter (Rd-e) circumstances of the sample is impossible as high-frequencies are measured as lower-frequencies in the image. In real space the under-sampling is represented by an increasing pixelation and thus loss of resolution.

A.14 FLAT-FIELDING UND UNITLESS IMAGES

For this section as well as the upcoming sections Section A.15 and Section A.16 a simple confocal setup with high numerical aperture $NA = 1.4$, excitation wavelength $\lambda^{(\mathrm{ex})} = 488\,\mathrm{nm}$ and emission wavelength $\lambda^{(\mathrm{em})} = 510\,\mathrm{nm}$ is assumed. Hence, the analysed images are scanned images and the counts represent detected photon emissions of the sample. In a natural 3D-sample wave-matter interaction processes like scattering, absorption or emis-

sion influence the effective system PSF and Out of Focus (ooF) information is present in the in-focus slice (Equation A.58). While ooF light might not be present in very thin or non-scattering objects, the 2D in-focus ($z = 0$) intensity distribution I of an ideal 3D-imaging process of a 3D sample on at the 2D detector can be approximated as:

$$I(\vec{\chi}) = I^{(\text{signal})}(\vec{\chi}) + I^{(\text{bg})}(\vec{\chi}) \tag{A.70}$$

Typically the in-focus signal $I^{(\text{signal})}$ consists of higher-frequencies and changes rather quickly laterally and axially, while the ooF information $I^{(\text{bg})}$ (=background) is composed of low-frequencies and changes rather slow w. r. t. the lateral field of view (FoV) and axial step-size. The used wording „rather"is quite imprecise, but cannot be made clearer without stating the used imaging method and hence whether inherent background-suppression exists. While in case of Confocal Laser Scanning Microscopy (CLSM) imaging the background $I^{(\text{bg})}$ can be locally confined to the region of the axial spread of the CLSM-PSF (Equation 1.22) it has to be extended to the whole volume of 3D-sample in case of a Widefield Microscopy (WF) system due to the missing axial sectioning.

Background can be assumed to be constant over the FOV at a given in-focus $z = z_l$ and can e. g. simply be estimated by:

- subtracting the mean of a Gaussian fitted to the lower frequencies of the histogram of the measured image from the image
- subtracting the mean calculated from low variance regions

Imperfections (physical trade-offs) of the detecting device during the amplification and read-out process lead to additional, mostly uncorrelated noise which typically is approximated by zero-mean Gaussian white noise. To avoid clipping on low-voltages, hence to reduce measurement bias due to missing sensitivity of the device, typically a voltage offset is applied to each sensor-element. This offset can be modeled into the imaging process (Equation 1.40) as additional term g_a which is a combination of the noise-free voltage-offset and the zero-mean Gaussian white noise. Hence, g_a can be assumed to be Gaussian distributed with $\mu^{(g_a)}$ as the voltage-offset.

A brief simulation of flat-fielding scenarios is displayed in Figure A.6. Here, 2D-WF-imaging of a 2D-sample (S) was applied and hence constant background at $I^{(\text{bg})} = \max\left(I^{(\text{illu})}\right)/10$ was added (Equation A.70), then Poisson-noise applied and finally the detector offset

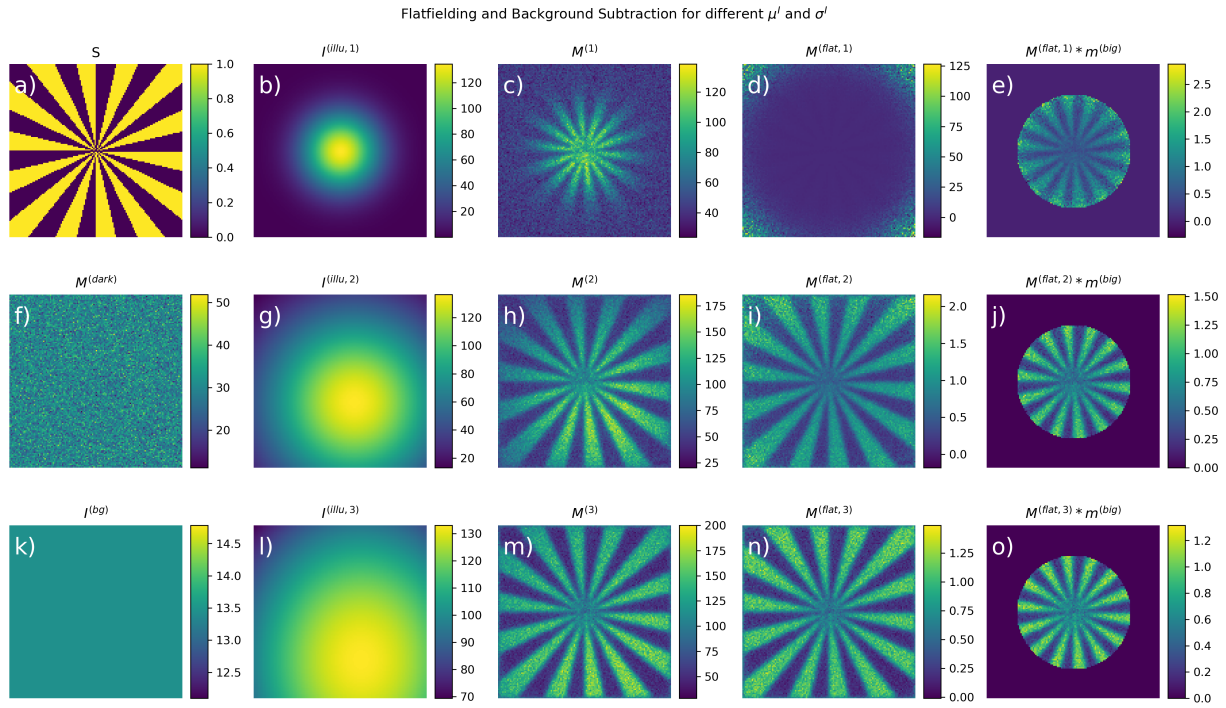


FIGURE A.6: **Flat-Fielding** Simulation to display effect of flat-fielding pre-processing on different levels of misaligned illumination and sizes. a) is the Sample while f),k) are examples of dark-image and background-intensity. 3 different illumination μ^I, σ^I combinations are compared. b)-e) illumination $I^{(\text{illu},1)}$, measured image $M^{(1)}$, flat-fielding corrected image $M^{(\text{flat},1)}$ and masked image $M^{(\text{flat},1)} \cdot m^{(\text{big})}$ for $\mu^I = (0,0)$, $\sigma^I = (20,20)$. g)-j) display the same for $\mu^I = (10,15)$, $\sigma^I = (50,50)$ and l)-o) for $\mu^I = (15,35)$, $\sigma^I = (100,120)$. 100 bright- (Sample=unit-plane) and dark-images where acquired and averaged. The image size used was 128×128 pixels of size $60\,\text{nm} \times 60\,\text{nm}$ and $NA = 1.4$ at $\lambda = 488\,\text{nm}$.

and noise added (Equation 1.40). The spokes-target serves as sample. The flat-field and offset-corrected image $M^{(\text{flat})}$ is calculated according to:

$$M^{(\text{flat})} = \frac{M - \overline{M}^{(\text{dark})}}{\overline{M}^{(\text{bright})} - \overline{M}^{(\text{dark})}} \overline{\overline{M}^{(\text{bright})} - \overline{M}^{(\text{dark})}} \tag{A.71}$$

where $\overline{M}^{(\text{bright})}$ is the average of all brightfield images taken without sample and $\overline{M}^{(\text{dark})}$ the average of all images taken without any illumination. The corrected image finally is rescaled by the average of difference of the averages of dark and bright-field images.

The simulation shows that naive flat-fielding of non-uniform illuminated samples (under condition of aberration-free PSF) can lead to severe amplification of non-signal regions, see borders of A.6 d) and its value range $\approx [0, 100]$. Even in case of more equal illumi-

nation g) the flat-fielded image still displays non-uniform radial brightness distribution and enhanced noise (compare i) with j) as well as value range $\approx [0,2]$). Even though g) does not suffer too much of noise-amplification at the borders due to a bigger and more equal illumination (value-range of i)) it is this time offset by $\mu^I = (15,35)$ pixels and thus higher noise amplification in the top-left as compared to the bottom-right corners becomes visible. Note that by limiting the FoV (e)-o)) even for severely bad illumination flat-fielding can be applied.

A.15 BINNING AND INTERPOLATION

During experiment design, particular care is taken to ensure that the pixel pitch used for detection satisfies at least the Nyquist criterion (Equation 1.37) for the assumed maximum frequency of image given aberration limited imaging. However, this criterion can be significantly exceeded by using a much smaller pixel pitch (higher maximum frequency k). Thus, due to the lower light yield per pixel on average the SNR of the imaging becomes worse. If an sCMOS camera was used for detection e. g. multiple pixels can be combined before conversion to digital units and thus readout noise is added only once (to the virtual, larger pixel). However, the photon noise remains the same. Since binning can be understood as a simple summation of the individual pixels, the SNR (Equation 1.33) per super-pixel according to Equation 1.40 is as follows:

$$\begin{aligned} SNR\left(N^{(\text{bin})}\right) &= \frac{\mathbb{E}\left\{\sum_{l=0}^{N^{(\text{bin})}} M_l\right\}}{\mathbb{V}\left\{\sum_{l=0}^{N^{(\text{bin})}} M_l\right\}} \\ &= \frac{\sum_{l=0}^{N^{(\text{bin})}} \mathbb{E}\left\{M_l\right\}}{\sum_{l=0}^{N^{(\text{bin})}} \mathbb{V}\left\{M_l\right\} + \sum_{l\neq m} \mathbb{W}\left\{M_l, M_m\right\}} \\ &= \frac{\overline{k}}{\sqrt{\overline{k} + N^{(\text{bin})} \cdot (\sigma^{(\text{read})})^2 + \sum_{l\neq m} \mathbb{W}\left\{M_l, M_m\right\}}} \end{aligned} \tag{A.72}$$

using independence between pixels with respect to the measurement process and zero-mean property of additional noise. $N^{(\text{bin})}$ is the number of pixels used for binning. In case of digital binning (=during pre-processing) $\mathbb{W}\left\{M_l, M_m\right\} \neq 0$, while it vanishes and $N^{(\text{bin})} = 1$ in the denominator if binning happens before voltage-conversion and digitalization (=read-out). Otherwise, $N^{(\text{bin})} \geq 1$ and thus the SNR gets worse.

Many interpolation schemes exist and come with their very individual costs. The information content acquired by the imaging system is limited by the sampling frequency (Equation 1.37) and no information beyond this limit can be regained without any prior-assumptions about the sample and particularly not by simple interpolation. Still, interpolation is used in everyday research to enlarge images and e. g. fit to varying display sizes and pixel-densities or to smoothen transitions and so on. In the case of image-processing routines, special attention must be paid to the influence of interpolation methods on pixel statistics and thus on error/variance propagation. While in iterative standard procedures such as e. g. linear, bi-linear or bicubic only mutual neighboring pixels contribute global interpolations such as FFT-based methods use all pixels. The consequences of this procedure will be briefly described using the example of 1) simple linear 2× and 2) Fourier interpolation.

1) Linear interpolation is a local, invertible and analytic interpolation-operation. For two neighbors such an interpolation scheme can be written as:

$$M_m = \begin{cases} M_m & m \in \mathbb{I}^{(\mathrm{l})} \\ \frac{1}{2}(M_{m-1} + M_{m+1}) & m \in \mathbb{I}^{(\mathrm{m})} \backslash \mathbb{I}^{(\mathrm{l})} \end{cases} \tag{A.73}$$

where the original-image indices $l \in \mathbb{I}^{(\mathrm{l})}$ are mapped to $m \in \mathbb{I}^{(\mathrm{m})}$ with size $\left|\mathbb{I}^{(\mathrm{m})}\right| = 2 \cdot \left|\mathbb{I}^{(\mathrm{l})}\right|$ and hence l are even indices in the new image. Assuming Poisson-noise only the statistics of the support-nodes (where $l \in \mathbb{I}^{(\mathrm{l})}$) stay unaltered while for the new created nodes $m \in \mathbb{I}^{(\mathrm{m})} \backslash \mathbb{I}^{(\mathrm{l})}$ it follows:

$$\begin{aligned} \mathbb{E}\{M_l\} &= \frac{1}{2}(\mu_l + \mu_{l+1}) \\ \mathbb{V}\{M_l\} &= \frac{1}{4}(\mu_l + \mu_{l+1}) \\ SNR(M_l) &= \mathbb{E}\{M_l\} / \sqrt{\mathbb{V}\{M_l\}} = \sqrt{\mu_l + \mu_{l+1}} \end{aligned} \tag{A.74}$$

Hence local linear interpolation has an averaging effect and (for Poisson distributed random-var) interpolated pixels have increased SNR.

2) Fourier-Interpolation is achieved by Fourier-transforming the input image $\widetilde{M} = \mathcal{F}\{M\}$, padding 0's around $\widetilde{M}$ and back-transforming it $M^{(\mathrm{pad})} = \mathcal{F}\left\{\widetilde{M}^{(\mathrm{pad})}\right\}$ in Fourier-space (continuous case)

$$M^{(\mathrm{pad})}(\vec{\chi}') = \mathcal{F}\left\{\mathcal{F}\{M\}(\vec{\kappa}) \cdot \mathrm{rect}\left(\vec{\kappa} \mid \mathrm{supp}\left(\widetilde{M}\right)\right)\right\} \tag{A.75}$$

$$M^{(\mathrm{pad})}(\vec{\chi}') = \sqrt{\frac{2\,|\vec{\kappa}|}{\pi}}\,(M(\vec{\chi}) \otimes \mathrm{sinc}(\vec{\kappa} \cdot \vec{\chi})) \tag{A.76}$$

where the window-function rect added zeros to the support of M according to:

$$\text{rect}\left(\vec{\kappa} \mid \text{supp}\left(\widetilde{M}\right)\right) = \begin{cases} 1, & \vec{\kappa} \in \text{supp}\left(\widetilde{M}\right) \equiv \vec{\kappa}_M \\ 0, & else \end{cases} \tag{A.77}$$

The statistical properties of such a convolution operation (Equation A.76) can be written in matrix notation, using R_{lm} as the resampling-kernel (here: sinc), as:

$$M_l = \sum_m R_{lm} M_m$$

$$\mathbb{E}\{M_l\} = \sum_m R_{lm} \mu_m \tag{A.78}$$

$$\mathbb{V}\{M_l\} = \sum_m R_{lm}^2 \mu_m \tag{A.79}$$

The convolution kernel is non-zero over the whole image and hence all pixels have a (diminishing) correlation. The resulting covariance matrix is of size $N^{(\text{cov})} = N \times N$ with N being the number of pixels of M. For big image stacks working with the full $N^{(\text{cov})}$ covariance-matrix is highly inefficient. According to Equation A.79 the variance then scales with the square of the convolution kernel and thus, in case of a fast vanishing kernel, only a small neighboring region (e. g. 1Airy Unit (AU) of the PSF could be 5×5 pixels and hence COV is 25×25 pixels) is a reasonable approximation.

Assuming two ideal, noise-free images with only 1 bead per image (and zero everywhere else) that are displaced by 1/4 pixel w. r. t. each-other, sub-pixel shifting needs to be applied to perfectly match the two images. This can be achieved by Fourier-shifting with a sub-pixel-phase, but if therefore a 4-fold interpolation is introduced according to Equation A.76, the side-lobes of the sinc not only spread over the whole image, but introduce negative values and hence change the assumed (underlying) statistic.

A.16 FOURIER-SELECTIVE FILTERING

Under the assumption of *band-limited information* systematic errors, e. g. like pickup of the field of the mains voltage due to poor shielding of the electronic components, could manifest with a strong modulus at frequencies outside the OTF-support. Unwanted contributions of this type are called pickup-noise. With Fourier-masking based methods like lowpass-filtering and selective frequency-filtering this contributions can be suppressed.

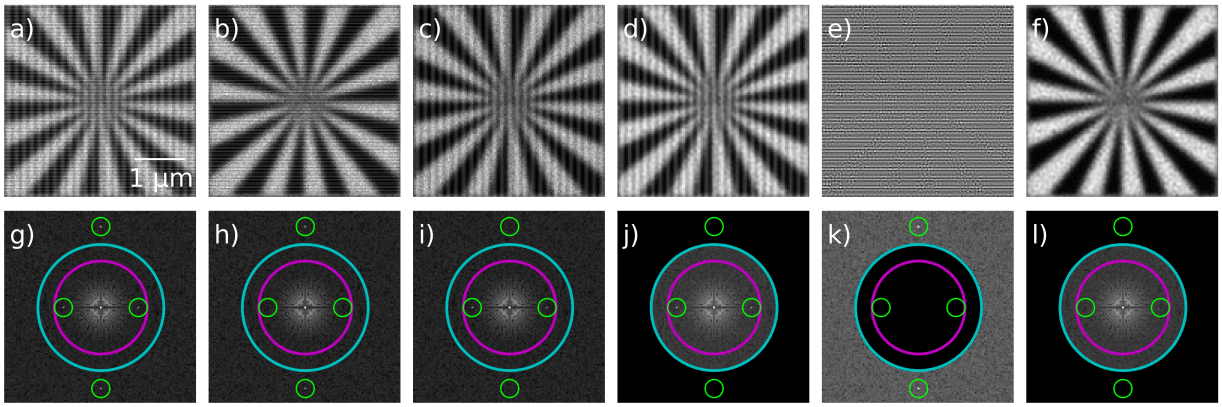


FIGURE A.7: **Fourier-Filtering strategies: Single Frequency- and Band suppression.** Continuation of the simulation from Section A.14. a)-f) show the Real- and d)-l) the Fourier-images. a)+g) overlay of noisy image and pickup-noise, b)+h) annealed out-of-band pick-up, c)+i) annealed in-band pickup, d)+j) low-pass filtered until OTF-support maximum, e)+k) high-pass filtered from OTF-support maximum, f)+l) low-pass filtered and in-band pickup annealed. OTF-support is displayed with cyan, effective (ideal) image-support with magenta and the pick-up noise positions are highlighted with lime.

Continuing the simulation from Section A.14 pickup-noise inside and outside the OTF-support has been added to the sample (see Figure A.7) and is marked with lime-green circles while the OTF-support is marked with a cyan-circle. While local annealing (horizontal pick-up in b)+h), vertical in c)+i)), i. e. replacing the values at certain Fourier frequencies with e. g. 0, can cure particular sinusoidal patterns in the image. Detecting and isolating these peaks can be difficult and within the OTF-support of the image, i. e. in-band annealing, can lead to loss of sample information. The combination of local- in-band annealing and low-pass filtering (f) + l)) already provides a reasonable cleanup of the image. e) displays the covariance introduced by this global (Fourier-based) processing. Even though no sample-information is left in the resulting image (compare k)) the object can still be recognized due to correlated noisy pixels, i. e. the covariance of two mutually neighboring pixel is unequal to zero. If such a high-pass filtering happened prior to the detection, e) would mostly consist of white uncorrelated noise. Due to this effect and the additional possibility to recover out-of-band information pre-processing measures should be used with care to not interfer with further processing steps [81].

A.17 ML OF SCALED GAUSSIAN

The same analysis as for Poisson will be investigated for Gaussian distributed random-variables (Equation 1.32). Two interesting cases can be analysed. 1) The gaussian distribution as limit to the spatially independent Poisson-distribution (multiplicative) 2) spatially independend additive Gaussian noise.

In case of 1) $\sigma_l^{(\mathrm{G})} = \sqrt{\mu_l}$ and thus it follows:

$$\begin{aligned}
\mathfrak{L}'^n = \partial_{S_n}\mathfrak{L} &= -\partial_{S_n}\ln\left(\prod_{l=0}^{N-1}\frac{1}{\sqrt{2\pi\mu_l}}\exp\left\{-\frac{1}{2\mu_l}\left(k_l-\mu_l\right)^2\right\}\right)\\
&= -\partial_{S_n}\sum_{l=0}^{N-1}\left(-\frac{\ln\left(2\pi\mu_l\right)}{2}-\frac{1}{2\mu_l}\left(k_l-\mu_l\right)^2\right)\\
&= -\sum_{l=0}^{N-1}\left(-\frac{2\pi}{4\pi\mu_l}\partial_{S_n}\mu_l-\frac{2\left(k_l-\mu_l\right)\mu_l-\left(k_l-\mu_l\right)^2}{2\mu_l^2}\partial_{S_n}\mu_l\right)\\
&= \frac{1}{2}\sum_{l=0}^{N-1}\left(\mu_l^{-1}+2\left(\frac{k_l}{\mu_l}-1\right)-\left(\frac{k_l}{\mu_l}-1\right)^2\right)\partial_{S_n}\mu_l \qquad ,\ \mu_l\neq 0\\
&= \frac{1}{2}\sum_{l=0}^{N-1}\left(\mu_l^{-1}+2b_l-b_l^2\right){H_l}^n \qquad ,\ b_l\equiv\frac{k_l}{\mu_l}-1\\
&= H^{ln}S'_l \qquad\qquad (A.80)\\
&\stackrel{!}{=} 0
\end{aligned}$$

with the partial derivative Equation 1.43 and $S'_l = \left(\mu_l^{-1}+2b_l-b_l^2\right)$. Equation A.80 could, in continuous coordinates, also be written as a convolution such that:

$$\mathfrak{L}'\left(\vec{x}\right) = \left[h^T\otimes S'\right]\left(\vec{x}\right) \qquad (A.81)$$

where the flip of the summation index of $\mathbf{H}$ in Equation A.80 corresponds coord-flipped PSF $h^T = h\left(-\vec{x}\right)$.

A.18 1/OTF DECONVOLUTION

As explained inSection A.11 the object information degrades through the imaging process. According to the imaging model given in Equation 1.41, it is reasonable to think of reconstruction by direct inversion of the forward model:

$$\tilde{S}_{k_l}^{(1/\mathrm{otf})} = \tilde{h}^{-1}_{a\,k_l}\widetilde{M}_{ak_l} = \tilde{S}+\tilde{h}^{-1}_{a\,k_l}\widetilde{\mathcal{N}}_{ak_l} \qquad (A.82)$$

where $\tilde{h}^{-1}$ is the inverse OTF. This approach partially succeeds for noise-free images ($\widetilde{\mathcal{N}}_{ak_l} = 0$) in case of limiting processing to the OTF-support, i. e. only allowing $k \leq k^C$. In the noisy case, however, row and column rank of the PSF matrix are equal to column and row number and the actually degraded PSF kernel is filled with random, potentially informationless values. This leads to a strong amplification at regions of weak transfer strength of the OTF when multiplying the image with the inverse OTF. Due to these two influences, a simple inversion of the imaging model fails and alternative strategies become necessary. Another obvious approach is to reduce the gain of the high frequency components of the OTF support by introducing a regularization parameter ε according to:

$$\tilde{S}_{k_l}^{(1/\mathrm{otfr})} = \begin{cases} \left(\tilde{h}_{a\,k_l} + \varepsilon\right)^{-1} \widetilde{M}_{ak_l} \approx \tilde{S}_C + \left(\tilde{h}_{a\,k_l} + \varepsilon\right)^{-1} \widetilde{\mathcal{N}}_{ak_l}, & \left|\tilde{h}\right| > \varepsilon \\ 0, & \left|\tilde{h}\right| \leq \varepsilon \end{cases} \tag{A.83}$$

where the latter approximation is true for a thresholded support of size C of S where $\varepsilon / \tilde{h}_{a\,k_l} \ll 1$. ε regularizes the problem by reducing the effective size of the OTF support, i. e. $k^{(\mathrm{C},1/\mathrm{OTF})} \leq k^{(\mathrm{C})}$ as the OTF is set to 0 for modulo-values smaller than ε, and thus the actual effect of deconvolution.

A.19 WIENER DECONVOLUTION

Assuming an information-limited measurement process, it can be conjectured that the best possible estimator $\hat{S}_l$ of each individual measured image pixel l is Gaussian distributed $\mathcal{G}_l$ around the original object distribution S_l and with Equation 1.32 here can be written as:

$$\mathcal{G}_l = A\, e^{-\frac{\left|\hat{S}_l - S_l\right|^2}{2\sigma_l^2}} \tag{A.84}$$

where $\mathbb{E}\{G_l\} = S_l$ and $\mathbb{V}\{G_l\} = \sigma_l^2$ are mean and variance of the Gaussian. The log-likelihood problem for an N-fold measurement process can then be written as:

$$\ln \mathcal{L} = \ln \prod_l^N \mathcal{G}_l = \sum_l^N \ln(\mathcal{G}_l) = -\sum_l^N \frac{\left|\hat{S}_l - S_l\right|^2}{2\sigma_l^2} \tag{A.85}$$

Introducing a normalized log-likelihood function $L = -\ln \mathcal{L} / 2\sigma_X^2$ and interpreting the sum over all possible evaluations as the expected value $\mathbb{E}$, we obtain:

$$L = \mathbb{E}\left\{\left|\hat{S}_l - S_l\right|^2\right\} = \mathbb{E}\left\{\left|\hat{\tilde{S}}_{k_l} - \tilde{S}_{k_l}\right|^2\right\} \equiv \tilde{L} \tag{A.86}$$

where for the second step Parseval-Plancherel's identity [82] was used. A filtering function $\widetilde{W}$ is to be found which reconstructs the object distribution as best as possible using the measured image according to:

$$\hat{\tilde{S}} = \widetilde{W}\widetilde{M} = \widetilde{W}\left[\tilde{h}\tilde{S} + \widetilde{\mathcal{N}}\right] \tag{A.87}$$

where here and in the following all indices k_l have been omitted for easier readability. For Equation A.86 thus results:

$$\tilde{L} = \mathbb{E}\left\{\left|\left[\widetilde{W}\tilde{h} - 1\right]\tilde{S} + \widetilde{W}\widetilde{\mathcal{N}}\right|^2\right\} \tag{A.88}$$

$$\tilde{L} = \left|\tilde{S}\right|^2\left[\widetilde{W}\tilde{h} - 1\right]\left[\widetilde{W}\tilde{h} - 1\right]^* + \left|\widetilde{\mathcal{N}}\right|^2\,\widetilde{W}\widetilde{W}^* - \underbrace{2\mathfrak{Re}\left\{\mathbb{E}\left\{\left[\widetilde{W}\tilde{h} - 1\right]\tilde{S}\widetilde{W}^*\widetilde{\mathcal{N}}^*\right\}\right\}}_{=0} \tag{A.89}$$

The last summand of Equation A.89 represents the expected value of the product of the uncorrelated random variables $\tilde{S}$ and $\widetilde{\mathcal{N}}$ which equals 0. Applying the expected value to the remaining summands and taking $\partial\widetilde{W}^*/\partial\widetilde{W} = 0$ of complex derivatives into account, the filter-function $\widetilde{W}$ can be determined from $\tilde{L}$ via extrema search, i. e. derivation of $\tilde{L}$ w. r. t. to $\widetilde{W}$ and solving for $\widetilde{W}$:

$$\begin{aligned} 0 \stackrel{!}{=} \frac{\partial\tilde{L}}{\partial\widetilde{W}} &= \left|\tilde{S}\right|^2\left[\widetilde{W}^*\left|\tilde{h}\right|^2 - 1\right] + \widetilde{W}\,\mathbb{E}\left\{\left|\widetilde{\mathcal{N}}\right|^2\right\} \\ \widetilde{W}^* &= \frac{\tilde{h}^*\left|\tilde{S}\right|^2}{\left|\tilde{S}\right|^2\left|\tilde{h}\right|^2 + \mathbb{E}\left\{\left|\widetilde{\mathcal{N}}\right|^2\right\}} \end{aligned} \tag{A.90}$$

The found filter function $\widetilde{W}$ Equation A.90 can be further reduced and finally leads to the Wiener Filter Equation 1.47.

A.20 WEIGHTED AVERAGING IN FOURIER SPACE

In the following k_l will be omitted for the ease of readability. To determine the real-valued SNR (Equation 1.33), the expected value (**??**) and variance (Equation 1.5) of Equation 1.48 are first determined:

$$\mathbb{E}\left\{\widetilde{M}^{(\mathrm{wa})}\right\} = \sum_a \widetilde{\omega}_a\mathbb{E}\left\{\widetilde{M}_a\right\} = \sum_a \widetilde{\omega}_a\tilde{h}_a\tilde{S} \tag{A.91}$$

$$\mathbb{V}\left\{\widetilde{M}^{(\mathrm{wa})}\right\} = \sum_a \widetilde{\omega}_a\widetilde{\omega}_a^*\mathbb{V}\left\{\tilde{M}_a\right\} + \sum_a\sum_{b\neq a}\widetilde{\omega}_a\widetilde{\omega}_b^*\widetilde{\Xi}_{ab} \tag{A.92}$$

with covariance (Equation 1.7) $\Xi_{ab} \equiv \mathbb{W}\left\{\widetilde{M}_a, \widetilde{M}_b\right\}$. Given that $\mathbb{V}\left\{\widetilde{M}_a\right\} = \tilde{\sigma}_a^2$ the SNR can be calculated according to:

$$SNR = \frac{\left|\mathbb{E}\left\{\widetilde{M}^{(\mathrm{wa})}\right\}\right|}{\sqrt{\mathbb{V}\left\{\widetilde{M}^{(\mathrm{wa})}\right\}}} = \sum_a \frac{\widetilde{\omega}_a \tilde{h}_a \tilde{S}}{\sqrt{\sum_a \widetilde{\omega}_a \widetilde{\omega}_a^* \tilde{\sigma}_a^2 + \sum_a \sum_{b \neq a} \widetilde{\omega}_a \widetilde{\omega}_b^* \widetilde{\Xi}_{ab}}} \tag{A.93}$$

The maximum of this measure can be determined by extrema search for the weights $\widetilde{\omega}_c$.

$$\begin{aligned} 0 &\stackrel{!}{=} \frac{\partial SNR}{\partial \widetilde{\omega}_c} \\ &= \tilde{h}_c \tilde{S} \mathbb{V}\left\{\widetilde{M}^{(\mathrm{wa})}\right\} - \sum_a \widetilde{\omega}_a \tilde{h}_a \tilde{S} \left[\widetilde{\omega}_c \tilde{\sigma}_c^2 + \sum_{b \neq c} \widetilde{\omega}_b \Xi_{cb}\right] \end{aligned} \tag{A.94}$$

Where symmetry of the covariance was used and the assumption that $\widetilde{\omega}_c$ is hermitian was made thereby omitting the factor 1/2 in the negative summand. After inserting Equation A.92 and sorting we obtain:

$$\tilde{h}_c \left[\sum_a \left|\widetilde{\omega}_a\right|^2 \tilde{\sigma}_a^2 + \sum_a \sum_{b \neq a} \widetilde{\omega}_a \widetilde{\omega}_b \widetilde{\Xi}_{ab}\right] = \left[\widetilde{\omega}_c \tilde{\sigma}_c^2 + \sum_{b \neq c} \widetilde{\omega}_b \Xi_{cb}\right] \sum_a \widetilde{\omega}_a \tilde{h}_a \tag{A.95}$$

For the two sides of Equation A.95 to be equal their prefactors (c-terms) and their sums (a-terms) must be the same, so e. g. if:

$$\tilde{h}_c \sum_a \left|\widetilde{\omega}_a\right|^2 \tilde{\sigma}_a^2 = \widetilde{\omega}_c \tilde{\sigma}_c^2 \sum_a \widetilde{\omega}_a \tilde{h}_a \tag{A.96}$$

$$\tilde{h}_c \sum_a \sum_{b \neq a} \widetilde{\omega}_a \widetilde{\omega}_b \widetilde{\Xi}_{ab} = \left[\sum_{b \neq c} \widetilde{\omega}_b \Xi_{cb}\right] \sum_a \widetilde{\omega}_a \tilde{h}_a \tag{A.97}$$

Assuming that the different views a are independed all $\Xi_{cb} = 0$ vanish and hence from Equation A.96 directly the weights Equation 1.49 follow.

A.21 ML OF POISSON AND THE RICHARDSON-LUCY DECONVOLUTION

In case of indepedent Poisson-distributed random-measurements $P_{K_l}(K_l = k_l \,|\, \mu_l)$ (Equation A.65) it follows for the derivative ∂_{S_n} of $\mathfrak{L}$ w. r. t. the parameters S_n:

$$\begin{aligned}
\partial_{S_n}\mathfrak{L} &= -\partial_{S_n} \ln\left[\prod_{l=0}^{N-1} P_{K_l}(K_l = k_l \,|\, \mu_l)\right] \\
&= -\partial_{S_n} \sum_{l=0}^{N-1} \ln\left\{\frac{\mu_l^{k_l}}{k_l!} e^{-\mu_l}\right\} = -\partial_{S_n} \sum_{l=0}^{N-1} \left[k_l \ln\{\mu_l\} - \ln\{k_l!\} - \mu_l \ln\{\exp\}\right] \\
&= -\sum_{l=0}^{N-1} \left[\frac{k_l}{\mu_l}\partial_{S_n}\mu_l - 0 - \partial_{S_n}\mu_l\right] \\
&= \sum_l H_{ln}\left[\mathbb{1}_l - \frac{k_l}{\mu_l}\right] = \sum_l H_{ln}\left[\mathbb{1}_l - \frac{k_l}{\sum_m H_{lm} S_m}\right] \qquad \text{(A.98)} \\
&\stackrel{!}{=} \mathbb{0}_n
\end{aligned}$$

with the partial derivative according to Equation 1.43, the (diagonal) unit-matrix $\mathbb{1} \in \mathbb{N}^{N_l/2 \times N_l/2}$, i. e. the diagonal is filled with ones, and $\mathbb{0}$ a diagonal-matrix of the same shape as $\mathbb{1}$, but filled with zeros. Equation A.98 is again a convolution, but this time by using a summation about the first index which corresponds to convolution with a flipped PSF $h(-\vec{x})$ in continuous notation. Even if h (or its matrix representation $\mathbf{H}$) is known an analytic solution for the optimal $\hat{S}$ is difficult to find. Hence iterative algorithms can be used to find an approximate solution. In case $\mathfrak{L}$ is a convex function a global minimum exists and it can be found by using duality formulations of $\mathfrak{L}$ to convert it to another problem-type solveable by highly efficient linear or quadratic problem solvers. Proving whether the problem is convex tends to be rather difficult [83]. It is beyond the scope of this thesis to dwell further on convexity and existance of solutions to a given inverse problem in optics and thus the reader is referred to e. g. [62].

Two standard iterative schemes are Gradient Descent (GD) and Fixpoint Iteration (FI). For each iterative step $r+1$, the unknown parameter $S^{(\mathrm{r}+1)}$ is estimated from the previous step with an update formula. For GD the idea is to start with a random or mean (of the input data) initialization of the parameter $S^{(0)}$ and iteratively update the guess $S^{(\mathrm{r})}$ by

using the the gradient thereby converging towards the desired optimal solution . The update scheme:

$$S^{(\mathrm{r}+1)} = S^{(\mathrm{r})} - h\partial_{S_n}\mathfrak{L}^{(\mathrm{r})} \tag{A.99}$$

$$S^{(\mathrm{r}+1)} \underset{l\to\infty}{\longrightarrow} \hat{S} \quad \text{if } \partial_{S_n}\mathfrak{L}^{(\mathrm{r})} \underset{l\to\infty}{\longrightarrow} 0 \tag{A.100}$$

tends towards the optimal estimator $\hat{S}$ in case that the gradient $\partial_{(S_l)}\mathfrak{L}$ (approximately) vanishes. By choosing the stepsize h a tradeoff between speed and convergence is achieved. While a bigger h allows for bigger steps along the derivative direction $\partial_{S_n}\mathfrak{L}^r$ it could overshoot around the optimum, but given the step-size never reach it and thus reducing the degree of convergence. It can be exchanged by inserting a line-search per step and hence have a variable step-size.

Now, taking Bayes theorem Equation 1.54 and reinterpreting $P(\vec{k} \mid \vec{\mu}) = h(\vec{-x}) \otimes M$ as imaging-probability distribution as well as the original sample distribution as $P(\vec{\mu}) = S^{(\mathrm{r})}$ it follows:

$$S_n^{(\mathrm{r}+1)} = \frac{\left(\sum_l H_{ln} k_l\right) S_n^{(\mathrm{r})}}{\sum_m H_{nm}\, S_m^{(\mathrm{r})}} \tag{A.101}$$

with $P(\vec{k}) = \sum_{\vec{\mu}} P(\vec{k} \mid \vec{\mu}) P(\vec{\mu})$, see [84]. The update scheme converges to the solution in case of

$$\frac{S^{(\mathrm{r}+1)}}{S^{(\mathrm{r})}} \underset{l\to\infty}{\longrightarrow} 1 \tag{A.102}$$

and is called Richardson-Lucy Deconvolution (RL) [85]. It was shown that the scheme converges towards the ML-optimal estimator, if existent [86].

B

APPENDIX: METRICS FOR SOFTWARE AUTOFOCUS

B.1 METRICS FOR SOFTWARE AUTOFOCUS

The selection of image sharpness measures as well as a short interpretation of their meaning is given for Section 2.2. The sharpness metrics Ψ used for the comparison are a selection from the analyses in the supplement of [87], from where the categorization was also adopted. Additional categories such as „histogram"-based or „intuitive thresholds"as in [49] were not added.

The image stacks used in the main section are mostly 4D or 5D datasets, e. g. $dim\,(M) = [X, Y, Z, \mathcal{P}_{pois}, \mathcal{A}_{spher}]$, however, the metrics $\forall n, o, p$ of the dataset M_{lkmn} are applied, allowing the simplified description $M_{l,m} = M_{lm}$ to be used to calculate the metrics. The comma is introduced to conveniently represent any shifts, e. g. along the x direction by 1 pixel via $l + 1$, where N_M represents the number of pixels in each 2D slice M_{lm}. By normalizing to the pixel number of each individual image $M_{l,m}$, the calculation is image size independent.

B.1.1 *Differential Measures*

Differential measures are the computationally simplest measures in this comprehension as only summations and simple image shifts are involved. They are derivative-based, where the derivatives are represented by simple addressing shifts (e. g. $M_{l+1,m}$) using difference quotient. Thus, Equation B.1 represents the central difference quotient representation of the first derivative along the X-direction, where mostly $\Delta x = 0.5$ is set.

$$\partial_x M \approx \frac{M_{l+1,m} - M_{l-1,m}}{2\Delta x} \tag{B.1}$$

By using the first or second derivatives (along arbitrary spatial directions) of the image, especially edges can be emphasized. The basic idea of differential filters is that in-focus elements have especially many high frequencies and thus sharp edges compared to defocus elements.

Brenner Stresses the edge-enhancing property of the first derivative by squaring it [88]. Although this implementation of the measure only uses the first (central) derivative along 1 direction it can work quite reliably if samples can be assumed to display features in equal amounts and sizes towards all (lateral) directions.

$$BRE(M) = \frac{1}{N_M} \sum_{lm} \left\{ M_{l,m-1} - M_{l,m+1} \right\}^2 \tag{B.2}$$

Tenengrad Originally „Tenenbaum gradient" [89] it is the sum of a convolution of the image with two different kernels, namely the horizontal Sobel SBH and the vertical Sobel kernel SBV [90], which allow to calculate a more stable isotropic gradient efficiently.

$$\begin{aligned} TEN(M) &= \frac{1}{N_M} \sum_{lm} \left(SBH_{lm}(M) + SBH_{lm}(M) \right) \\ SBH_{lm}(M) &= M_{l+1,m-1} + 2M_{l+1,m} + M_{l+1,m+1} \\ &\quad - M_{l-1,m-1} - 2M_{l-1,m} - M_{l-1,m+1} \\ SBV_{lm}(M) &= M_{l-1,m+1} + 2M_{l,m+1} + M_{l+1,m+1} \\ &\quad - M_{l-1,m-1} - 2M_{l,m-1} - M_{l+1,m-1} \end{aligned} \tag{B.3}$$

Total Variation Typically Total Variation is used to e. g. calculate a linear measure for the distance between a measurement and a model. The same idea can be reinterpreted as a measure for sharpness along the individual X- and Y-direction [36]. The filter is non-differentiable, but pseudo isotropic as diagonals are not explicitely taken into account.

$$TOV(M) = \frac{1}{N_M} \sum_{lm} \sqrt{\left(M_{l+1,m} - M_{l-1,m} \right)^2 + \left(M_{l,m+1} - M_{l,m-1} \right)^2} \tag{B.4}$$

B.1.2 *Correlative Measures*

Correlative measures are based on self-similarity assumptions within an image. While typically a correlation can be expressed as a product using Fourier-Transformations Equation B.5 correlative measures are trying to identify

$$M \star M' = \mathcal{F}^{-1} \left\{ \mathcal{F}(M) \cdot \mathcal{F}(M')^* (-k) \right\} \tag{B.5}$$

The Vollath metrics used here are based on the approach that neighboring pixels differ only by their relative amplitude and the local, (potentially) uncorrelated noise. Vollath has presented 2 combination metrics, F4 and F5, which are independent of the (additive) noise terms as well as the image mean [91].

Symmetric Vollath F4 The differences between neighboring pixel values are weighted by the local amplitude, causing the filter to reach higher values as the edge sharpness or number of edges increases. Pixel values around the mean, on the other hand, hardly contribute to the filter value. Even though lower frequencies are weighted less than higher spatial frequencies the highest frequencies of the image-frequency-support are damped as well.

$$\begin{aligned} VO4(M) = \frac{1}{N_M}(&V4A(M;1,0) + V4A(M;-1,0) \ldots \\ &+ V4A(M;0,1) + V4A(M;0,-1)) \end{aligned} \tag{B.6}$$

$$V4A(M;\Delta l,\Delta m) = \left|\sum_{lm} M_{l,m}\left(M_{l+\Delta l,m+\Delta m} - M_{l+2\Delta l,m+2\Delta m}\right)\right|$$

Vollath F5 This metric increases less and less as the image becomes sharper and the relative weighting of the Fourier components actually decreases steadily, making this metric more suitable for flat or smooth edges.

$$VO5(M) = \frac{1}{N_M}\left(\sum_{lm} M_{l,m}M_{l+1,m} - \mathbb{E}\{M\}\right) \tag{B.7}$$

B.1.3 *Statistical Measures*

Statistical measures operate on statistical properties of the underlying image. Histogram properties like maximum, minimum or median are typical as well as statistical moments like mean, variance or kurtosis. **Maximum** Assuming that the sharpest edges and thus the largest localization of brightness are present at the focus, the maximum provides a measure to detect this point. It is obviously susceptible to offset (e. g. due to local fluctuating illuminances) as well as noise (e. g. pick-up, hot-pixels or cosmic noise).

$$MAX(M) = max(M_{lm}) \tag{B.8}$$

Variance Although the normalized variance $NVA = \frac{\mathbb{V}\{\cdot\}}{\mathbb{E}\{\cdot\}^2}$ is more independent of image size and image detail (e. g. at varying magnification) [92] stable results can already be achieved with the direct variance. For an image stastically normalized to $\mathbb{E}\{M\} = 0$, the variance reduces to the total mean of the pixel squares of the image and thus the mean square deviation of the image from 0.

$$VAR(M) = \frac{1}{N_M}\sum_{lm}|M_{lm} - \mathbb{E}\{M\}|^2 \tag{B.9}$$

B.1.4 *Spectral Measures*

Spectral Measures operate on representations of the input data in alternative bases or spaces. As basic basis decompositions, instead of the Fourier transform $\mathcal{F}\{\cdot\}$ used so far, the discrete cosine transformation (DCT) $\mathcal{F}_C\{\cdot\}$ is used here, where the DCT is defined according to Equation B.10.

$$\widetilde{M}_l = \sum_{m=0}^{M_N-1} \cos\left(\left(n+\frac{1}{2}\right)\frac{\pi k}{N}\right) \tag{B.10}$$

Kristan's Entropy Combines a sum-normalized-to-1 DCT of the input image with Bayes entropy function $H^2 = 2\left(1-\sum_l p_l^2\right)$, with H being the resulting entropy measure and p_l the probabilities of the according pixel-values. It assumes that an image when increasing focus tends from having low-frequencies only towards a uniform amplitude distribution across the DCT-space. While any entropy function would suffice, the choice of Bayes entropy suppresses the influence of uniformly distributed (e. g. Poisson) noise in lower DCT-frequencies. The measure is inherently independent to brightness variations between images and higher frequencies outside the assumed or calculated image-support are left-out. The described procedure is applied for every tile of the non-overlapping tiles $M^{(\text{tiled})}$ of each picture. Fast implementations of DCT are used eg for JPEG compression in consumer digital cameras, where the tiling is typically chosen to be quadratic and of size $N_M^{\text{tile}} = 2^4 = 16$ which results into limiting the DCT-support to a radius of $r_0 = 6$ as chosen originally [93]. Still, within this research better results where achieve with only changing the tile-size slightly to $N_M^{\text{tile}} = 2^3 = 8$ and choosing $r_0 = 2$. Finally, the mean of the tiles yields the resulting measure .

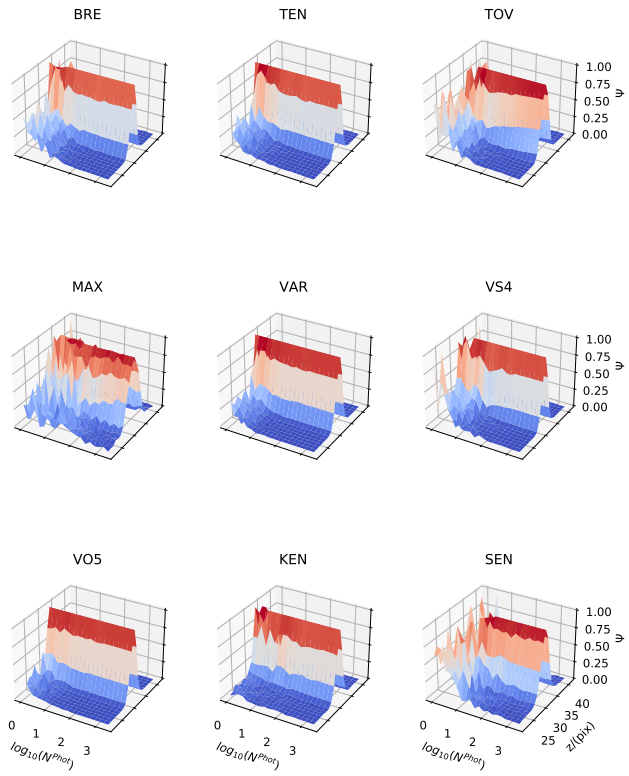
$$KEN(M) = -\mathbb{E}\left\{1 - \frac{\sum_{l+m<r_0} \mathcal{F}_C\left\{M^{(\text{tiled})}\right\}_{lm}^2}{\left(\sum_{l+m<r_0} \mathcal{F}_C\left\{M^{(\text{tiled})}\right\}_{lm}\right)}\right\} \tag{B.11}$$

Shannon Entropy Based on Kristan's approach, this measure is a normalized DCT-based version of Shannon Entropy, which is a measure of the energy of a distribution. The calculation can be understood as the product of significance (prefactor) and uncertainty (log term), where the logarithm is secured to the range $(0,\infty)$ [87].

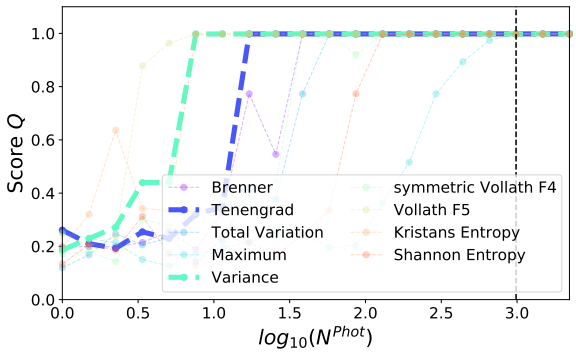
$$SEN(M) = -\frac{2}{r_0^2} \sum_{l+m<r_0} \left|ENT(M)_{l,m}\right| abslog_2\left[ENT(M)_{l,m}\right] \tag{B.12}$$

$$ENT(M)_{l,m} = \frac{\mathcal{F}_C\{M\}_{l,m}}{L_2\left(\mathcal{F}_C\{M\}_{l,m}\right)}$$

(I)



(II)



(III)

	$\tilde{q}^{\text{accu}}$	$\tilde{q}^{\text{Nfalse}}$	$\tilde{q}^{\text{range}}$	$\tilde{q}^{\text{FWHM}}$	Q
BRE	1.00	1.00	1.00	0.98	1.00
TEN	1.00	1.00	1.00	0.98	1.00
TOV	1.00	1.00	1.00	0.97	1.00
MAX	1.00	1.00	1.00	0.96	1.00
VAR	1.00	1.00	1.00	0.98	1.00
VS4	1.00	1.00	1.00	0.98	1.00
VO5	1.00	1.00	1.00	0.98	1.00
KEN	1.00	1.00	1.00	0.98	1.00
SEN	1.00	1.00	1.00	0.97	1.00

FIGURE B.1: **2D Noise testing** of the sharpness metrics using *spokes2d* thin object (Figure 2.10$a) - d)$). a) Surface plot display of resulting (non-filtered) sharpness values per metric for different Photon-counts $N^{(\text{Phot})}$ and Z-positions, thresholded with $u = 0.07$. b) Scoring Results Q for each metric Ψ with exemplary score-calculation display at position $b = 0.3$ (black dotted line) and hence $N^{(\text{Phot})} \approx 11$.

B.2 FURTHER METRIC PERFORMANCE EVALUATION

In this section first the behavior w. r. t. noise, spherical aberrations and astigmatism will be analyzed individually.

Using the *spokes2d* target as an example, the metrics are first tested with respect to noise level stability. The (pixel-wise) expected value λ desired for the application of Poisson noise was calculated by normalizing the simulated 3D stack to its 3D maximum and subsequent multiplication with the photon level. Photon levels were calculated via $N^{(\text{Phot})} = 1.5^b$ with $b \in [0, 20)$ and thus $N^{(\text{Phot})} \in [1, 2217)$. The metric results are shown in Figure B.1I as a 3D-surface plot. The scoring is listed using the example of noise level

$b = 3$ ($N^{(\mathrm{Phot})} \approx 3$) in Figure B.1III. The calculated scores for all noise levels and metrics are shown in Figure B.1*b*). Despite extremely low photon counts $N^{(\mathrm{Phot})}$, most metrics already reach a score $Q \geq 0.6$, whereas $\Psi^{(\mathrm{KEN})}$ and $\Psi^{(\mathrm{VAR})}$ can already predict the exact maxima position almost reliably with a score of $Q \approx 1$. While $\Psi^{(\mathrm{VO5})}$ (by definition) and $\Psi^{(\mathrm{KEN})}$ (by appropriate choice of filter mask size $r_0 = 2$) can maintain good contrast between maximum and background even in the low photon range, the contrast of the other metrics improves proportionally with increasing $N^{(\mathrm{Phot})}$. The maximum metric can generally be identified as hardly suitable, since it does not determine a clear maximum in the case of low $N^{(\mathrm{Phot})}$, the contrast remains low even for high $N^{(\mathrm{Phot})}$ and the scoring converges slower than the other metrics towards $Q \approx 1$.

The spherically aberrated PSFs required for the simulation are shown in Figure B.2. Due to an increasing level of aberration the PSF quality degrades strongly thereby giving the impression that instead of one central focus position the position of highest brightness and local sharpness splits into two (or more) slices above and below the nominal focus. The application of a sharpness metric Ψ will lead to a 1D Z-curve with two (or more) maxima. The evaluation measures $\bar{q}^{(\mathrm{i})}$ then will result in a very low score Q as they base their calculations on one central maximum and not axially shifted multiple maxima. Hence even though the sample might still be in perfect focus (even though the system PSF is severely aberrated) the scoring algorithm might state that its not. This could happen to a sample that is a set of axially ($d \geq d^{(\mathrm{1AU,ax})}$) spaced thin objects as well. Here the scoring algorithm would be used to assure that the measured individual slices stay at their respective focus position.

To take these effects into account the algorithm is extented to the application with multiple maxima. To keep this as reproducible as possible, the focus positions were calculated directly by applying the $\Psi^{(\mathrm{VO5})}$ sharpness filter onto the generated PSFs, subsequent $F^{(\mathrm{SGF})}$ filtering, and determination of the maxima positions and thresholding ($u = 0.07$). The resulting list of focal slices is then manually compared to the optical impression and inserted as a lean dashed line in Figure B.2 and used for further scoring calculations. Assuming a thin sample axially centered in the measured 3D-stack, the presented scoring algorithm together with the just this determination of focal positions is usable. For non-axially centered as well as extended samples, an offset in the calculation of the forward model and thus a changed focus position can occur. However, an error arising here can be limited by the final manual focus position check. It can be seen that the focus position for

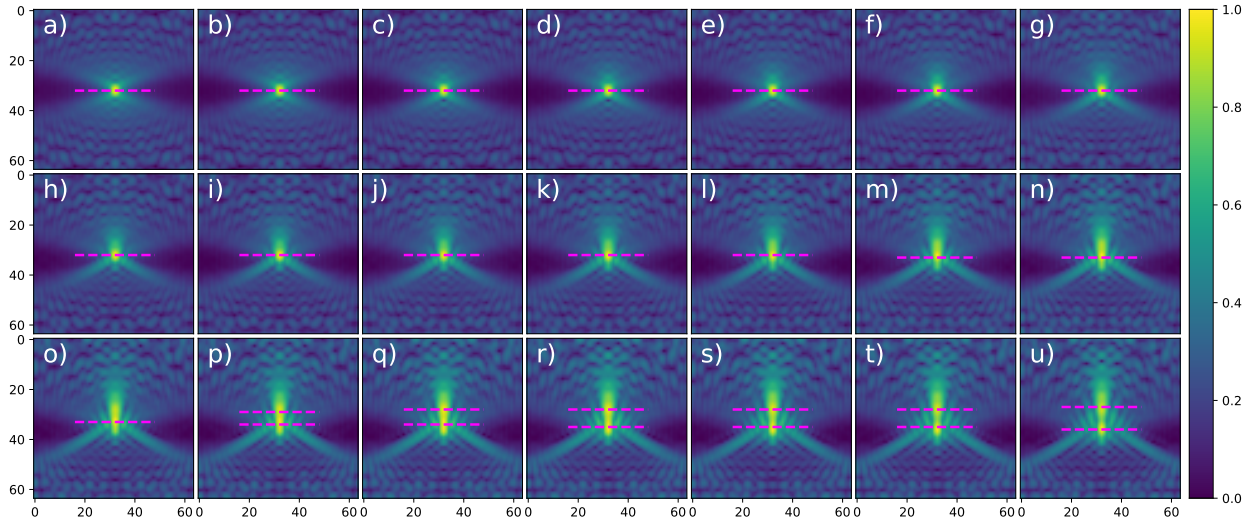
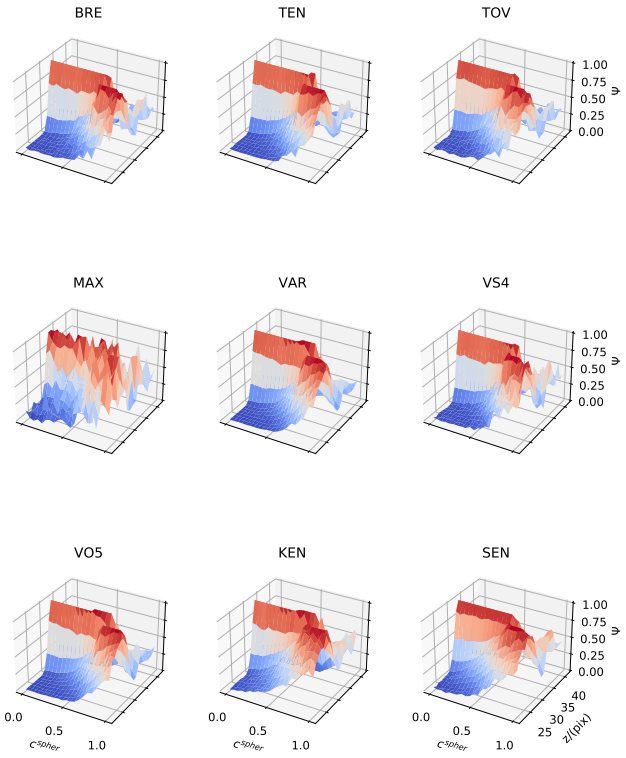


FIGURE B.2: **Aberrated PSFs** for different levels of spherical aberration severeness $c^{(\text{spher})} \in [0.0, 1.0]$ with step-size of $\Delta c^{(\text{spher})} = 0.05$ and direction of increase is sorted alphabetically $a) \rightarrow u)$. The simulated PSF is normalized to fulfill $\sum h = 1$ and for displaying purpose only a X-Z-cut at $y = N_y/2$, with $\sigma = 0.2$ and hence $h \rightarrow h^{0.2}$ for visibility of side-orders, is given. Magenta dashed lines mark the focus position.

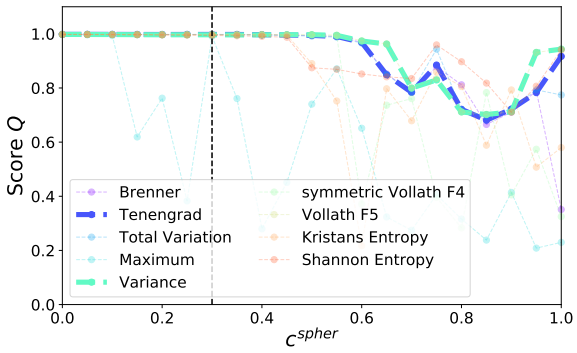
an increasing measure of spherical aberrations initially remains unchanged at the same Z-position ($f_z = N_z/2 = 32$), moves to $f_z = 33$ from $z = 12$, and then splits into 2 positions of best focus, which finally lie around positions $f_{z_1} = 28$ and $f_{z_2} = 35$. Even if the definition of a clear focus position is difficult in case of such strong aberrations, the determined split of the focus positions agrees with the theoretically expected division of the caustics, i. e. a distant focus for small angles (=E-field close to the optical axis) and a near focus for big angles (=E-field far from the optical axis) w. r. t. the objective lens arises ([94], p.21ff).

As shown in Figure B.3II, the performance of most of the metrics is uniformly high in this noise-free aberrated case and starts to drop uniformly from $c^{(\text{spher})} = 0.6$. $\Psi^{(\text{MAX})}$ at $c^{(\text{spher})} = 0.35$, $\Psi^{(\text{KEN})}$ at $c^{(\text{spher})} = 0.4$ as well as $\Psi^{(\text{SEN})}$ at $c^{(\text{spher})} = 0.45$ drop off earlier. This is due to the choice of the weighting of the score parameters Equation 2.7, the limited pixel accuracy of the calculation as well as filter-inherent differences regarding the sharpness of the expression of a maximum. The threshold-factor here was chosen to be $u = 0.14$ because otherwise the maximum-search routine would have found too

(I)



(II)



(III)

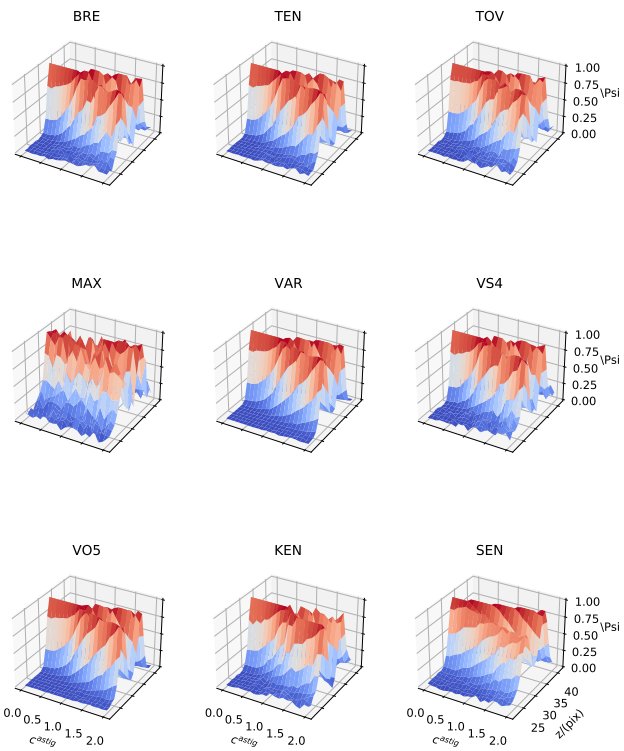
	$\tilde{q}^{\text{accu}}$	$\tilde{q}^{\text{Nfalse}}$	$\tilde{q}^{\text{range}}$	$\tilde{q}^{\text{FWHM}}$	Q
BRE	1.00	1.00	1.00	0.98	1.00
TEN	1.00	1.00	1.00	0.97	1.00
TOV	1.00	1.00	1.00	0.97	1.00
MAX	0.99	1.00	1.00	0.96	0.99
VAR	1.00	1.00	1.00	0.97	1.00
VS4	1.00	1.00	1.00	0.98	1.00
VO5	1.00	1.00	1.00	0.97	1.00
KEN	1.00	1.00	1.00	0.97	1.00
SEN	0.99	1.00	1.00	0.96	0.99

FIGURE B.3: **Spherical Aberration testing** of the sharpness metrics using *spokes2d* thin object (Figure 2.1 0a) $-$ d)) for different degrees of severity $c^{(\text{spher})} \in [0.0, 1.0]$ with steps of $\Delta c^{(\text{spher})} = 0.05$ and Z-positions, thresholded with $u = 0.14$. a) Surface plot of resulting (non-filtered) sharpness values per metric. b) Scoring Results Q for each metric Ψ with c) exemplary score-calculation display at position $c^{(\text{spher})} = 0.85$ (black dotted line).

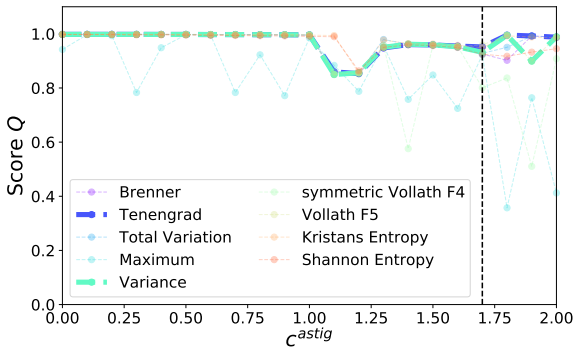
many irrelevant side maxima e. g. in the case of $\Psi^{(\text{KEN})}$ and thus biased the evaluation[1]. Around $c^{(\text{spher})} \approx 0.5$ the aberration level starts to induce a (sub-pixel) shift to the nominal focus-position which might already be detected by $\Psi^{(\text{KEN})}$ and $\Psi^{(\text{SEN})}$, but not yet by the $\Psi^{(\text{VO5})}$ (used for reference focus-position estimation) or manually, thus inducing decreasing levels of precision as compared to the other metrics. This can be understood

1 While it is tempting to only use the biggest maximum it is not a reasonable option here as the behavior of the different sharpness metrics w. r. t. to noise and aberration shall be studied and therefore eventual side-maxima of the sharpness Z-curve are relevant. Further, the algorithm shall be ready to use for multi-focus applications.

(I)



(II)



(III)

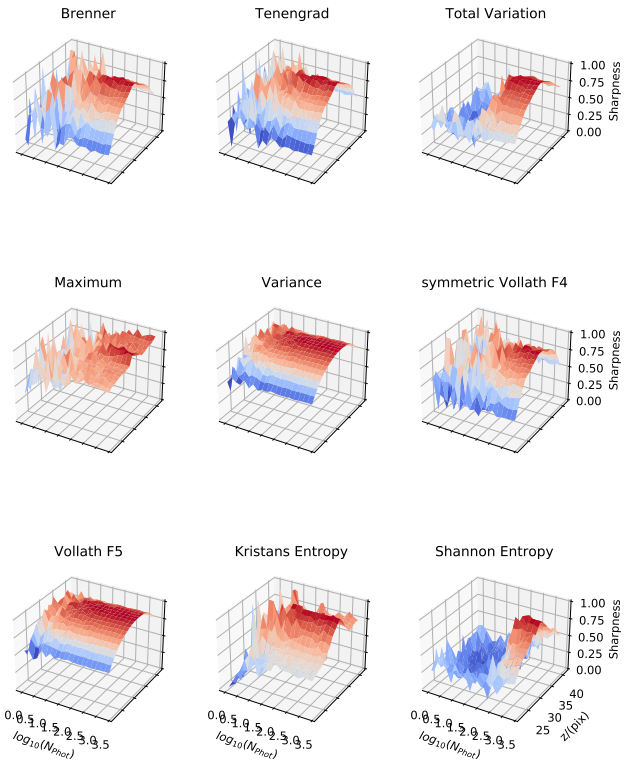
	$\tilde{q}^{(\text{accu})}$	$\tilde{q}^{(\text{Nfalse})}$	$\tilde{q}^{(\text{range})}$	$\tilde{q}^{(\text{FWHM})}$	Q
BRE	0.86	1.00	1.00	0.93	0.93
TEN	0.92	1.00	1.00	0.92	0.95
TOV	0.87	1.00	1.00	0.93	0.93
MAX	0.86	1.00	1.00	0.93	0.92
VAR	0.88	1.00	1.00	0.93	0.93
VS4	0.81	0.88	0.68	0.98	0.80
VO5	0.93	1.00	1.00	0.92	0.96
KEN	0.91	1.00	1.00	0.92	0.95
SEN	0.86	1.00	1.00	0.93	0.92

FIGURE B.4: **Astigmatic Aberration testing** of the sharpness metrics using *spokes2d* thin object (Figure 2.1 $0a) - d)$) for different degrees of severity $c^{(\text{astig})} \in [0.0, 2.0]$ with steps of $\Delta c^{(\text{astig})} = 0.1$ and Z-positions, thresholded with $u = 0.14$. a) Surface plot of resulting (non-filtered) sharpness values per metric. b) Scoring Results Q for each metric Ψ with exemplary score-calculation display at position $c^{(\text{astig})} = 1.6$ (black dotted line).

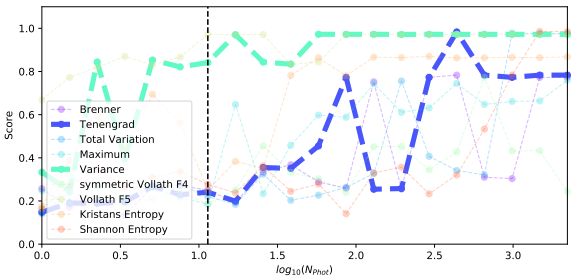
as false positive result as these metrics seem to perform better than the reference data-set, but the calculated score is lower. The focal positions than finally split into two peaks with increasing contrast w. r. t. to surrounding Z-slices and finally manifest at position and contrast that seem to yield the same detectability for most (5 out of 9) of the metrics.

In case of astigmatism Figure B.4 nearly all metrics perform equally well and mostly range around the maximum score. A deviation from maximum scoring can be observed for all metrics at $c^{(\text{astig})} \geq 1$, where a wave-front deformation of bigger than 1λ ($c^{(\text{astig})} \geq 1$) induces a shift and split of the focal-plane that can be detected reliably from $c^{(\text{astig})} = 1.3$

(I)



(II)



(III)

	$\tilde{q}_{accu}$	$\tilde{q}_{Nfalse}$	$\tilde{q}_{range}$	$\tilde{q}_{FWHM}$	Q
BRE	0.25	0.19	0.13	0.75	0.25
TEN	0.08	0.44	0.17	0.99	0.24
TOV	0.07	0.44	0.13	0.97	0.22
MAX	0.03	0.38	0.10	0.99	0.19
VAR	0.95	0.94	0.56	0.83	0.84
VS4	0.03	0.44	0.06	0.98	0.19
VO5	0.98	1.00	1.00	0.79	0.97
KEN	0.07	0.62	0.11	0.99	0.25
SEN	0.05	0.75	0.17	0.99	0.28

FIGURE B.5: **3D Noise testing** of the sharpness metrics using Obj3d sparse thick object (Figure 2.1 3a) $-$ d)). a) Surface plot display of resulting (non-filtered) sharpness values per metric for different Photon-counts N_{Phot} and Z-positions, thresholded with $u = 0.07$. b) Scoring Results Q for each metric Ψ with exemplary score-calculation display at position $b = 0.3$ (black dotted line) and hence $N_{phot} \approx 11$.

on. The limitations in pixel-accuracy and sensitivity to this split and shift lead to the slight performance variation of the individual metrics. As usual $\Psi^{(\mathrm{MAX})}$ is rather unreliable and $\Psi^{(\mathrm{VS4})}$ is strongly influenced due to side-maxima.

B.3 SCORE EVALUATION OF A 3D SAMPLE

All simulations performed in Section 2.3 evaluated the usability of the scoring algorithm to hold the focus position in the case of a thin sample in different noise and aberration scenarios. Now, testing the metrics and scoring on a sparse but thick in-silico sample *obj3d* (Figure 2.1 3a) $-$ d)), follows and is displayed in Figure B.5. Here, not 1 but many slices are intendend to be *in-focus* at the same time. Therefore, the procedure here is now

different: In the case of a thick 3D sample, the metrics are used to determine a Z-curve characteristic of the sample distribution. The sample *obj3d* used for the simulation has the most features in the central Z-slice, which results in the largest score Q when applying the sharpness metrics. Starting from the central Z-slice, the amount of pixels that are part of an edge, decreases. The sample intersects a fluorescent rod at $z = z^{(\mathrm{F})}$. However, since this rod has very little features its effect on the result of the sharpness metrics is rather small. Therefore, the presented scoring algorithm can be directly applied to this thick sample as well, since there is a central maximum and none, or only one side maximum (due to the fluorescent rod) is expected. For the calculation of Q, the plane $z = z^{(\mathrm{o})}$ was allowed as the only focus and no side maxima/foci.

$\Psi^{(\mathrm{VAR})}$ and $\Psi^{(\mathrm{VO5})}$ achieve a reliable scoring of $Q \geq 0.8$ for $log_{10}(N_{Phot}) \geq 0.7$. For higher photon-counts $\Psi^{(\mathrm{TEN})}$ ($log_{10}(N_{Phot}) > 1.9$) and $\Psi^{(\mathrm{KEN})}$ ($log_{10}(N_{Phot}) > 1.5$) perform reliable ($Q \approx 0.8$), while most of the others need rather high photon levels to achieve the same scoring.

C

APPENDIX: JITTER CORRECTION

C.1 HAF AND SAF COMPARISON

A brief summary of the findings of Bian *et al.* [49] follows.

In the field of Hardware Autofocus (HAF), confocal pinhole detection and triangulation with line-detector seem to have gained acceptance in principle. In the case of confocal pinhole detection, the reflection signal is focused on a pinhole via a beam splitter with an extra lens and then integrated with a photo detector (e. g. PMT, Avalanche Photo Diode (APD)). The reflection at different boundary layers, e. g. the air-glass transition in case of air lenses, is used as reference and a whole Z-stack has to be recorded to obtain a reliable statement about the type and position of the reflection. In the triangulation method, an additional illumination and detection unit is attached to the microscope. Here, only half the BFP of the objective is illuminated, such that the interfacial back-reflection (e. g. at the front of the cover-slip) has a lateral offset to the illumination. The detection path is adjusted such that the back-reflection is centered on a quadrant photo-diode and the difference signal resulting from the measured photovoltages disappears. If there is an axial offset of the sample, there is a lateral offset in the diode array plane and the difference signal is different from 0. A further development of this approach includes an additional lens group for analog correction of the optical offset (=desired focus position) set by the user to the reference focus position as well as a CCD line sensor with which the lateral profile can be recorded per axial position. In case of a defocus, the PSF maximum shifts towards the defocus direction. The upgraded (i. e. line-sensor and lens-group added) method is more expensive and more complex to integrate than the diode-array based method, but allows to adjust for arbitrary focus position offsets prior to the measurement. Thus, HAF methods are not automatically compatible with arbitrary sample chambers and measurement methods, but can mostly be adapted to them. However, due to their sample independence they have the potential to be more stable than Software Autofocus (SAF) methods. Since HAF e. g. only requires the detection (and degree) of a change in

the position of the center-of-mass of an intensity distribution on the additional detector(-array), very fast switching times with very low data streams can be achieved here, and thus even an implementation despite severely limited computing power.

While HAF requires modifications of the existing experimental setup, a SAF can be implemented directly into any automated measurement system with computer-based control on the software side. SAF methods operate directly on the measured image data and are thus highly dependent on the sample, recording quality, and noise. In addition, inherently much more data must be recorded and processed, making the application of complex algorithms or methods with large data stacks on limited hardware even partially unusable. SAF approaches can be divided into classical (=focus map) and real-time methods. Classical methods do not require any changes to the optical setup and can therefore be used directly for any optical system. Real-time methods are based on the metrics and algorithms of the classical methods and can be seen as hardware optimizations for live applications of these. Five different methods Methods presented in [49] use: 1) a second sensor to parallelize the sharpness metric evaluation for image acquisition, 2) a beam-splitter array in the finite beam-path to evaluate different Z-regions in parallel, 3) a tilted sensor in the infinite beam-path to image different Z-regions in parallel on one sensor, 4) phase detection which cuts out two opposing regions of the focussing sensor lens' BFP and estimates axial shift from the offset of the focused images of the filtered regions and 5) illumination from different opposite spatial points and angles using e. g. dual-led such that axial defocus leads to lateral defocus and thus double images. Furthermore, deep-learning based approaches exist, which either determine the distance to the optimal focus position from a defocus-stack or calculate an optimal (virtual) focus image.

C.2 ALGORITHMS IN PSEUDOCODE

For the algorithms described in Section 3.2, pseudocode notation is given in this section to achieve an overview of the code flow. Algorithm 1 describes the *in-silico* jittered data generation and Algorithm 2 describes the dejitter algorithm.

Algorithm 1 In-Silico Datastack Generation.

drao ≡ „draw random amount of“
procedure *generate_data*$(N, NA, \lambda^{(\mathrm{ex})}, \lambda^{(\mathrm{em})}, d^{(\mathrm{pitch})})$
    Generate PSF h $= h^{(\mathrm{ex})} \cdot h^{(\mathrm{det})}$ with $\sum h = 1$
    Generate Sample S create $N^{(\mathrm{z})}$ copies of $S^{(\mathrm{2D})}$
    Calculate Image M $= \mathcal{P}\{h \otimes S\}$
    Add jitter $M^{(\mathrm{jit})} = generate\_jitter(M, \ldots)$
end procedure
procedure *generate_jitter*$(M, N^{(\mathrm{lines})}, \Delta^{(\mathrm{lines})}, s^{(\mathrm{max})}, d^{(\mathrm{block})})$
    Get Amount $\mathbf{N_l^{(lines)}}$ drao lines to be shifted from $[0, N^{(\mathrm{y})} \cdot N^{(\mathrm{lines})}]$ per Z-slice
    for all $N_l^{(\mathrm{lines})}$ **do**
        Get Lines $\mathbf{L_{lm}}$ drao line-numbers from range $[0, N^{(\mathrm{y})} \cdot \Delta^{(\mathrm{lines})}]$
        Get shift $\mathbf{s_{lm}}$ drao shift for each line within $[-N^{(\mathrm{x})} \cdot s^{(\mathrm{max})}, N^{(\mathrm{x})} \cdot s^{(\mathrm{max})}]$
        Get retardance $\mathbf{s_{lm}^{(ret)}}$ drao lines around selected lines L_{lm} from $[0, N^{(\mathrm{x})} \cdot d^{(\mathrm{block})}]$
        Add retardance apply (smoothened) retardance along Y-direction
        Apply X-Shiftmap using the generated (random retardance) shiftmap
    end for
end procedure

C.3 PARAMETER SEARCH FOR GS DECONVOLUTION

Like is demonstrated in Section 3.3 in case of Total Variation (TV) regularization a parameter search Figure C.1 in case of using Gradient Square (GS) regularization on the same data as used in Figure 3.4 yields a maximum $C = 0.88$ at $\gamma^{(\mathrm{GS})} = 10^{-3.5}$. It is reasonable to assume that although the regularization maximum has not yet been reached, the analyzed region is close to it. The patterning visible in (d-g) is a residual of the tiling routine and can be reduced by a different choice of padding parameters.

Algorithm 2 Movement Correction

procedure *dejitter_and_shift*($M_{lmn}^{(\text{raw})}$,...)
  Preprocess $\mathbf{M_{lmn}} \equiv pre(M_{lmn}^{(\text{raw})})$
  Correct Axial Drift $\mathbf{M_{lmn}^{(\text{dcor})}}, M_{mn}^{(\text{ref})} = correct\_zdrift(M_{lmn})$
  Get Dejitter Shiftmap $\mathbf{C_{l'm'}^{(1D)}} = dejitter\_with\_reference(M_{lmn}^{(\text{dcor})}, M_{mn}^{(\text{ref}')})$
  Apply Shifts $\mathbf{M_{l'm'n'}^{(\text{dej})}} = \mathcal{FS}(M_{l'm'n'}^{(\text{dcor})}, C_{l'm'}^{(1D)})$
  Deconvolve $\mathbf{M_{lmn}^{(\text{decon})}}$ using $tiled_{deconv}(M_{l'm'n'}^{(\text{dej})}, h, \vec{t})$
end procedure
procedure *correct_zdrift*(M_{lmn})
  Calculate Reference $\mathbf{M_{mn}^{(\text{ref})}}$ Z-stack using 'mean','median' or 'SG'
  Get slicewise Shifts $\mathbf{s}^{(\text{lin})}$ via $\underset{x}{argmax}(M_{lmn} \circledast M_{mn}^{(\text{ref})})$
  Apply Shifts $\mathbf{M_{lmn}} = \mathcal{FS}(M_{lmn}, s^{(\text{lin})})$ using Fourier-shift theorem
end procedure
procedure *dejitter_with_reference*($M_{lmn}, M_{mn}^{(\text{ref})}$)
  Damp Edges $\mathbf{M_{l'm'n'}^{(\text{damp})}}$ of M_{lmn} by adding onto outside
  Get Correlation Map $\mathbf{L_{l'm'n'}}$ via $M_{l'm'n'} \circledast_x M_{m'n'}^{(\text{ref})}$
  Multiply Window $\mathbf{L_{l'm'n'}^{(\text{win})}} = \Omega_{l'm'n'} \cdot L_{l'm'n'}$ ('Gauss' or 'Rect' or 'None')
  if use Savitkzy-Golay pre-filter **then**
    apply SG-filter $\mathbf{L_{l'm'n'}^{(\text{win})}} = SG(L_{l'm'n'}^{(\text{win})})$
  end if
  Find linewise Maxima $\mathbf{L_{l'm'}^{(\text{max})}} = \underset{x}{argmax}(L_{l'm'n'}^{(\text{win})})$
  if use center of mass **then**
    Create Binary Mask $\mathbf{B_{l'm'n'}}$ set 1 at found $L_{lm}^{(\text{max})}$ and 0 elsewhere
    Dilate Mask $\mathbf{B_{l'm'n'}^{(\text{dil})}} = dilate(B_{l'm'n'})$
    Apply to Correlation Map $\mathbf{L_{l'm'n'}^{(\text{dil})}} = B_{l'm'n'}^{(\text{dil})} \cdot L_{l'm'n'}^{(\text{win})}$
    Offset Correct $\mathbf{L_{l'm'n'}^{(\text{dil}')}}$
    Calculate 1D center of Mass $\mathbf{C_{l'm'}^{(1D)}}$
  else
    Use linewise Maxima $\mathbf{C_{l'm'}^{(1D)}} \equiv L_{l'm'}^{(\text{max})}$
  end if
end procedure

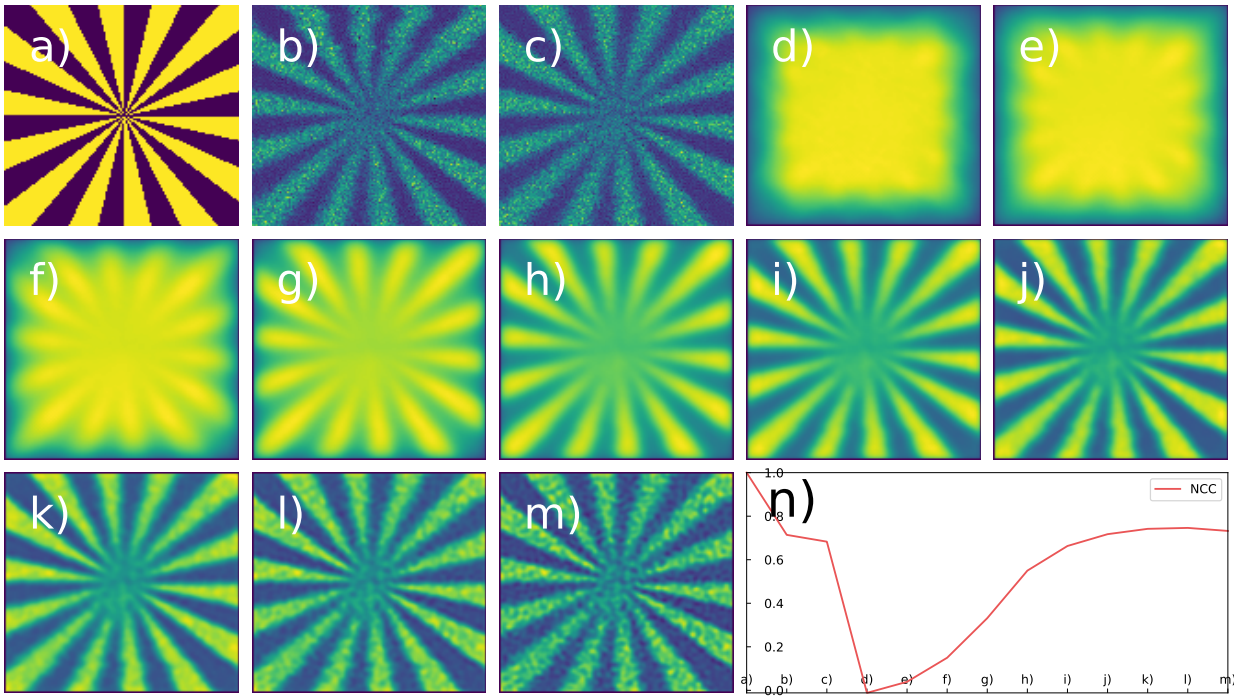


FIGURE C.1: **GS-Deconvolution lambdas.** Results $\mathbf{M}^{(\text{decon})}_{\mathbf{l}=\mathbf{z}^{(\text{o})},\mathbf{mn}}$ of processing and tiled deconvolution of a) 3D obj, b) jittered image, c) dejittered image d-m) using GS regularizer with $\gamma^{(\text{GS})} \in 10, 10^{0.5}, \ldots, 10^{-3.5}$.

D

APPENDIX: THICK SLICE UNMIXING

D.1 BEHAVIOUR OF ISM PSF

The system PSF h is shown in Figure D.1II for a series of increasing x-shifts of the detection PSF, with $\delta x = 0.083\,AU$. The amplitude decreases rapidly in this case. A cut through $y = 15$ is shown for the first 8 shifts (a-h,Figure D.1II) in Figure D.1I. The PSF becomes more asymmetric for increasing shifts dx. Thus, e. g. the left sidelobe for $dx = 0.66\,AU$ is larger than the right sidelobe and both in particular are larger than the central maximum of the corresponding PSF.

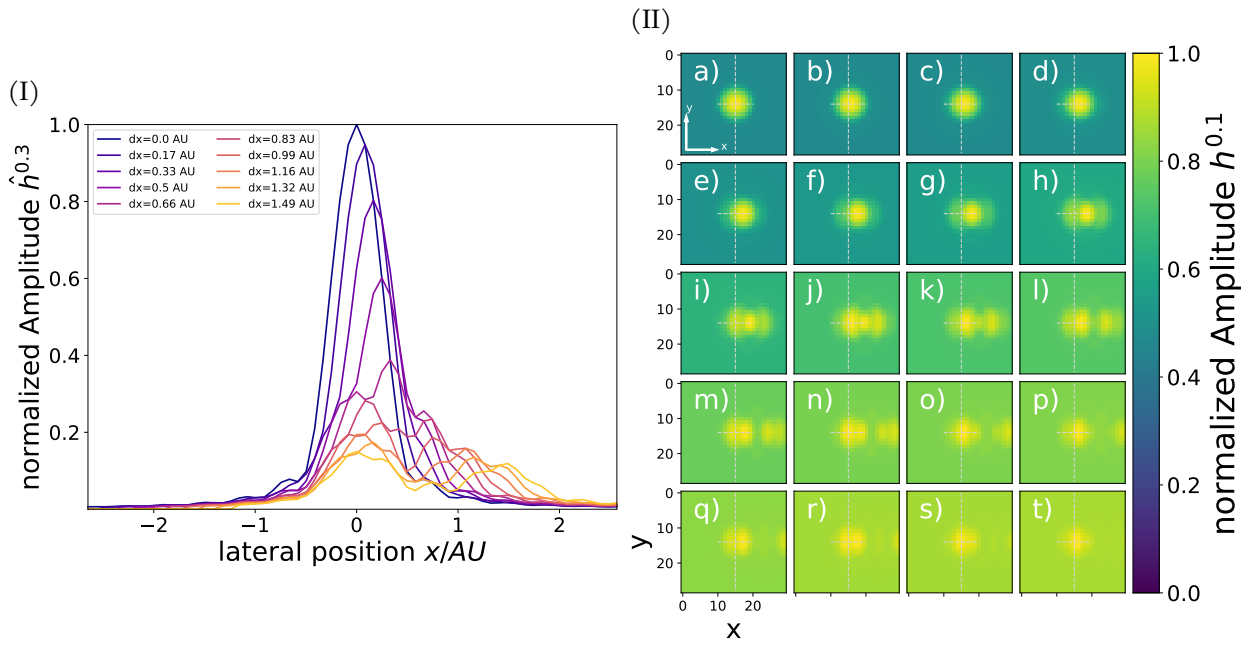


FIGURE D.1: **DSAX FWHM Shift Dependence.** (I) Linecuts through lateral ideal PSF at $y = y^{(c)}$ for different detectors spaced with dx. (II) According shifted ideal 2D PSFs. While a-h) are 2D representations of the respective linecuts in (I), i-t) are the further spaced pinholes with individual increase of spacing by $\Delta x = 0.17\,AU$. Images are individually max-normalized.

D.2 RANK OF THE OTF MATRIX

In fluorescence Image Scanning Microscopy (ISM) the PSF h is a real-valued, (theoretically unbounded) function and thus its Fourier transform $\tilde{h}$ satisfies the continuous Hermitian property $\tilde{h}^*(k) = \tilde{h}(-k)$. In case of the presented forward model Equation 4.1 the discrete OTF $\tilde{h}$ not necessarily fulfills $\tilde{h}^T = \tilde{h}^*$ due to the chosen sampling parameters and dataset sizes. Yet, if the OTF $\tilde{h}_{ak_y}$ $(\forall k_l, k_n)$ is square it is diagonalizable, has real eigenvalues, and the lower triangular matrix can be generated from the upper triangular matrix by complex conjugation [95]. From diagonalizability it follows that the trace of the square matrix $\tilde{h}_{ak_y}$ is the sum of its eigenvalues and thus its rank, if symmetric, can be estimated via:

$$rank(\tilde{h}) \geq \frac{\left(tr(\tilde{h})\right)^2}{tr(\tilde{h}^2)} \tag{D.1}$$

It follows that its rank is not automatically bounded by half its number of lines.

D.3 OUT-OF-FOCUS REJECTION CAPABILITY COMPARISON

For the reconstruction of the in-focus slice $(z^{(0)})$, different methods are compared in Figure D.2 for 3 different choices of detector pinhole masks (Figure D.2II a,g,k). While direct sum $M^{(\text{sum})} = \sum_a^{N^{(a)}} M_a$ (Figure D.2Ia-c) and Pixel Reassignment (PiRe) (Figure D.2Id-f) can be used without additional information, the remaining methods always require the effective system PSF $h^{(\text{ISM})}$. Weighted Averaging in Fourier Space (WAVG) (Figure D.2Ig-i) requires an OTF support boundary parameter $(\epsilon^{(\text{mask})})$. Noise normalized WAVG (Figure D.2Ij-l) and Wiener Deconvolution (WD) followed by summation (Figure D.2Im-o, see Section A.19) additionally require a regularizing parameter for the normalization step (γ). Thick slice Unmixing (TU) (Figure D.2Ip-r) additionally allows choosing the number of Singular Values (SV)s to use (directly via $N^{(\tilde{\xi}\,\text{lim})}$ or indirect via $\Delta\tilde{\xi}^{(\text{rel})}$). Finally, 2D deconvolution using the in-focus slice and a 2D-PSF (Figure D.2Is-u) as well as 3D deconvolution using the full 3D-stack and 3D-PSF (Figure D.2Iv-x) with selection of the central slice $z^{(0)}$, which require a variety of other parameters depending on the choice of the optimizer, are depicted. Provided that a constant shift factor is used for the PiRe, the

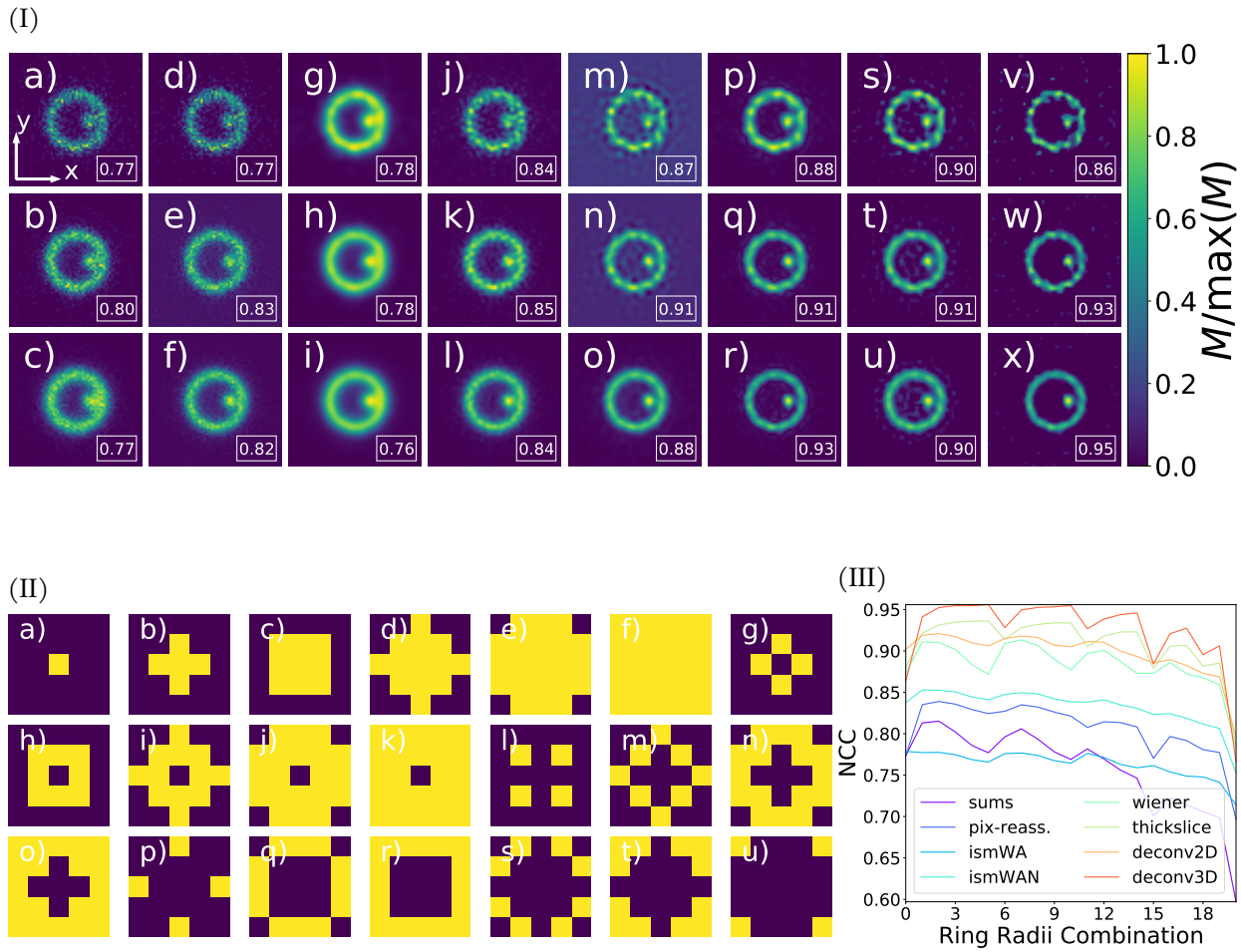


FIGURE D.2: **ooF Rejection Comparison.** (I) comparison of ooF rejection and image reconstruction quality for different methods (columns) and different pinholes: upper row uses IIa), middle row uses IIg) and lower row uses IIk). Methods used are a-c) $M^{(\text{sum})}$, d-f) $M^{(\text{ISM})}$, g-i) $M^{(\text{wavg})}$, j-l) $M^{(\text{nwavg})}$, m-o) $M^{(\text{wd})}$, p-r) $M^{(\text{tu})}$, s-u) 2D $M^{(\text{mDEC})}$, v-x) 3D deconvolution $M^{(\text{mDEC})}(z = z^{(\text{o})})$ with selection of central slice. Normalized Cross Correlation (NCC) similarity is displayed in lower right corner of each tile. (II) Available masks for the 5×5 detector pixels. Active pixels are marked yellow, inactive are black. (III) NCC results for all different methods (I) and all available masks (II).

column-wise listing of a-v) can also be read as an ascending sorted list w. r. t. necessary processing time, system resources and algorithmic complexity (for this, however, column 5 must be at the 3rd position).

The reconstruction quality quantified by the NCC depends on the number and position of the used detector pinholes (Figure D.2III). Depending on the imaging device more/bigger/faster pixels might cost exponentially more money or time and additionally generates more data that needs to be processed. A balance between costs/computational overhead and gain in reconstruction quality, especially if limited hardware is used for processing, is thus desired. While an increasing number of used pinholes improves the reconstruction quality of TU multi-view DEConvolution (mDEC), the NCC maximum for the other methods is mostly reached for a balance case between number of pinholes and distance to the central pinhole. This is the case for selection of the first pinhole ring (e. g. Figure D.2IIb,c). TU and mDEC proceed very similarly and the average absolute deviation is $\mathbb{E}\left\{C^{(\mathrm{TU})}/C^{(\mathrm{mDEC})}|\right\} \approx 2\%$. It can be further concluded that more pinholes lead to a steadily higher NCC and thus better reconstruction quality. Nevertheless, especially for pinmasks (Figure D.2II) b-f), the relative NCC changes are in the range of 1.55%, but the numerical effort increases up to a factor of 5 (compare b and f) given that the used amount of active pinholes is 5 times bigger for f instead of b. Here, the jump from pinhole configuration a) to b) is much more rewarding. For a visual comparison in the case of TU see Figure D.4. Thus, it is reasonable to assume that for the parameters used here, reconstruction with pinhole configuration b) represents a good balance of numerical cost and reconstruction quality[1].

D.4 DETECTOR PIXEL PITCH

The obvious question is whether there is an NCC optimal detector pixel pitch in the context of w. r. t. the number of pinholes used as well as the lateral size of the system PSF (Figure D.3). For this purpose, different lateral distances Δs between the detector pinholes (unidistal = equal distances in both lateral directions) are simulated and given in sample coordinates, thus $1pix = d^{(\mathrm{SPL,lat})} \approx 42\,nm \approx 0.085 \times d^{(\mathrm{1AU,det})}$. A selection of the effective PSFs of the 25 pinholes is shown for $\Delta s \in [1, 3, 5, 7]\,pix$ (Figure D.3I). Despite increasing lateral pixel spacing, the pixel pitch was assumed unchanged to be δ-like and thus any pixel shape specific properties were not further included. With increasing spacing, the

1 Note, that this result further depends on the detector pixel pitch which here was $2/3\,d^{(\mathrm{1AU,eff})}$ of the effective system PSF.

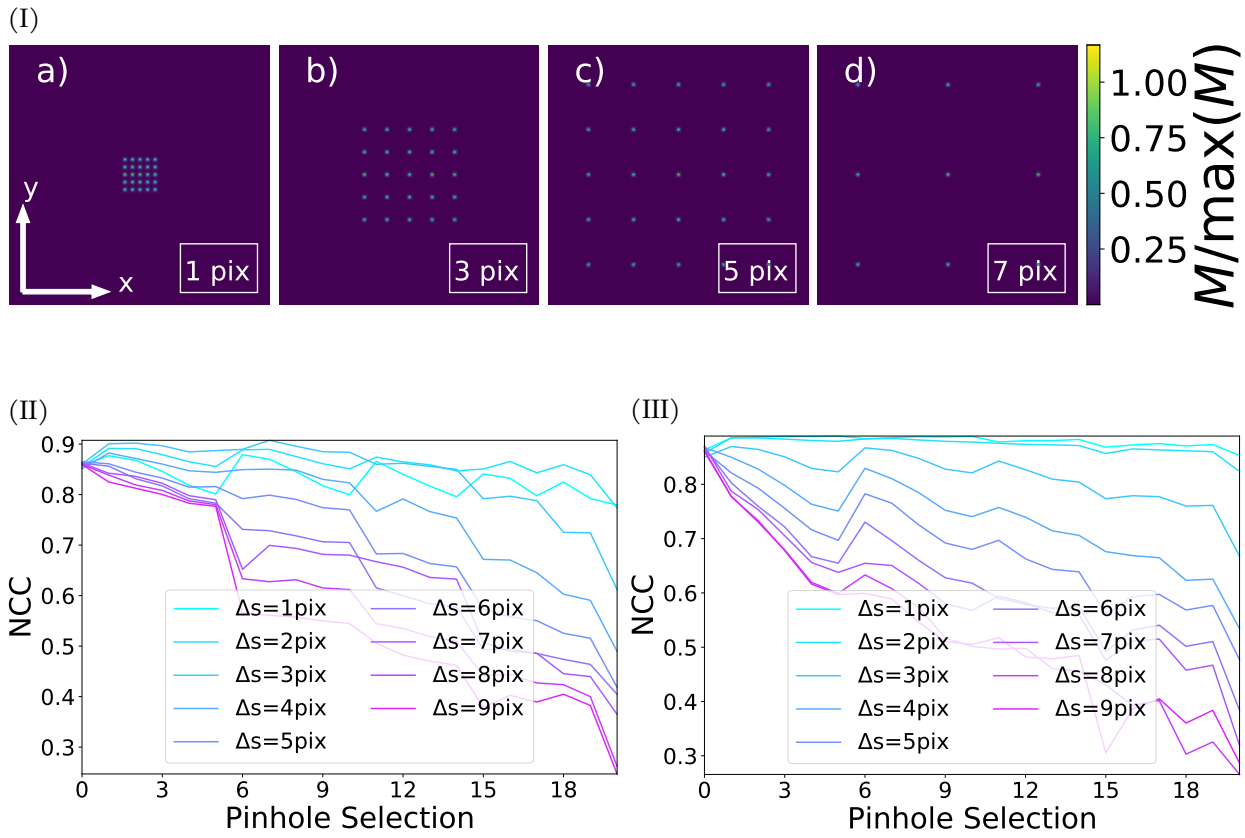


FIGURE D.3: **Detector Pixel Pitch** for unirectangular detector-pixelgrid with unit-vectors $\hat{\vec{x}} = [\Delta s, 0]^T, \hat{\vec{x}} = [0, \Delta s]^T$ for (II)Thickslice-Unmixing and (III) 3D Deconvolution. (I) Effective PSF h for all detector-pinholes for in-focus slice ($z^{(o)}$) and at different spacings Δs = a) 1, b) 3, c) 5, d) 7 pix. Spacings are displayed in white box on lower right of each tile. PSFs are normalized to their total sum and rescaled by $(log_{10}(h))^{0.1}$.

amplitude decreases more radially symmetrically and there may also be amplification of sidelobes due to overlap of $h^{(\mathrm{ex})}$ with sidelobes of $h^{(\mathrm{det})}$ (and vice-versa) and thus a strong change of the effective PSF. For each detection geometry the forward model is calculated again which always results in a local-varying noise distribution (with the same applied statistics), which is especially obvious in the NCC of the reconstructions for detector pinhole selection $a^{(o)}$ (=central pinhole).

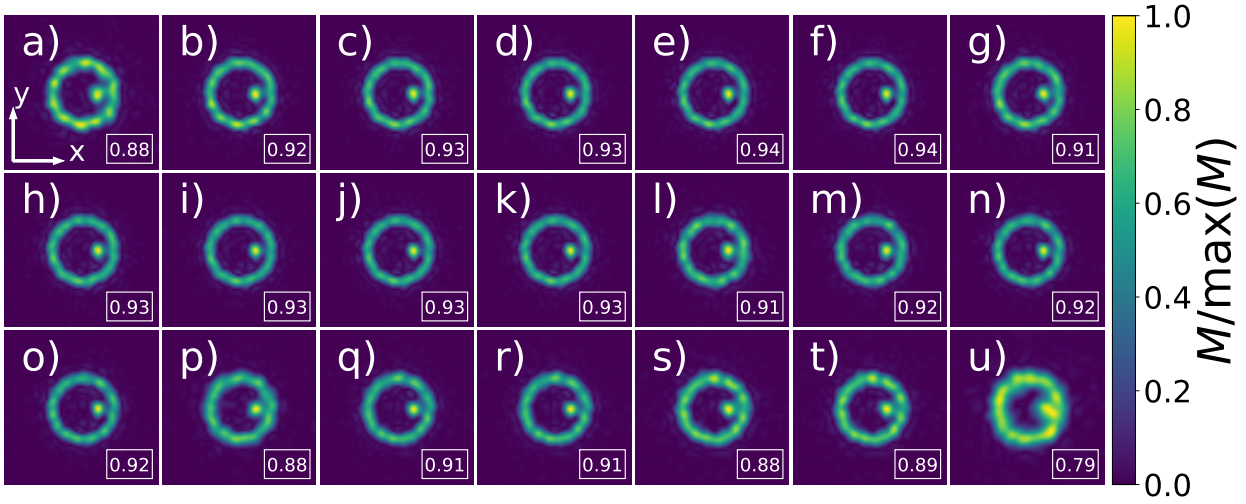


FIGURE D.4: **Thickslice NCC Visualization** for visual comparison of the NCC optimal result for each individual pinmask as given in Figure D.2II.

D.5 PIXEL EFFECTIVITY IN TU

Camera systems with many pixels are available directly and relatively inexpensive, but their pixel dwell time and register readout speed is on average significantly slower than from PMT or APD arrays. In order to reduce the measurement times for recording the multi-view ISM data, the transition to PMT arrays or the construction of a dedicated detector array is an obvious choice. In contrast to a camera produced in large quantities, the cost of such a custom array e. g. scales rapidly with the number of individual detector elements, so a brief analysis of how many pixels are suitable using a rectangular detector geometry will be given. The pinhole masks used are shown in Figure D.2II.

In the comparison number of pinholes over achieved NCC value, the central pinhole (Figure D.2IIa) is unbeatable, but the configurations are subjectively a better choice, since the achieved NCC is close to the maximum achievable NCC after TU processing in this comparison and on top of that the axial distribution of the sample can be reconstructed better due to the significantly larger view-angles.

Within this work, two different existing detector geometries were used for reconstruction, but construction of an own detector according to the simulated findings is left for future work.

N_z	$S_{cpu}[MiB]$	R_{cpu}^{gpu}	$S_{gpu}[MiB]$	$S_{gpu}[\%]$	$T_{gpu}[°C]$	$t_{ext}[s]$	$t_{dec}[s]$
0	0.00	3.00	0.00	0.15	71.00	0.00	0.00
2	0.12	1424.00	178.00	8.90	71.00	2.80	0.90
4	0.25	744.00	186.00	9.30	71.00	1.70	1.40
8	0.50	404.00	202.00	10.10	72.00	3.60	3.10
16	1.00	234.00	234.00	11.70	74.00	6.60	5.70
32	2.00	149.00	298.00	14.90	75.00	13.10	11.30
64	4.00	106.50	426.00	21.30	77.00	25.80	22.50
128	8.00	85.25	682.00	34.00	81.00	50.10	43.50
256	16.00	74.62	1194.00	59.60	83.00	101.80	87.40
512	32.00	60.16	1925.00	96.10	84.00	195.80	166.10

TABLE D.1: **mDEC GPU-RAM usage test.** Input Image size $N = [N^{(z)}, 128, 128]$, necessary storage for image on RAM S_{cpu} as datatype is $float32$, storage used on GPU for processing S_{gpu}, relative storage usage of GPU to total on GPU-available storage $S_{GPU}[\%]$, temperature of GPU T_{gpu}, $R_{cpu}^{gpu} = S_{gpu}/S_{cpu}$ ratio of memory on GPU vs RAM, time needed for running the full optimization routine t_{ext} (excludes graph-buildup time and storage allocation preparation on GPU) and time needed for on-GPU optimization including copy-process to↔from GPU t_{dec}. First row is only for displaying the „empty"GPU-state before first mDEC process.

D.6 PERFORMANCE COMPARISON OF TU AND MDEC

While the previous comparisons have demonstrated the reconstruction quality of TU compared to mDEC, the memory and processing time requirements will be briefly discussed here. The properties of the test system are given in Section 1.8.

Only the central pinhole for different scanning stack sizes $N = [N^{(z)}, 128, 128]$ and a relative border region $\Delta b = [\Delta b_z, \Delta b_y, \Delta b_x] = [0\%, 10\%, 10\%]$ were used as the data set for mDEC. The border is thereby plotted symmetrically around the existing data points, i. e. are appended to the lateral slices on both sides about $6pix$ each. The reconstruction was fully 3D, i.e. from a 3D measurement dataset of $N^{(p)} = N^{(z)}$ to a reconstructed 3D object of axial size $N^{(z)}$. The performance measurements for the different input datasets are displayed in Table D.1. S_{cpu} marks the necessary memory of an array of size N of

N_z	$N_{a'}$	$S_{cpu}^{(\text{theo})}[MiB]$	$S_{cpu}^{(\text{resid})}[MiB]$	$S_{cpu}^{(\text{peak})}[MiB]$	$t^{(\text{OTFu})}[s]$	$t^{(\text{unmixing})}[ms]$
1	5	2	481	1008	11	8
2	10	6	487	1008	9	7
4	20	23	508	1008	10	13
8	40	86	588	1008	11	36
16	80	331	898	1384	12	126
32	160	1302	2117	4049	15	468
64	320	5164	6955	14659	18	2108

TABLE D.2: **Thickslice RAM usage test.** Input Image size $N^{\widetilde{M}} = [N^{(a')}, 1, 128, 128]$ and $N^{\tilde{h}} = [N^{(a')}, N^{(z)}, 128, 128]$. $S_{cpu}^{(\text{theo})}$ theoretically expected size on RAM of all recon-arrays $(\widetilde{M}, \tilde{h}, \tilde{h}^{\dagger}, \hat{S}^{(\text{TU})})$ together, $S_{cpu}^{(\text{resid})}$ measured resident RAM-usage, $S_{cpu}^{(\text{peak})}$ peak usage during reconstruction including all array-overheads and $t^{(\text{unmixing})}$ time needed for unmixing. Datatype for real-space data was $float32$, for fourier-space $complex64$.

datatype $float32$ on system Random-Access Memory (RAM), S_{gpu} Graphics Processing Unit (GPU) memory used, $S_{GPU}[\%]$ the relative used memory of GPU w. r. t. to total available GPU memory, T_{gpu} GPU temperature, $R_{cpu}^{gpu} = S_{gpu}/S_{cpu}$ the ratio of used GPU memory and system RAM, t_{ext} the time needed for running the full optimization routine (excluding graph-buildup time and storage allocation preparation on GPU) and t_{dec} the time needed for on-GPU optimization including copy-process to↔from GPU. The memory overhead for creating the Tensorflow graph (Section 1.8) used for the GPU based mDEC can be read directly from the storage used on GPU $S_{gpu}(N^{(z)} = 2)$, since the memory required for the dataset is still significantly smaller even with all 13 copies needed for the gradient calculations. Here the ratio of memory used on the GPU and on system RAM is $R_{cpu}^{gpu}/13 \approx 110$. The maximum dataset possible on the test system's GPU is $N^{(\text{gpu,max})} = [512, 128, 128]$ pixels of $float32$ datatype in real space and the processing time required for this is $\approx 166\,s$. As the data set size doubles, the relative memory overhead reduces and the processing time t_{dec} roughly doubles.

The performance analysis of TU for a complete data set, i. e. for $N^{(\widetilde{M})} = [N^{(a')}, 1, 128, 128]$ and $N^{(\tilde{h})} = [N^{(a')}, N^{(z)}, 128, 128]$, is shown in Table D.2. Here $S_{cpu}^{(\text{resid})}$ measured resident RAM-usage, $S_{cpu}^{(\text{peak})}$ peak memory usage during reconstruction including all array-

overheads and $t^{(\text{unmixing})}$ the time needed for the unmixing part of the TU routine. Since for the test $N^{(\text{z,meas})} = N^{(\text{z})}$ was chosen the necessary memory for the determination of the inversion matrix scales with $O\left(\left[N^{(\text{z})}\right]^2\right)$. The theoretical expected memory requirement $S^{(\text{theo})}$ for the algorithm for the four arrays necessary for reconstruction, the image $\widetilde{M}$, the OTF $\tilde{h}$, unmixing matrix $\tilde{h}^{\dagger}$, and the reconstructed object $\hat{S}^{(\text{TU})}$, is above the average (via the processing routine) memory $S^{(\text{resid})}$ allocated to RAM starting at $N^{(\text{z})} = 16$, which is due to both sectional memory frees and partial swaps to cache. For $N^{(\text{z})} \leq 8$ the size of the compiler including all program overheads necessary for the used transformations can be read off ($S^{(\text{resid})} \gg S^{(\text{theo})}$). The approximately 4× overhead of the maximum RAM required during the algorithm runtime $S^{(\text{peak})}$ comes from the necessary transformations and regularization masks. The necessary time $t^{(\text{OTFu})}$ for computing the unmixing matrix is dominated by the lateral image size for $N^{(\text{z})} < 16$. For the final time $t^{(\text{unmixing})}$ necessary to reconstruct the image data, after determining the unmixing matrix $\tilde{h}^{\dagger}$ once, $t^{(\text{unmixing})} \ll t^{(\text{OTFu})}$ holds. For $N^{(\text{z})} = 64$ even $t^{(\text{unmixing})}/t_{dec} \approx 0.09$ holds. Note that $t_{dec} \propto O\left(N^{(\text{z})} \log N^{(\text{z})}\right)$ and $t^{(\text{unmixing})} \propto O\left(\left[N^{(\text{z})}\right]^2\right)$.

Typical ISM mDEC stacks are of the order $N = [32, 1024, 1024, 40]$, in which case TU can be used with significant time gains over mDEC.

E

APPENDIX: DSAX-ISM

E.1 MOTIVATION

Increasing the transferable information through the optical system by saturating the fluorescent sample was already analysed in the basic case of confocal microscopy [96]. In their work not only a way of describing the relation between illumination and saturated fluorescent emission, but the weak illumination radiances that are already sufficient to saturate the sample are mentioned. Since then, the approach has been further refined and analyzed in various techniques. For example, in saturated patterned excitation microscopy (SPEM), in which higher orders (in addition to the linear excitation grid) can be generated for nonlinear structured illumination reconstruction due to the nonlinear interaction with the sample, thus extending the $supp(OTF)$ [97], [98]. The resolution gain in these articles is achieved by order stitching during the reconstruction of the measured data. More measurement images are needed according to the number of higher orders to be included.

Saturation leads to an effective excitation PSF broadening which can initially be considered a loss of resolution and thus to be avoided if possible [99]. In CLSM-based Saturated Excitation Microscopy (SAX) however, it is the center that exhibits an emission depending nonlinear on the excitation that higher resolution is realized. In SAX the exciting laser light is modulated by an interferometric combination of two differently Doppler-shifted beams (using Acousto Optical Modulator (AOM)s) to receive a sinusoidal temporal intensity modulation e. g. at $f^{(\mathrm{ex})} = 10\,kHz$. The intensity is chosen such that it is sufficiently high to saturate the sample, but not yet sufficient for multi-photon excitation. The fluorescence-saturation causes higher temporal harmonics $f^{(\mathrm{NL,i})}$ for $i \in 1, 2, \ldots$ of the modulation frequency $f^{(\mathrm{ex})}$ in the emitted signal which can be separated from the linearly responding fundamental signal using a lock-in amplifier driven at this harmonic frequency. This achieves higher resolution [100]. Since we aim to follow the ideas of SAX, we analyze our recorded data in the framework and nomenclature of SAX and extract *non-linear orders* (NL) even though the data is acquired by recording images at three fixed

illumination-intensity values rather than applying a continuous intensity modulation. As shown below, our approach is better suited to low-noise scenarios while shrinking the setup-complexity or in general improving the versatility of the method.

In dynamic saturation optical microscopy (DSOM), the fluorescence signal temporarily sampled with high resolution and thus the sample fluorophore distribution is reconstructed from the temporal course of the fluorescence response [101]. Using pulsed excitation with a set of *ps*-pulses and repetition-times in the 100 *ms*-range the influence of the triplet state and hence the tendency towards bleaching can be reduced while sticking to the SAX imaging scheme [102]. Infinite resolution is limited by photostability and brightness. Regaining e. g. the second and third nonlinear fluorescence order[1], can also be achieved via the imaging scheme of differential Saturated Excitation (dSAX) [103]. With this method instead of excitation intensity modulation by using two AOMs and a lock-in for post-detection extraction of higher harmonics, multiple images at different excitation intensities were taken. Now, only one AOM together with a function generator for excitation intensity switching is used and no lock-in is required anymore. The new approach to increase SNR of dSAX by implementing a ISM imaging scheme has been part of a joint publication [104]. In this thesis, the fundamentals of the approach and further data processing will be presented.

E.2 THEORY AND METHODS OF DSAX-ISM

E.2.1 *Two-State Saturated Fluorescence Model*

For the fluorophores excited in the linear range it has been analyzed in the previous sections how much information can be obtained from the measured data. The reconstructable object data are limited despite the use of assumptions (priors) about the original distribution of the OTF support. By manipulating the imaging process or the imaging geometry, this support can be changed. For the following sections, a nonlinear fluorophore response is generated by saturated excitation. The model of the excitation behavior of a fluorophore introduced in Section 1.5 is further simplified for the purposes of this section and thus reduced to a two-state model. For modeling purposes, this corresponds to Figure 1.2

1 Wording and usage established in works concerning with SAX.

without $S^{(2)}$, $T^{(0)}$ as well as the corresponding transitions. Thus, the ground state $S^{(0)}$, the excited state $S^{(1)}$, the absorption rate $k^{(a)}$ ($S^{(0)} \longrightarrow S^{(1)}$), and the fluorescence rate $k^{(f)}$ ($S^{(1)} \longrightarrow S^{(0)}$) remain. With $N^{(0/1)}$ being the number of molecules in the $S^{(0/1)}$ state from the rate equation solutions Equation E.26 (see Section E.5) the fluorescent photonflux $\Phi^{(F)}(t)$ (Equation 1.29) can be rephrased as:

$$\begin{aligned} \Phi^{(F)}(t) &= Q^{(e)} k^{(f)} N^{(1)}(t) \\ \overset{t \to t^{(e)}}{\to} \Phi^{(F)} &= Q^{(e)} k^{(f)} N^{(1,e)} = \frac{Q^{(e)} k^{(f)} k^{(a)}}{k^{(f)} + k^{(a)}} = Q^{(e)} \frac{\tau^{-1} I^{(ex)}}{I^{(ex)} + \left(\tau \eta^{(a)}\right)^{-1}} \equiv Q^{(e)} \Phi^{(F')} \end{aligned} \tag{E.1}$$

with $k^{(a)}$ the intensity dependent absorption rate (Equation 1.28), $\eta^{(a)}$ the fluorophore optical absorption cross-section, $k^{(f)} = \tau^{-1}$ the emission rate constant, τ the excited state lifetime, $Q^{(e)}$ the fluorophore quantum efficiency and $I^{(ex)}$ the excitation intensity. Note that $I^{(ex)}$ can be expressed in units of $photons/cm^2/s$ by making use of the Planck-Einstein relation $E^{(ex)} = h\lambda^{ex}/c$ with $[E] = J/photon$, thereby converting $J \to photons$ and thus both summands in the denominator of Equation E.1 have the same physical units for each individual photon and molecule. The spatial dependence (e. g. $\Phi^{(F')}(\vec{x})$) is not visualized in the equations above and only one type of fluorophore (i. e. only one τ) with the same absorption cross section over the whole object was assumed for the model. The degree of saturation of a fluorophore $\psi^{(max)}$ at a given (maximum) intensity $\max(I^{(ex)})$ can be defined by Equation E.1 as:

$$\psi^{(max)} = \tau \cdot \Phi^{(F')}(\max(I^{(ex)})) \tag{E.2}$$

According to Equation 1.30 by multiplying with the spatial sample-density distribution $C^{(F)}(\vec{x})$ the final fluorescent emittance $I^{(F)}$ can be calculated:

$$I^{(F)} = \Phi^{(F')}(\vec{x}) Q^{(e)} C^{(F)}(\vec{x}) = \left[\Phi^{(F')} S'\right](\vec{x}) \tag{E.3}$$

with $S' \equiv Q^{(e)} C^{(F)}(\vec{x})$. For simulations in this chapter the nonlinearity of the sample response is modelled into a modified illumination $\Phi^{(F')}(\vec{x})$ according to Equation E.1.

A brief example for the paramaters in case of Fluorescein will be given based on [105] . Here $\eta^{(a)} = 3.06 \cdot 10^{-16}\, cm^2/molecule$, $\tau = 4.5 \cdot 10^{-9}\, s$ yielding

$I^{(ex,S)} \approx 7.3 \cdot 10^{23}\, photon/cm^2/s$ where the assumption of a perfect conversion efficiency $Q^{(e)} = 1\, photon/molecule$ was used. In case of typical CLSM a Gaussian beam, i. e. a beam with shape like a Gaussian distribution in its lateral profile, with e. g. waist radius

$w_0 = 250\,nm$ focused by an objective lens with $NA = 1.25$ and oil-immersion has a focus-spot intensity of $I^{(\mathrm{ex})} = 5.1 \cdot 10^5\, W/cm^2$. The example of Fluorescein was chosen as most of the dyes used in this thesis are of type Alexa which was created from Fluorescein [30].

E.2.2 *Reconstruction Scheme of dSAX-ISM*

For intensities $I^{(\mathrm{ex})} \ll (\tau\eta^{(\mathrm{a})})^{-1}$ the transition $\Phi^{(\mathrm{F}')} \longrightarrow I^{(\mathrm{ex})}$ holds and the linear excitation process studied so far in this chapter can be reconstructed. In the following, cases for the range $I^{(\mathrm{ex})} = O\left(\left[\tau\eta^{(\mathrm{a})}\right]^{-1}\right)$ are considered. One speaks of a saturated state when $\geq 50\%$ of the molecules are (permanently) in the excited state $S^{(1)}$, i.e. $N^{(1,\mathrm{e})} = 0.5N$. This will be the case at the saturating excitation intensity $I^{(\mathrm{ex,S})}$:

$$I^{(\mathrm{ex,S})} = \left(\tau\eta^{(\mathrm{a})}\right)^{-1} \tag{E.4}$$

and with Equation E.1 it follows $\Phi^{(\mathrm{F}')}\left(I^{(\mathrm{ex,S})}\right) = (2\tau)^{-1}$. A Taylor series expansion of $\Phi^{(\mathrm{F}')}$ around $I^{(\mathrm{ex})} = I^{(\mathrm{ex,S})}$ yields:

$$\Phi^{(\mathrm{F}')}\left(I^{(\mathrm{ex})}\right) = \sum_n^{\infty} \frac{1}{n!} \partial^{(n)}_{I^{(\mathrm{ex})}} \left[\Phi^{(\mathrm{F}')}\right]_{I^{(\mathrm{ex})}=I^{(\mathrm{ex,S})}} \left(I^{(\mathrm{ex})} - I_s^{(\mathrm{ex})}\right)^{(n)}$$

$$\Phi^{(\mathrm{F}')}\left(I^{(\mathrm{ex})} + I^{(\mathrm{ex,S})}\right) = \Phi^{(\mathrm{F}')}\left(I^{(\mathrm{ex,S})}\right) + \sum_{n=1}^{\infty} \frac{(-\Gamma)^{n+1}}{(n-1)!} \left(I^{(\mathrm{ex})}\right)^n \tag{E.5}$$

$$\Gamma^n = \frac{\tau^{n-1}}{\left[1 + I^{(\mathrm{ex,S})}\tau\right]^n} \tag{E.6}$$

Equation E.5 can be further simplified by normalization and another coordinate transformation to:

$$\hat{I}^{(\mathrm{F}')}\left(\hat{I}^{(\mathrm{ex})}\right) \equiv \Phi^{(\mathrm{F}')}\left(\hat{I}^{(\mathrm{ex})}\right) - \Phi^{(\mathrm{F}')}\left(I^{(\mathrm{ex,S})}\right) \approx n\Gamma \sum_{n=1}^{\infty} (-1)^{n+1} \left(\hat{I}^{(\mathrm{ex})}\right)^n \tag{E.7}$$

$$\hat{I}^{(\mathrm{ex})} \equiv \frac{\Gamma}{\sqrt[n]{2\pi n}} \left(\frac{e}{n}\right) \left(I^{(\mathrm{ex})} - I^{(\mathrm{ex,S})}\right)$$

where Stirling's approximation $n! \approx \sqrt{2\pi n}(n/e)^n$ was used for simplifying the factorial. According to Equation A.58, for a recorded image using a 2D detector at position $z = z^{(\mathrm{o})}$:

$$M \propto I^{(\mathrm{M})}(\vec{\chi}) = \left[I^{(\mathrm{F})} \otimes h_{det}\right]_{z=z^{(\mathrm{o})}} (\vec{\chi}) \tag{E.8}$$

Due to linearity of the convolution operation (with h_{det}) and multiplication (with S') further calculations only focus on the $I^{(\mathrm{ex})}$ terms. The goal of the next steps is to isolate higher order fluorescent responses and thereby increase spatial image-resolution, i. e. finding a scheme to individually isolate the $\left(I^{(\mathrm{ex})}\right)^n$ terms.

According to Equation E.5, the Taylor series expansion for a chosen excitation intensity $I^{(\mathrm{ex})}$ vanishes rapidly $\Gamma^{k+1}\left(I^{(\mathrm{ex})}\right)^k/(k-1)! \longrightarrow 0$ for increasing higher orders k and its residual estimation $R \propto O\left(\left[I^{(\mathrm{ex})}\right]^k\right)$ vanishes for any chosen k within this thesis. To e. g. reconstruct the saturated fluorescent responses up to $k = 3$, taking three images $M^{(\mathrm{k})}$ at different excitation intensities $I^{(\mathrm{ex\ k})}$ ($k \in [1, 2, 3]$) equates to:

$$M^{(1)} \propto \hat{I}^{(\mathrm{ex1})} \tag{E.9}$$

$$M^{(2)} \propto \hat{I}^{(\mathrm{ex2})} - \left(\hat{I}^{(\mathrm{ex2})}\right)^2 \tag{E.10}$$

$$M^{(3)} \propto \hat{I}^{(\mathrm{ex3})} - \left(\hat{I}^{(\mathrm{ex3})}\right)^2 + \left(\hat{I}^{(\mathrm{ex3})}\right)^3 \tag{E.11}$$

By solving the system of equations, the 1^{st} and 2^{nd} nonlinear fluorescence orders e. g. can be obtained in terms of $M^{(\mathrm{NL1})}$ and $M^{(\mathrm{NL2})}$ [103]:

$$M^{(\mathrm{NL1})} = \frac{\hat{I}^{(\mathrm{ex2})}}{\hat{I}^{(\mathrm{ex1})}} M^{(1)} - M^{(2)} \tag{E.12}$$

$$M^{(\mathrm{NL2})} = \left(\frac{\hat{I}^{(\mathrm{ex3})}}{\hat{I}^{(\mathrm{ex2})}}\right)^2 M^{(\mathrm{NL1})} - \frac{\hat{I}^{(\mathrm{ex3})}}{\hat{I}^{(\mathrm{ex1})}} M^{(1)} + M^{(3)} \tag{E.13}$$

Since the nonlinear orders can be extracted by linear combination of the measured data, the method is named dSAX. Assuming a Gaussian-like illumination distribution $I^{(\mathrm{ex})}$, it follows from Equation E.7 (or even more clearly Equation E.12) that the gain in resolution over the unsaturated image scales with $\sqrt{n}$. Note, that while the saturated images could be directly used within e. g. a nonlinear iterative reconstruction like mDEC or the saturation model could even directly be baked into a semi-blind deconvolution model, the existing SAX and dSAX approaches try to extract these higher nonlinear orders $M^{(\mathrm{NL1})}$, $M^{(\mathrm{NL2})}$, ... in a linear manner. Thus, for comparability of existing works the theoretical and experimental analysis within this thesis will focus around the reconstructibility of these higher nonlinear orders.

Using the example of the two-state model derived in Section E.2.1, the influence of saturation on the individual PSFs and OTFs is shown in Figure E.1. Optical parameters chosen were $NA = 1.3$, $\lambda^{(\mathrm{ex})} = 488\,nm$, $\lambda^{(\mathrm{em})} = 520\,nm$. For the spatial image sampling a strong undersampling of $d^{(\mathrm{SPL,lat})} = 10\,nm$ and $d^{(\mathrm{SPL,ax})} = 50\,nm$ was chosen for better visualization. This gives $d^{(\mathrm{Abbe,lat})} = 198\,nm$ for the detection PSF $h^{(\mathrm{det})}$ according to Equation 1.12 and thus by Table 1.1 $d^{(\mathrm{1AU,lat})} \approx 483\,nm$ and $d^{(\mathrm{1AU,ax})} \approx 1715\,nm$. Accordingly, the Abbe frequency in Fourier space according to Equation 1.13 is $k^{(\mathrm{Abbe,lat})} = 31.7\,rad/\mu m$ and $k^{(\mathrm{Abbe,ax})} = 8.9\,rad/\mu m$. Saturation parameters were calculated according to the Fluorescein example to $\tau = 4.5{\cdot}10^{-9}\,s$, $\eta^{(\mathrm{a})} = 3.06{\cdot}10^{-16}\,cm^2/molecule$, $Q^{(\mathrm{e})}C^{(\mathrm{F})} = 3.2^{15} molecules/cm^2$

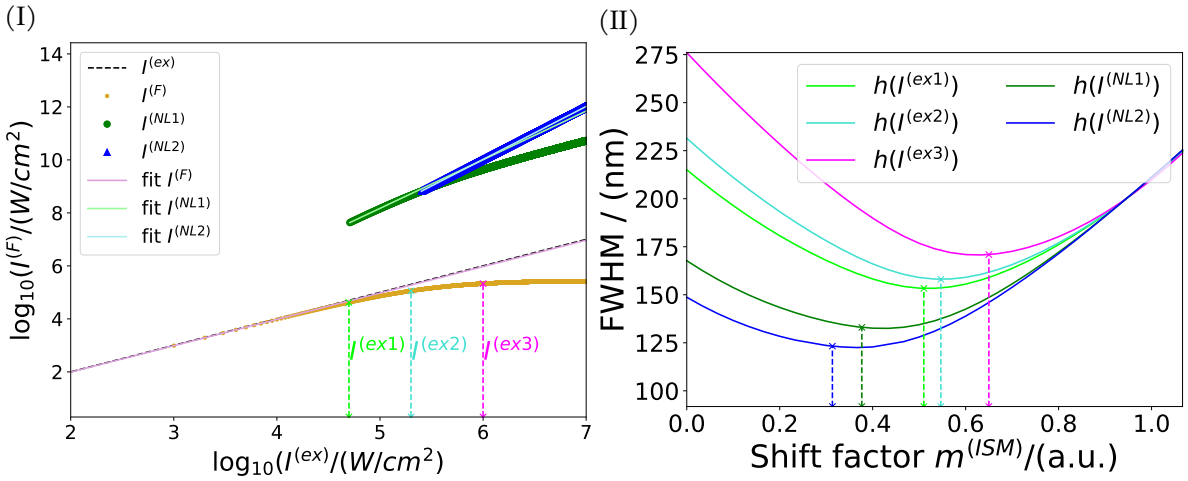


FIGURE E.1: **DSAX Saturation Curve.** (I) Simulated saturation curve (golden) and applied excitation intensity (black dotted) as log-log-plot. Chosen saturation intensities ($I^{(\mathrm{ex1\ldots3})}$) are displayed for reference. Calculated nonlinear first $I^{(\mathrm{NL1})}$ (green) and second (blue) order $I^{(\mathrm{NL2})}$ and linear fits to the measured and calculated curves. (II) Dependency of FWHM of measured images and extracted higher orders on the use lateral reassignment shift-distance. Correlation based calculated shift-values and achieved FWHM are displayed with a dotted line and a cross at its ends.

(per unit-depth dz) chosen, giving $\psi^{(\max)} = 1/2$ at $I^{(\mathrm{ex})} \approx 2800\,kW/cm^2$. Figure E.1I demonstrates the fluorescent signal $I^{(\mathrm{F})}$ (golden) over an excitation intensity range of $I^{(\mathrm{ex})} \in [1, \ldots, 10^7]$ with step size $\Delta I^{(\mathrm{ex})} = 1000$. For the reconstruction of the first and second nonlinear orders, three different excitation intensities $I^{(\mathrm{ex})}$ must be chosen according to Equation E.12 and Equation E.13. Analogous to the procedure of Nawa *et al.* [103], a linear fit (fit $I^{(\mathrm{F})}$, see Figure E.1I) is applied to the saturated fluorescence curve $\log_{10}(I^{(\mathrm{F})})$, in log-log representation. From the linear range of $\log_{10}(I^{(\mathrm{F})})$, the largest possible $I^{(\mathrm{ex1})}$ is now chosen to achieve the lowest possible noise measurement signal. From the chosen $I^{(\mathrm{ex1})}$ and $I^{(\mathrm{F1})} = I^{(\mathrm{F})}(I^{(\mathrm{ex1})})$, the fluorescence curve $I^{(\mathrm{F})}(I^{(\mathrm{ex})} > I^{(\mathrm{ex1})})$ is used, according to Equation E.12, to calculate the curve $I^{(\mathrm{NL1})}$. Another linear fit to $\log_{10}(I^{(\mathrm{NL1})})$ (green, see Figure E.1I) allows us to estimate the linear range of $\log_{10}(I^{(\mathrm{NL1})})$ and thus choose the largest possible $I^{(\mathrm{ex2})}$ with associated $I^{(\mathrm{FNL1})}$. With the tuple $\{I^{(\mathrm{ex1})}, I^{(\mathrm{F1})}\}$, $\{I^{(\mathrm{ex2})}, I^{(\mathrm{FNL1})}\}$ chosen in this way, it is now possible to extract $I^{(\mathrm{F})}(I^{(\mathrm{ex})} > I^{(\mathrm{ex2})})$ from the remaining

	$h^{(1)}$	$h^{(2)}$	$h^{(3)}$	$h^{(NL1)}$	$h^{(NL2)}$
FWHM/AU	0.77	0.81	0.91	0.67	0.62
$h^{(i)}/h^{(1)}$	1.00	1.05	1.17	0.87	0.81

TABLE E.1: **FWHM of dSAX-ISM PSF Simulation.** Results are based on the chosen excitation intensities of Figure E.1I.

fluorescence curve. According to Equation E.13 $I^{(\mathrm{NL2})}$ can now be determined. A new linear fit to the calculated curve $\log_{10}(I^{(\mathrm{NL2})})$ allows the choice of $I^{(\mathrm{ex3})}$. In the simulation, $I^{(\mathrm{ex1})} = 50\,kW/cm^2$, $I^{(\mathrm{ex2})} = 200\,kW/cm^2$, and $I^{(\mathrm{ex})} = 1000\,kW/cm^2$ were chosen.

The lateral and axial PSFs broaden with increasing saturation by up to 17% in the case of $h^{(3)}/h^{(1)}$ simulation (see Table E.1). $h^{(\mathrm{NL1})}$ and $h^{(\mathrm{NL2})}$ were calculated according to Equation E.12 and Equation E.13. The isolated nonlinear orders on the other hand achieve a resolution gain of up to 19% in the case of $h^{(\mathrm{NL2})}$.

The dependence of Full Width at Half Maximum (FWHM) on the selected shift vector along the X-axis is shown in Figure E.1II. The simulation parameters remain unchanged, with only the sampling adjusted to $d^{(\mathrm{SPL,lat})} = 40$ nm, $d^{(\mathrm{SPL,ax})} = 100$ nm and extenting the detector to consist of $[16, 16]$ pixels. The dashed lines mark location and FWHM after determination of the shift value using method 2. The deviation of the of the FWHM curves' individual minimum from the X-marked shifts determined using method 2, although the same FWHM is obtained, might indicate numerical inaccuracies of the correlation or shift routine for subpixel shifts. Since no appreciable influence of this inaccuracy could be detected in other simulation and processing scenarios, no deeper analysis of this circumstance was performed. The values given in Figure E.1II also illustrate that the optimal lateral shift values depend on the degree of saturation and offsetting. Thus, e. g. the optimal shift values of $h^{(\mathrm{Iex3})}$ and $h^{(\mathrm{NL2})}$ differ by $\Delta x^{(\mathrm{Iex3,opt})}/\Delta x^{(\mathrm{NL2,opt})} \approx 73\%$. If e. g. the shift value $\Delta x^{(\mathrm{Iex3,opt})}$ would be used for $h^{(\mathrm{NL2})}$, the FWHM would be $143\,nm$ instead of $122\,nm$, resulting in a deviation of $\approx 17\%$. The order of reconstruction, 1) pixel reassignment followed by 2) extraction of the nonlinear orders or 2) and then 1), has no significant effect on the FWHM in this case, since the calculation (Equation E.13) is linear.

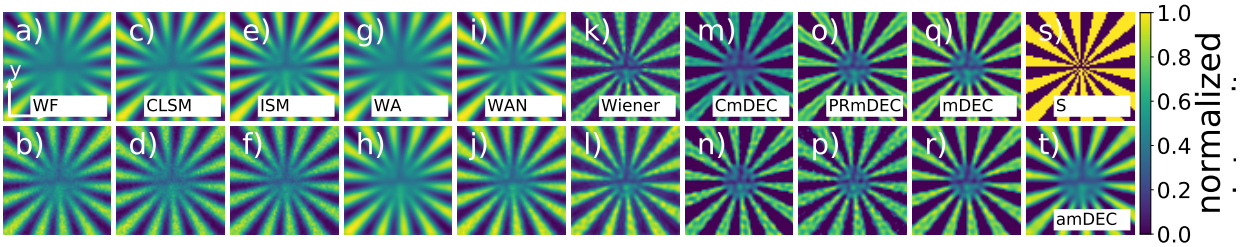


FIGURE E.2: **DSAX Comparison of Reconstruction Methods for High SNR.** Reconstruction results of differential Saturated Excitation Image Scanning Microscopy (dSAX-ISM) images. Upper row: applied to $M^{(1)}$, lower row: applied to $M^{(\mathrm{NL2})}$. a+b) WF, c+d) CLSM at 1 *AU* pinhole size, e+f) PiRe at 1 *AU* pinhole size, g+h) WAVG i+j) noise normalized Weighted Averaging in Fourier Space (nWAVG), k+l) generalized WD, m+n) deconvolution of CLSM data (CmDEC), o+p) deconvolution of ISM data (PmDEC), q+r) multi-view deconvolution , s) ground truth (=object), t) multi-view deconvolution using all measured images $M^{(\mathrm{i})}$ for $i \in 1,2,3$ (amDEC). Images are normalized to their individual zero frequency.

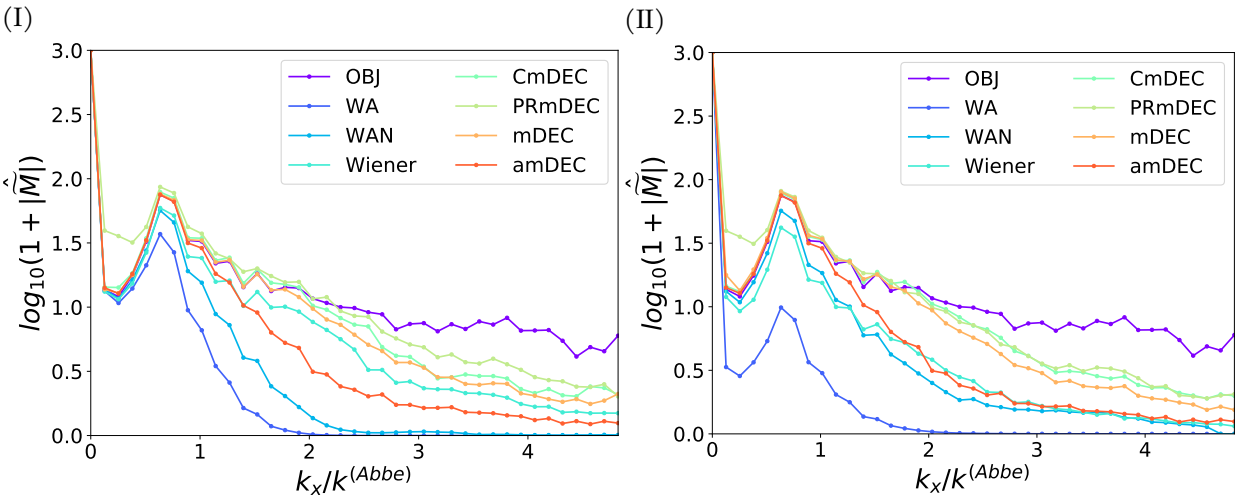


FIGURE E.3: **Reconstruction of DSAX Data for High SNR.** Fourier radial sum example for (I) $\widetilde{M}^{(\mathrm{o})}$, (II) $\widetilde{M}^{(\mathrm{NL2})}$.

E.2.3 *Comparison of Reconstruction Methods on dSAX-ISM*

Further processing by means of WAVG (Section 1.7) and a set of deconvolution techniques will be applied to the dSAX-ISM data in case of high- and low-SNR. For legend entries the following nomenclature is used: OBJ=object, Wiener=WD, WA = WAVG, WAN = nWAVG, mDEC = mDEC and amDEC = multi-view DEConvolution using all available

	$M^{(1)}$	$M^{(2)}$	$M^{(3)}$	$M^{(NL1)}$	$M^{(NL2)}$
WF	0.78	0.76	0.71	0.82	0.83
CLSM	0.81	0.80	0.77	0.84	0.85
ISM	0.83	0.83	0.81	0.85	0.86
WA	0.78	0.77	0.76	0.81	0.82
WAN	0.84	0.83	0.83	0.86	0.87
Wiener	0.92	0.93	0.94	0.90	0.89
CmDEC	0.90	0.88	0.85	0.92	0.91
PRmDEC	0.93	0.92	0.90	0.92	0.91
mDEC	0.93	0.92	0.91	0.88	0.91
amDEC	0.87	0.87	0.87	0.87	0.87

(a)

	$M^{(1)}$	$M^{(2)}$	$M^{(3)}$	$M^{(NL1)}$	$M^{(NL2)}$
WF	0.77	0.75	0.70	0.68	0.52
CLSM	0.80	0.79	0.77	0.67	0.50
ISM	0.82	0.83	0.81	0.69	0.52
WA	0.78	0.77	0.76	0.80	0.78
WAN	0.84	0.83	0.83	0.82	0.79
Wiener	0.88	0.89	0.90	0.83	0.79
CmDEC	0.87	0.87	0.85	0.81	0.77
PRmDEC	0.89	0.91	0.90	0.82	0.77
mDEC	0.90	0.91	0.90	0.82	0.77
amDEC	0.93	0.93	0.93	0.93	0.93

(b)

TABLE E.2: **NCC Values of DSAX Processing.** For processing of dataset (a) high-SNR (b) low-SNR. Compared techniques abbrevations are: OBJ=object, Wiener=WD, WA = WAVG, WAN = nWAVG, mDEC = mDEC and amDEC = mDECusing all available (saturated) images $M^{(i)}$ for $i \in 1, 2, 3$.

images (amDEC) using all available (saturated) images $M^{(i)}$ for $i \in 1, 2, 3$. Thus in case of amDEC only 1 reconstruction for the whole set of images exists. Simulation parameters are the same as in the previous section.

Figure E.2 shows a comparison of the different reconstruction methods applied to $M^{(1)}$ (top row) and $M^{(NL2)}$ (bottom row) in the case of high-SNR. The last column is isolated and shows the underlying object and the amDEC reconstruction.

From the visual impression, almost all post-processed and non-post-processed results based on the nonlinear third order $M^{(NL2)}$ (Figure E.2, lower row) isolated by means of dSAX-ISM appear sharper, since the struts of the *spokes2d* target can be traced further inwards, they are filled more evenly and their edge gradient appears steeper. Only in the case of WD (g) a better result seems to be achieved by means of the dataset $M^{(1)}$ (Figure E.2, upper row) used for comparison.

Calculation of NCC by comparison to the simulation object (o) confirms this observation, since all values of the last column (see Table E.2a) are larger than those of the first column, except in the case of the WD. Since amDEC uses all underlying measurement data for

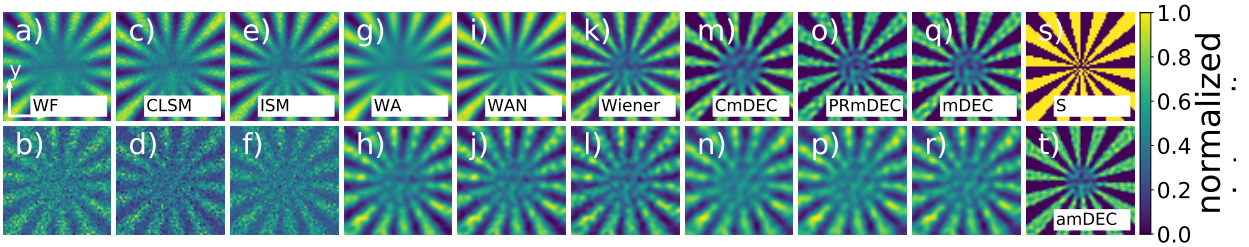


FIGURE E.4: **DSAX Comparison of Reconstruction Methods for Low SNR.** Reconstruction results of dSAX-ISM images for low-SNR. Upper row: applied to $M^{(1)}$, lower row: applied to $M^{(\mathrm{NL2})}$. Numbering and normalization analog to Figure E.2.

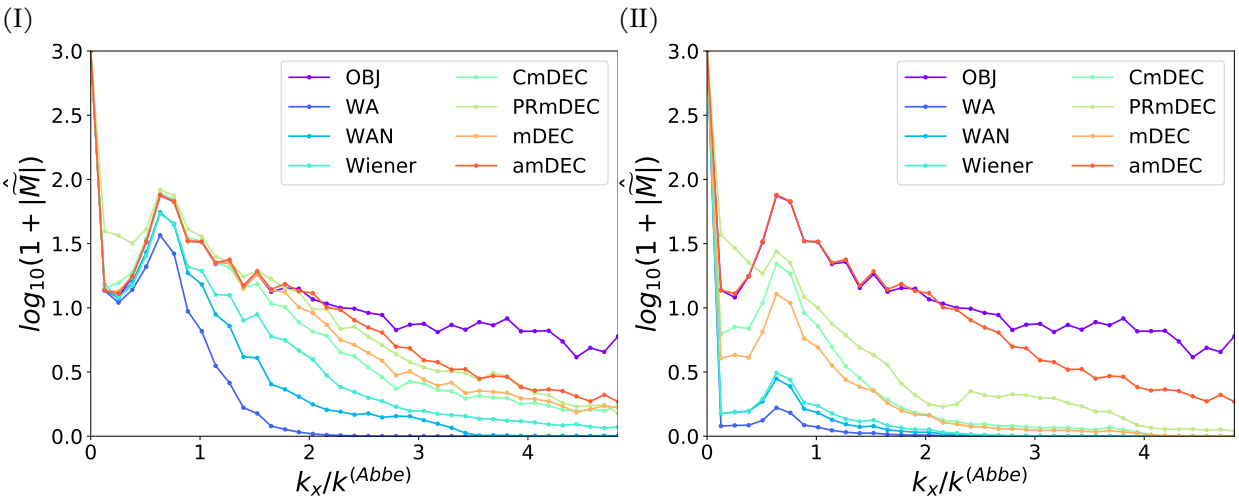


FIGURE E.5: **Reconstruction of DSAX Data for Low SNR.** Fourier radial sum example for (I) $\widetilde{M}^{(\mathrm{o})}$, (II) $\widetilde{M}^{(\mathrm{NL2})}$.

multi-view deconvolution only one correlation value exists, but it was entered in all columns. ISM always achieves NCC greater similarities to the object than CLSM at 1 AU or WF. WD is the most NCC performant method in this reconstruction comparison.

A comparison of the amplitudes of the frequency components of the reconstructions of $M^{(1)}$ (see Figure E.3I; Equation E.9) and $M^{(\mathrm{NL2})}$ (see Figure E.3II) does not suggest the previous observation that WD reconstructs the data better than e. g. mDEC. The data are normalized to $1000/k_0$, with k_0 their individual zero frequency. The plots suggest that mDEC and amDEC are nearly identical to the original object spectrum until just below $k_x = 2k^{(\mathrm{Abbe})}$, and below that mDEC is closest to the original frequency distribution. For $k_x \geq 2k^{(\mathrm{Abbe})}$, the amplitudes of the frequency spectra of all reconstructions drop below the original object distribution. While WD, mDEC and amDEC are an estimate of the noise-free

object, WAVG and nWAVG are still noise PSF-weighted summations of the measured data. They are included in this comparison image to illustrate the noise-cutoff and frequency-dependent decay of the OTF of the underlying imaging. For $k_x \geq 2.5 \cdot k^{(\mathrm{Abbe})}$, WAVG and nWAVG are negligible and thus above this limit the comparison of the frequencies reconstructed by e. g. mDEC with the original frequency spectrum of OBJ are of particular interest. In the case of $M^{(\mathrm{NL2})}$, nWAVG and WD are very similar in shape, while mDEC and amDEC are much closer to the original frequency spectrum, see Figure E.3II. This difference manifests itself only slightly in the calculated NCC values (Table E.2a), but in the pictorial comparison (Figure E.2) by larger unevenness in the struts (areas of equal pixel values) of the *spokes2d* target at WD compared to mDEC.

In the low-SNR case, a visual SNR gain of ISM (f) over CLSM (d) is noticeable, see Figure E.4. However, the quality of the reconstructions based on the nonlinear 2^{nd} order $M^{(\mathrm{NL2})}$ is lower than the reconstructions of the non-saturated fluorescence data (top row).

This conjecture is supported by the determined NCC values, see Table E.2b. All reconstructions achieve better results for the direct simulation data $M^{(\mathrm{i})}$ compared to the extracted nonlinear orders $M^{(\mathrm{NLi})}$. ISM is also NCC-closer to the original object than CLSM. amDEC and mDEC now achieve the NCC-best results here and are NCC-closer to the original object than WD. In the case of $M^{(\mathrm{NL2})}$ processing, WD, WAVG, nWAVG and mDEC achieve the same NCC value of 0.78. This suggests that the frequency cut $k^{(\mathrm{C,NL2})}$ of the underlying simulation data is lower than e. g. at $M^{(\mathrm{o})}$ and thus less frequencies can be reconstructed from the noise floor. Figure E.5II supports this conjecture, as the normalized amplitude values for all reconstructions (amDEC is displayed here for reference only) run well below the distribution of the original frequency spectrum.

For a SNR comparison of CLSM-based dSAX and dSAX-ISM the reader is referred to Section E.6.

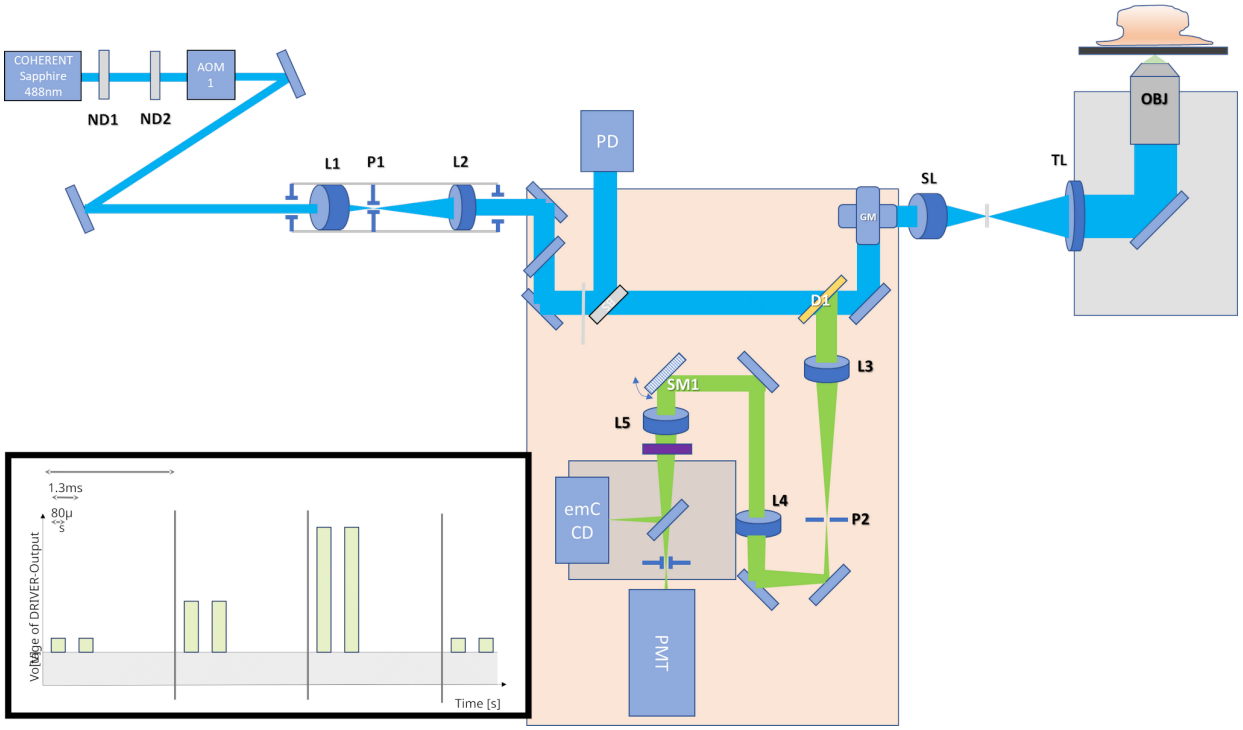


FIGURE E.6: **DSAX Experimental Setup.** Insert displays pulse-timing used.

E.3 DSAX-ISM EXPERIMENTS

E.3.1 *dSAX-ISM Experimental Setup*

The experimental setup is basically a CLSM setup extended by a camera (emCCD) and AOM, see Figure E.6. A continuous wave laser (Sapphire-488 nm,Coherent) is used as the light source. For precise intensity adjustment, two ND filters (ND1,ND2) follow after the laser. For pulse length and intensity control, a AOM (AOM-405-AF1-J, IntraAction Corp. with driver ME-405-J, IntraAction Corp.) is used. The beam is then focused onto a pinhole ($d = 50\,\mu$m) using an achromatic lens $f^{(\mathrm{L1})} = 50$ mm (AC254-050,Thorlabs) where it gets spatially cleaned and afterwards magnified by the telescope assembly using a second lens $f^{(\mathrm{L2})} = 150$ mm 3× to fill the BFP of the objective lens. A beam splitter BS (BSN10R, Thorlabs) reflects 10% of the excitation light onto a Photodiode (PD) (S3399, Hamamatsu with amplifier C8366, Hamamatsu) to always measure the excitation intensity present. A pair of galvanometric mirrors (6210HBR, Cambridge Technology) is centered around the BFP of the scanning lens $f^{(\mathrm{SL})} = 50$ mm, allowing the angle of the beam in the BFP and thus the lateral position of the excitation spot in the object plane to be changed. Due to the larger focal length of the subsequent tube lens $f^{(\mathrm{TL})} = 180$ mm, the beam diameter is expanded

3.6× and focused into the sample plane by means of an objective lens (UPlanSApo 60×, 1.30 Sil, Olympus Corporation). The fluorescent emission is collected through the same objective, propagated back into the BFP of the SL using TL and SL, and descanned there. A dichromatic short-pass beam splitter D1 (FF505-SDi01-25×36, Semrock) separates the fluorescent emission from the excitation beam path. Another telescope, $f^{(L3)} = 200$ mm and $f^{(L4)} = 200$ mm serves as a relay and for coarse ooF filtering through the pinhole ($d = 50\,\mu$m) located between the lenses. A further achromatic lens $f^{(L5)} = 200$ mm ultimately focuses the beam onto a PMT (H7422-40, Hamamatsu), in the case of CLSM use, or emCCD camera (proEM 1024B, Princeton Imaging) in the case of ISM. The PMT is used to search for a suitable sample position. One pixel cycle, i.e. dwell, readout and dead-time, of the camera requires $t^{(Ccycle)}5$ ms. In sample coordinates, one lateral AU of the detection beam has the size of $d^{(1AU,sample)} = 483$ nm. With a system total magnification of $M^{(total)} = 144$ and the camera pixel pitch of $d^{(Pcam)} = 13\,\mu$m, this corresponds to $d^{(1AU,sample)} \approx 5.35\,pixels$. The camera pixel pitch expressed in AU is thus approximately $d^{(Pcam)} \approx 0.2\,AU$. To record the first three AU of the fluorescent emission, a central area of the camera of 16 × 16 pixels is used.

Constant excitation in the nonlinear response regime of the fluorophores, in the case of the presenting setup geometry, can easily lead to undesired side effects such as thermal heat aggregation or bleaching [102, 106]. Since the degree of fluorophore saturation depends on peak photon flux and less on flux duration (Equation E.1), a pulse-like excitation is generated using a function generator (AFG3102C, Tektronics) with the AOM. The confocal scanning unit is parked on a spatial position for the subsequent acquisition of the three images (Equation E.13) necessary to reconstruct the third order in the dSAX-ISM scheme, where one image is recorded at each of three different peak intensities (see insert lower left of Figure E.6). During one recording cycle $t^{(Ccycle)}$ two rectangular pulses with $t^{(Pw)} = 80\,\mu s$ duration and $t^{(Pp)} = 1.3\,\mu s$ period are used.

E.3.2 *Measuring the In-Sample Saturation Curve*

In order to determine the appropriate excitation intensities for the particular sample, in case of the dSAX-ISM scheme, the fluorescent response of the sample must first be studied as a function of excitation intensity $I^{(ex)}$. Here, the voltage applied to the PD (PD, see

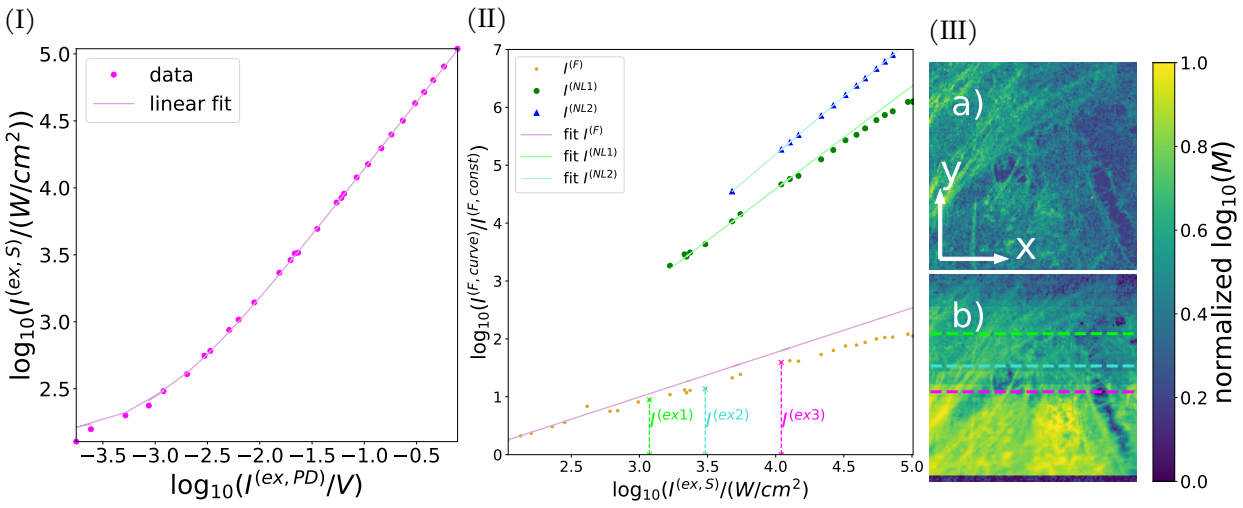


FIGURE E.7: **DSAX Experimental Saturation Curve.** (I) Linear Fit to PD measurements. Displayed as log-log-curve. (II) Saturation curve for Alexa-488 stained HeLa-cell. Measured fluorescent response (magenta points), calculated first order (NL1, green thick points) and second order (blue triangles) are displayed together with their respective linear fits. (III) Logarithm of Measured images with a) constant excitation intensity b) steps of increasing intensity. Chosen excitation intensities, $I^{(\mathrm{ex1})}$ (lime), $I^{(\mathrm{ex2})}$ (turquoise) and $I^{(\mathrm{ex3})}$ (magenta) for further DSAX-measurements are marked with a ticked line.

Figure E.6) must be converted to the power present in the sample plane. To measure the power in the sample plane, a longitudinal measurement is performed using a powermeter (PM100D, Thorlabs). The measured power is converted to the applied intensity via the theoretical focal spot size of $d^{(\mathrm{1AU,exc})}$ (Table 1.1) according to Equation E.14:

$$I^{(\mathrm{ex,S})} = \frac{P^{(\mathrm{ex,S})}}{\pi \left(d^{(\mathrm{1AU,exc})}/2\right)^2}, \quad \left[I^{(\mathrm{ex})}\right] = \frac{W}{cm^2} \tag{E.14}$$

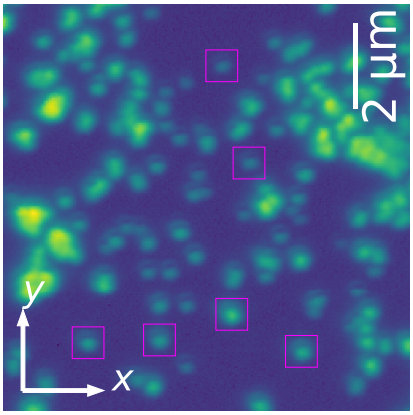
A linear fit to the intensity measurement series yields the relation Equation E.15:

$$I^{(\mathrm{ex,S})} = \frac{1.38 \cdot 10^5}{V} \cdot I^{(\mathrm{ex,PD})} + 1.39 \cdot 10^2 \frac{W}{cm^2} \tag{E.15}$$

The fit including the measurement series is shown as a log-log plot for better visibility of the used excitation intensity steps $\Delta I^{(\mathrm{ex})}$, see Figure E.7I.

Using Alexa488 phalloidin conjugate (A12379, ThermoFischer) stained Actin filaments of Henrietta Lacks Cervical Cancer (HeLa) cells in an anti-photobleaching agent (ProlongGlass Antifade Mountant, ThermoFischer) as an example, the procedure will be

(I)



(II)

	$M^{(1)}$	$M^{(2)}$	$M^{(3)}$	$M^{(NL1)}$	$M^{(NL2)}$
CP	205	208	216	199	203
CLSM	222	228	241	216	237
ISM	219	224	234	206	201

FIGURE E.8: **DSAX FWHM of PSF Measurements.** (I) Acquired image. Selected beads are marked with a magenta box. (II) Calculated FWHM from fitted PSF measurements. All calculated sizes in *AU*.

demonstrated. A second, similar PD after the BS in the detection beam-path is used to measure the fluorescence response. A sample area filled as uniformly as possible with stained object structure is imaged with a constant, low intensity $I^{(\mathrm{F,const})}$ (see Figure E.7III a) as well as with an excitation intensity $I^{(\mathrm{F,curve})}$ (see Figure E.7III b) that increases every four lines. The mean is calculated over four lines each for the ratio $I^{(\mathrm{F,curve})}/I^{(\mathrm{F,const})}$ and plotted over the excitation intensity $I^{(\mathrm{ex,S})}$ in a log-log plot, see Figure E.7II. In particular, since the SNR of the first image $M^{(1)}$ has a large influence on the overall SNR and the nonlinear orders (Equation E.13) that can be reconstructed with it, $I^{(\mathrm{ex1})}$ must be taken as close as possible to the boundary of the linear to nonlinear response range of the sample. The procedure for finding these excitation irradiances and creating the fits is described in Section E.2.2. For the present example, $I^{(\mathrm{ex1})} = 1.18 kW/cm^2$ (green dotted), $I^{(\mathrm{ex2})} = 3.04 kW/cm^2$ (turquoise dotted) and $I^{(\mathrm{ex3})} = 10.99 kW/cm^2$ (magenta dotted), see Figure E.7II.

E.3.3 *Measurement of the dSAX-ISM PSF*

Beads with $d^{(\mathrm{Bead})} = 100$ nm (F8803, ThermoFisher scientific) in silicon oil were used to determine the experimental PSF. The beads used are about half of one airy unit of the effective PSF of a camera pixel $d^{(\mathrm{1AU})}\left(h_0^{(\mathrm{ISM})}\right) = 235$ nm. A simulation with the present

system parameters yields a FWHM, half of an AU, of $d^{(\mathrm{FWHM})}\left(h_0^{(\mathrm{ISM})}\right) = 152$ nm for the effective PSF and $d^{(\mathrm{FWHM})}\left(M_0^{(\mathrm{ISM})}\right) = 191$ nm for the PSF convolved with the 100 nm bead. By using a bead of this size, the actual FWHM is broadened by 26%. Since in the following the methodologies shall be compared relative to each other and not with absolute values, this broadening influence is accepted in exchange for a better SNR, due to the larger fluorescence volume.

The excitation intensities were chosen as $I^{(\mathrm{ex1})} = 91W/cm^2$, $I^{(\mathrm{ex2})} = 263W/cm^2$ and $I^{(\mathrm{ex3})} = 850W/cm^2$. The images where processed w. r. t. extraction of non-linear higher orders and CLSM/PiRe summation. From the reconstructed images a set of Region of Interest (ROI)s containing single beads was selected (magenta boxes) and a bead-representative calculated, see Figure E.8I. The bead-representatives (points) were then fitted by a 2D Gaussian (respective line).The FWHM is calculated from the fitted Gaussians using the relations Table 1.1, see Table E.3. Here, dSAX-ISM achieves a resolution gain of $1 - d^{(\mathrm{FWHM})}\left(M^{(\mathrm{NL2,ISM})}\right)/d^{(\mathrm{FWHM})}\left(M^{(\mathrm{NL2,CLSM})}\right) = 15\%$ using ISM scheme compared to CLSM at 1 AU sized virtual pinhole. In case of ISM the achieved resolution of $M^{(\mathrm{NL2})}$ is 2.6% better than for $M^{(\mathrm{NL1})}$. The dataset demonstrates, that despite the weak SNR of the useable beads the ISM based dSAX-ISM reconstruction routine is able to improve resolution by succesfully extracting the second higher order while this was not possible with the SAX scheme where $d^{(\mathrm{FWHM})}(M^{(\mathrm{NL1})}) < d^{(\mathrm{FWHM})}(M^{(\mathrm{NL2})})$. Further, the overall calculated FWHM

E.3.4 *dSAX-ISM Measurement of a Biological Sample: HeLa cells*

To analyze the method in biological cell material, a 2D multi-view image at 3 different excitation intensities of actin filaments of HeLa (for staining and fixation see Section E.3.2) cells were recorded. Instead of a PD in the detection beam-path, an emCCD camera (Figure E.6) is now used for imaging. The recorded section of a HeLa cell is shown for the central pinhole (CP), CLSM with $d^{(\mathrm{pin,CLSM})} = 1\,AU$ and PiRe with $d^{(\mathrm{pin,ISM})} = 1\,AU$ in Figure E.9.

Subsequently, the data were processed using deconvolution analogous to Section E.2.3. Figure E.10 shows the results for $M^{(0)}$ (top row) and $M^{(\mathrm{NL2})}$ (bottom row) for the ROI marked with a white rectangle in Figure E.9. Since amDEC (Figure E.10 t) is calculated

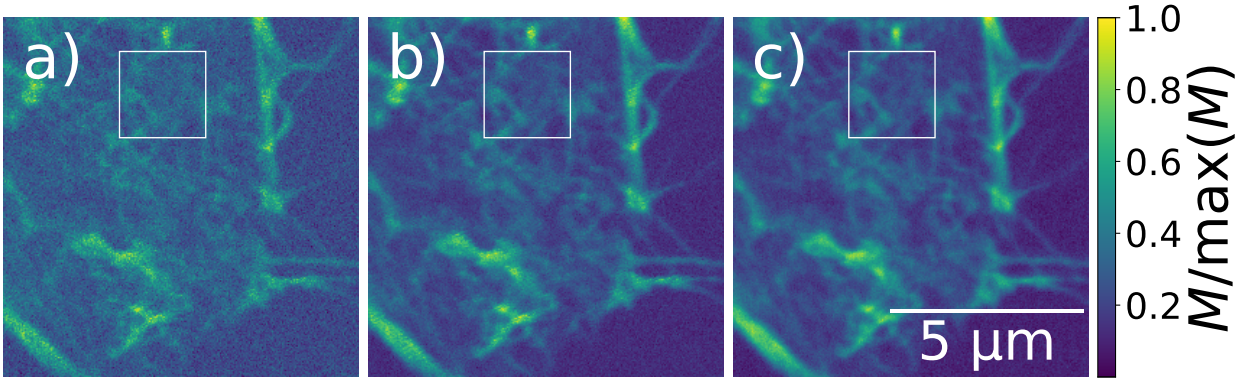


FIGURE E.9: **DSAX HeLa Image.** Reconstructed 2nd order $M^{(\mathrm{NL}_2)}$ for a) central pinhole, b) CLSM and c) ISM processing. Small boxes mark ROIused in Figure E.10.

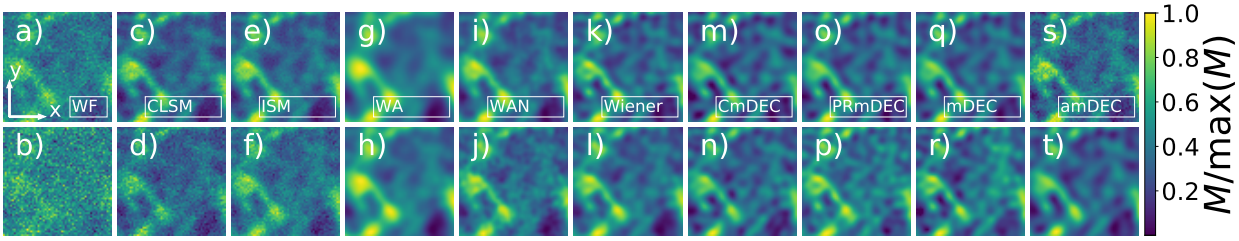


FIGURE E.10: **DSAX Reconstruction of HeLa Actin Filaments.** Reconstruction results of ROI (see Figure E.9). Reconstruction results of dSAX-ISM images. Upper row: applied to $M^{(1)}$, lower row: applied to $M^{(\mathrm{NL}_2)}$. a+b) WF, c+d) CLSM at 1 AU pinhole size, e+f) PiRe at 1 AU pinhole size, g+h) WAVG i+j) nWAVG, k+l) generalized Wiener deconvolution, m+n) confocal single-view DEConvolution (cDEC), o+p) pixel reassigned single-view DEConvolution (prDEC), q+r) mDEC , s) central pinhole image, t) amDEC. Images are normalized to their individual zero frequency.

from all available views and images $M^{(1\cdots3)}$, it is used as a reference for the NCC calculation. In particular, since for the nonlinear first and second order $M^{(\mathrm{NL}_{1\cdots2})}$ many of the available detector pixels only contribute noise to the reconstruction, the same pinhole selection was used for the multi-view deconvolutions as well as for the PiRe.

E.4 CONCLUSION AND OUTLOOK

The existing fluorophore saturation-based methodology dSAX was extended from CLSM to ISM geometry, thereby coining the name dSAX-ISM, and the influence of different processing was investigated.

In the simulations, a theoretical resolution gain of $h^{(\mathrm{NL}_2)}$ over $h^{(1)}$ of about 11% in low SNR case using the presented processing method was predicted. A simulation of different deconvolution strategies led to the remarkable result, that in case of an in-focus 2D multi-view processing the NCC-quality of the single-view cDEC as well as prDEC deconvolutions achieve about the same reconstruction quality like mDEC. Since only one view is used instead of the available 25, larger measurement datasets can be processed on more compact GPUs without compromising quality. In case of low-SNR scenarios, reducing the amount of detector pixels led to better reconstruction results.

The theoretical analyses were implemented experimentally in a dedicated setup. Here, it proved advantageous to select the excitation intensities tending to be from the already adjacent nonlinear region to achieve better SNR. The dSAX-ISM scheme was able to achieve a resolution gain for the extracted non-linear PSF $h^{(\mathrm{NL}_2)}$ over $h^{(1)}$ while dSAX did not. The reconstruction quality of cDEC and prDEC could be confirmed experimentally.

It was shown that the extension of dSAX by ISM can extract a $M^{(\mathrm{NL}_2)}$ with higher noise cutoff at the same excitation intensity. This allows comparatively lower excitation intensities to be chosen, thus reducing phototoxicity while maintaining the same resolution. Extending dSAX with ISM leads to a decrease in scan speed compared to CLSM imaging using PMT, but this could be compensated in the future by replacing the chosen emCCD camera with single photon avalanche diode (SPAD) arrays, or a PMT array (as in the commercial Zeiss Airyscan system). In future, a parameter search for an optimal SNR and FWHM as a function of excitation intensity and using five-state model (including triplet state and bleaching) could be helpful.

E.5 2 STATE DERIVATION

For analysis of SAX instead of the schematic 4-state model introduced in Section 1.5 a straightforward to implement 2-state model (Section E.2.1) is used. Assuming a lossless system (i. e. number and availability of excitable molecules remains the same), the rate equations are:

$$\partial_t N^{(0)} = -k^{(\mathrm{a})} N^{(0)} + k^{(\mathrm{f})} N^{(1)} \tag{E.16}$$

$$\partial_t N^{(1)} = -\partial_t N^{(0)} \tag{E.17}$$

where $N^{(0/1)}$ denotes the number of molecules in the $S^{(0/1)}$ state. $k^{(a)}(\vec{x}) = \eta^{(a)}(\vec{x})I^{(ex)}(\vec{x})$ the intensity dependent absorption rate, $\eta^{(a)}$ the fluorophor optical absorption cross-section, $k^{(f)} = \tau^{-1}$ the emission rate, τ the excited state lifetime, $Q^{(e)}$ the fluorophor quantum efficiency and $I^{(ex)}$ the excitation intensity. Furthermore, since the number of molecules remains the same, Equation E.18 applies and we choose $N^{(0)} \equiv N, N^{(1)}(0) \equiv 0$, since at the beginning of the excitation process all molecules are still in the ground state S_0.

$$N^{(0)}(t) + N^{(1)}(t) = N^{(0)}(0) + N^{(1)}(0) = N + 0 \tag{E.18}$$

Plugging Equation E.17 into Equation E.16 yields the condition at equilibrium $t = t^{(e)}$:

$$\frac{N^{(0)}}{N^{(1)}} = \frac{k^{(f)}}{k^{(a)}} \tag{E.19}$$

i.e. if more than two states would have been used within the model, eventual hidden-states would vanish and a balanced relation of excitation vs spontaneous relaxation will be established. By putting Equation E.19 into Equation E.18 we find:

$$N^{(1,e)} \equiv N^{(1)}(t = t^{(e)}) = \frac{k^{(a)}}{k^{(a)} + k^{(f)}} N \tag{E.20}$$

with Plugging Equation E.18 and Equation E.20 into Equation E.17 yields:

$$\partial_t N^{(1)}(t) = -k^{(f)} N^{(1)}(t) + k^{(a)}(N - N^{(1)}(t)) \tag{E.21}$$

$$= a\left(N^{(1,e)}(t) - N^{(1)}(t)\right) \tag{E.22}$$

with $a \equiv k^{(a)} + k^{(f)}$. Equation E.22 is a ordinary, inhomogeneous differential equation of 1st order and the individual solutions to the homogeneous $N^{(1,H)}(t)$ and inhomogeneous part $N^{(1,P)}(t)$ are:

$$N^{(1,H)}(t) = A\, e^{-at} \tag{E.23}$$

$$N^{(1,P)}(t) = N^{(1,e)} \tag{E.24}$$

$$N^{(1)}(t) = N^{(1,H)}(t) + N^{(1,P)}(t) \tag{E.25}$$

which together with the starting condition $N^{(1)}(t = 0) = 0$ and Equation E.25 yields:

$$N^{(1)}(t) = N^{(1,e)}\left(1 - e^{-kt}\right) \tag{E.26}$$

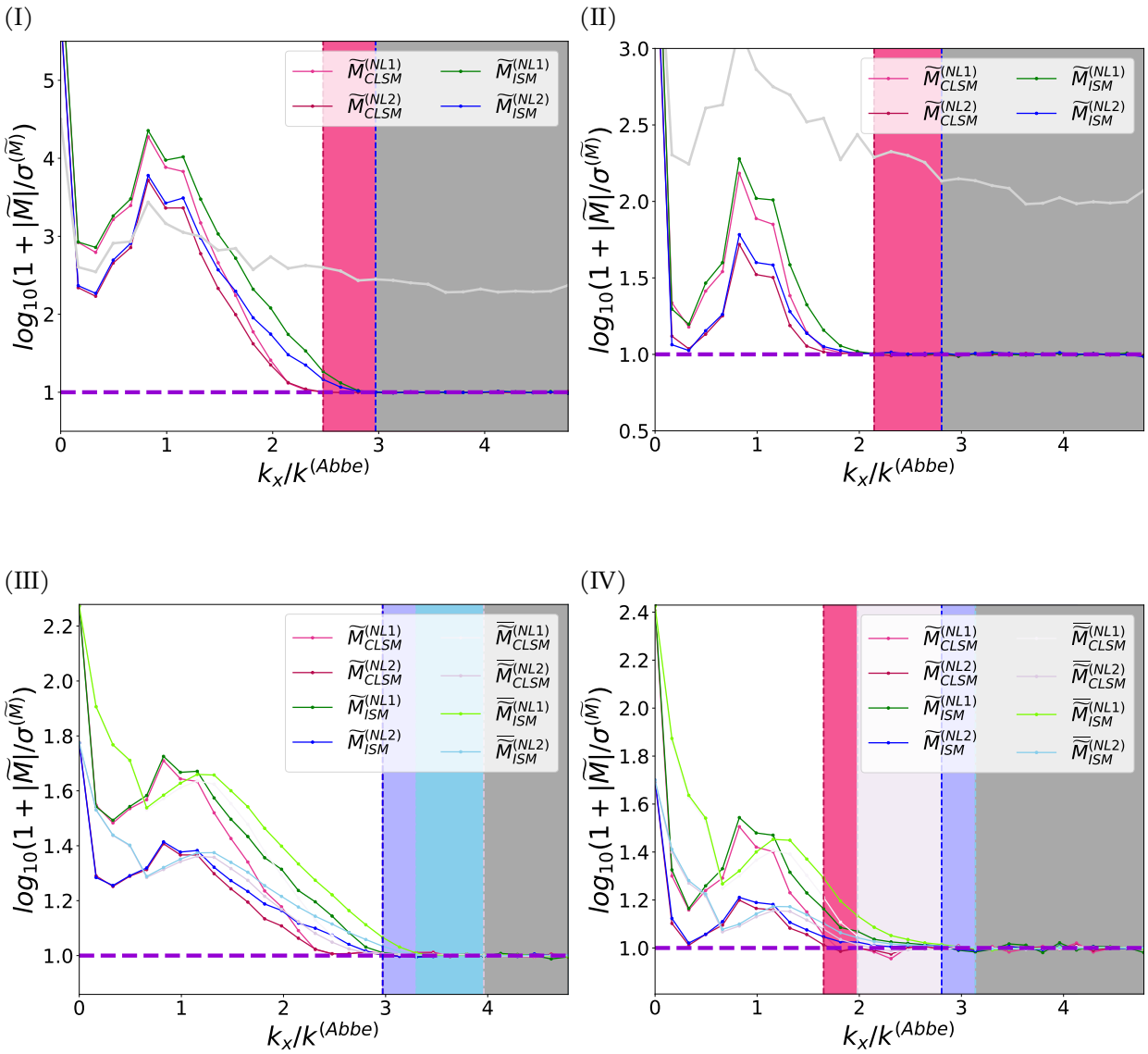


FIGURE E.11: **DSAX SNR Improvement by ISM.** (I-II) radialsum profile over average images. (III-IV) radialsum with moving average from 1 noisy image.

E.6 SNR IN DSAX-ISM

Previous SAX implementations are based on a CLSM measurement setup as described in Section E.1. In this section, the combination of SAX with a ISM scheme, leading to the dSAX-ISM method, to improve SNR will be explored. Thus, SAX and dSAX-ISM results will be compared.

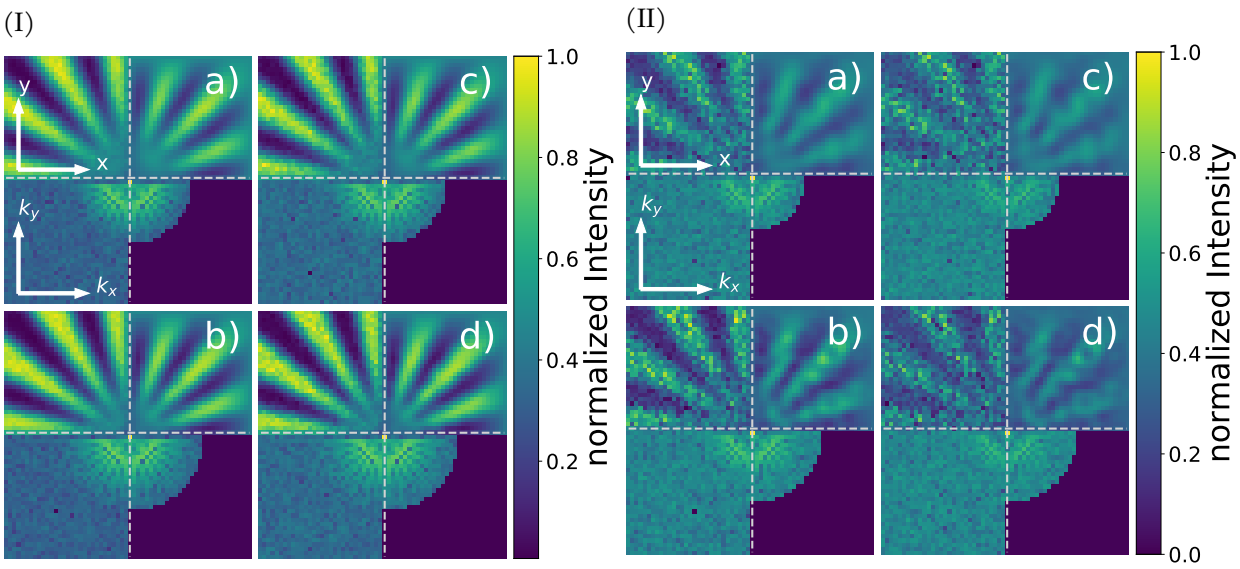


FIGURE E.12: **DSAX SNR Improvement Images.**Images used for Figure E.11. (I) for high-SNR, (II) low-SNR case. a) $M^{(\mathrm{NL1})}_{CLSM}$, b) $M^{(\mathrm{NL2})}_{CLSM}$ c) $M^{(\mathrm{NL1})}_{ISM}$, d) $M^{(\mathrm{NL2})}_{ISM}$. Per Image the upper row: real space, lower row: Fourier transform of respective upper row image-parts. In each for a)-d) left colum: simulated, noisy image; right: low-pass-filtered at cut-off frequency of $M^{(\mathrm{NL2})}$.

In the simulation, some parameters have been changed compared to Section E.2.2 as follows: the detector array now has $[5, 5]$, the detector pixel pitch is $d^{(\mathrm{pin})} = 0.25\,AU$ and the quantum efficiency fluorophore density product $Q^{(\mathrm{e})}C^{(\mathrm{F})}$ is reduced by 100× for higher SNR and increased by 10× for lower SNR. For noise statistics analysis, 100 realizations of the same measurement process are performed.

For analysis of the SNR (Equation 1.33) improvement by the ISM over the CLSM scheme, the cut-off frequency in Fourier space is investigated, see Figure E.11II. For this, the images are Fourier transformed, the transforms are divided into radial bins and averaged within them. Here $N^{(\mathrm{bins})} = 32$, i.e. half of the lateral image size, was used. The projection thus determined was then noise-normalized (noise-baseline, purple dashed line) and the first value of $\widetilde{M}^{(\mathrm{NL2})}_{CLSM}$ (dark red curve) lying below the noise-baseline is considered to be indistinguishable from the noise-floor. Therefore, the corresponding cutoff frequency $k^{(\mathrm{C,CLSM})}$ (vertical red dashed line) is chosen one simulate frequency point lower. Analogously, $\widetilde{M}^{(\mathrm{NL2})}_{ISM}$ (dark blue curve) with $k^{(\mathrm{C,ISM})}$. The area, starting from $k^{(\mathrm{C,CLSM})}$ is colored in a lighter red tone, correspondingly the area starting from $k^{(\mathrm{C,ISM})}$ in

	$M^{(1)}$	$M^{(2)}$	$M^{(3)}$	$M^{(NL1)}$	$M^{(NL2)}$
CLSM	179	190	218	149	136
ISM	150	154	165	130	121
1-ISM/CLSM	16	19	24	13	11

TABLE E.3: **DSAX FWHM of SNR Improvement by ISM.** Calculated FWHM from simulated, noise-free PSF. First two rows in *nm*, last in %.

a lighter blue tone. Thus, it is directly evident which curve has a higher cutoff frequency. From this cutoff area, the resolution and SNR improvement between the two schemes can be read. The area from the highest cutoff frequency achieved is colored gray, indicating that these frequencies cannot be distinguished from the noise floor for either technique. The light gray curve in the upper part of the image indicates the underlying object spectrum and is freely scaled.

The left column of Figure E.11 (i. e. (I),(III)) shows the results for the high-SNR simulation, the right column the results for lower-SNR simulation. For 100 realizations in the high-SNR case, the cutoff frequencies are $k^{(\mathrm{C,CLSM})}/k^{(\mathrm{C,ISM})} = 0.83\times$ (see Figure E.11I) while for the low-SNR case for the cutoff frequencies $k^{(\mathrm{C,CLSM})}/k^{(\mathrm{C,ISM})} = 0.77\times$ holds (see Figure E.11II). $k^{(\mathrm{C,ISM})} \approx 3k^{(\mathrm{Abbe})}$ for high-SNR and $k^{(\mathrm{C,ISM})} \approx 2k^{(\mathrm{Abbe})}$ for low-SNR. For e. g. $k_x \approx 2k^{(\mathrm{Abbe})}$, the transfer ratio of the two techniques in the high-SNR case is $\widetilde{M}_{ISM}^{(\mathrm{NL2})}/\widetilde{M}_{CLSM}^{(\mathrm{NL2})} \approx 1.059\times$ and in the low-SNR case at $\widetilde{M}_{ISM}^{(\mathrm{NL2})}/\widetilde{M}_{CLSM}^{(\mathrm{NL2})} \approx 1.007\times$.

In the case of evaluating a single noisy image, the maximum reconstructable cutoff frequency decreases by 8% for ISM and 16% for CLSM in both noise analyses (see Figure E.11III, Figure E.11IV). In addition, moving averages $\overline{\widetilde{M}}$ were calculated for these plots with a window-length $w = 4\,pixels \approx 0.67k^{(\mathrm{abbe})}$. In particular, due to the sharp drop of the frequency curve to the respective cutoff frequency, the weighted averaging gets a momentum and falls below the noise floor only at higher frequencies than the non-averaged curves. In the case of high-SNR, the cutoff frequency of $\overline{\widetilde{M}}_{CLSM}^{(\mathrm{NL2})}$ is $k^{(\mathrm{C,CLSM,avg})} \approx 3.47 \cdot k^{(\mathrm{Abbe})}$ and of $\overline{\widetilde{M}}_{ISM}^{(\mathrm{NL2})}$ at $k^{(\mathrm{C,ISM,avg})} \approx 3.63 \cdot k^{(\mathrm{Abbe})}$. The transfer ratio at $k_x \approx 2k^{(\mathrm{Abbe})}$ is $\overline{\widetilde{M}}_{ISM}^{(\mathrm{NL2})}/\overline{\widetilde{M}}_{CLSM}^{(\mathrm{NL2})} \approx 1.036\times$. For the low-SNR case, the cutoff frequency of $\overline{\widetilde{M}}_{CLSM}^{(\mathrm{NL2})}$ is $k^{(\mathrm{C,CLSM,avg})} \approx 1.98 \cdot k^{(\mathrm{Abbe})}$ and of $\overline{\widetilde{M}}_{ISM}^{(\mathrm{NL2})}$ at $k^{(\mathrm{C,ISM,avg})} \approx 3.14 \cdot k^{(\mathrm{Abbe})}$. The transfer ratio at $k_x \approx 2k^{(\mathrm{Abbe})}$ is $\overline{\widetilde{M}}_{ISM}^{(\mathrm{NL2})}/\overline{\widetilde{M}}_{CLSM}^{(\mathrm{NL2})} \approx 1.015\times$.

The simulated data are shown in Figure E.12I for high-SNR and in Figure E.12II for low-SNR. Each individual image, e. g. Figure E.12I a), consists of four subimages, all normalized to their individual maximum. Top left: Image in real space, top right: SNR low-pass filtered with the particular $k^{(C)}$, bottom row: Fourier transformed representations of the top row. Hence the panels display: a) SNR from 100 realizations (see e. g. Figure E.11I) for CLSM, b) same for ISM, c) one randomly chosen image from the 100 realizations (see e. g. Figure E.11III) for CLSM, and d) same for ISM. Due to the higher SNR of ISM compared to CLSM, the struts of the *spokes2d* target can still be well separated, especially for smaller radii.

A comparison of the FWHM of simulated, noise-free PSFs shows the relative resolution gain of ISM over CLSM, see Table E.3. For increasing saturation of PSF, the resolution gain increases by up to 24% for $M^{(3)}$ by ISM, while it increases progressively less towards higher reconstructed nonlinear orders. Thus, for $M^{(NL_2)}$ only a resolution gain of 11% is achieved.

F

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] Ron Milo and Rob Phillips. *Cell Biology by the Numbers*. 1st ed. Garland Science, 2015, p. 400. ISBN: 978-0815345374.

[2] E. Abbe. "Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung". In: *Archiv für Mikroskopische Anatomie* 9.1 (Dec. 1873), pp. 413–418. ISSN: 0176-7364. DOI: 10.1007/BF02956173.

[3] Cecie Starr, Christine A. Evers, and Lisa Starr. *Biology - Concepts and Applications*. 10th ed. Cengage Learning, 2018, p. 885. ISBN: 9781305967335.

[4] Marvin Minsky. "US3013467a, M. Minsky Microscopy Apparatus". In: (1957), p. 5. URL: https://patents.google.com/patent/US3013467A/en?oq=US3013467.

[5] C. J.R. Sheppard. "Super-resolution in confocal imaging". In: *Optik (Jena)* 80.2 (1988), pp. 53–54. ISSN: 00304026. DOI: 10.1117/12.967201.

[6] Claus B. Müller and Jörg Enderlein. "Image scanning microscopy". In: *Physical Review Letters* 104.19 (2010), pp. 1–4. ISSN: 00319007. DOI: 10.1103/PhysRevLett.104.198101.

[7] Andrew G York et al. "Instant super-resolution imaging in live cells and embryos via analog image processing." In: *Nature methods* 10.11 (2013), pp. 1122–6. ISSN: 1548-7105. DOI: 10.1038/nmeth.2687. URL: http://dx.doi.org/10.1038/nmeth.2687.

[8] Giulia M.R. De Luca et al. "Re-scan confocal microscopy: scanning twice for better resolution". In: *Biomedical Optics Express* 4.11 (Nov. 2013), p. 2644. ISSN: 2156-7085. DOI: 10.1364/BOE.4.002644. URL: https://www.osapublishing.org/abstract.cfm?URI=boe-4-11-2644.

[9] Stephan Roth et al. "Optical photon reassignment microscopy (OPRA)". In: *Optical Nanoscopy* 2.1 (2013), p. 5. ISSN: 2192-2853. DOI: 10.1186/2192-2853-2-5. arXiv: 1306.6230. URL: http://www.optnano.com/content/2/1/5.

[10] Takuya Azuma and Takayuki Kei. "Super-resolution spinning-disk confocal microscopy using optical photon reassignment". In: *Optics Express* 23.11 (2015), p. 15003. ISSN: 1094-4087. DOI: 10.1364/oe.23.015003. URL: http://www.osapublishing.org/viewmedia.cfm?uri=oe-23-11-15003%7B%5C&%7Dseq=0%7B%5C&%7Dhtml=true.

[11] Olaf Schulz et al. "Resolution doubling in fluorescence microscopy with confocal spinning-disk image scanning microscopy". In: *Proceedings of the National Academy of Sciences of the United States of America* 110.52 (Dec. 2013), pp. 21000–21005. ISSN: 00278424. DOI: 10.1073/pnas.1315858110. URL: http://www.pnas.org/cgi/doi/10.1073/pnas.1315858110.

[12] Clemens Roider et al. "Deconvolution approach for 3D scanning microscopy with helical phase engineering". In: *Optics Express* 24.14 (2016), p. 15456. ISSN: 1094-4087. DOI: 10.1364/oe.24.015456.

[13] Stephan Roth, Colin J. R. Sheppard, and Rainer Heintzmann. "Superconcentration of light: circumventing the classical limit to achievable irradiance". In: *Optics Letters* 41.9 (May 2016), p. 2109. ISSN: 0146-9592. DOI: 10.1364/OL.41.002109. URL: https://www.osapublishing.org/abstract.cfm?URI=ol-41-9-2109.

[14] Jerome Mertz. *Introduction to Optical Microscopy*. 2nd ed. Cambridge University Press, 2019. ISBN: 978-1108428309.

[15] Joseph W. Goodman. *Introduction to Fourier Optics*. 4th ed. W. H. Freeman, 2017. ISBN: 978-1319119164.

[16] Max Born and Emil Wolf. *Principles of Optics - Electromagnetic Theory of Propagation, Interference and Diffraction of Light*. 7th ed. Cambridge University Press, 1999. ISBN: 978-0521642224.

[17] Dimitri P. Bertsekas and John N. Tsitsiklis. *Introduction To Probability*. 2nd. Athena Scientific, 2008. ISBN: 978-1886529236.

[18] Athanasios Papoulis and Joseph G. Hoffman. "Probability, Random Variables, and Stochastic Processes". In: *Physics Today* 20.1 (Jan. 1967), pp. 135–135. ISSN: 0031-9228. DOI: 10.1063/1.3034123. URL: http://physicstoday.scitation.org/doi/10.1063/1.3034123.

[19] Wolfgang Walter. *Analysis 2*. 5th. Springer-Lehrbuch. Berlin, Heidelberg: Springer Berlin Heidelberg, 2002. ISBN: 978-3-540-58666-1. DOI: 10.1007/978-3-642-97614-8. URL: http://link.springer.com/10.1007/978-3-642-97614-8.

[20] Rainer Heintzmann. "Resolution enhancement of biological light microscopic data". In: (1999).

[21] A. Lipson, S.G. Lipson, and H. Lipson. *Optical Physics*. 4th ed. CUP, 2010, p. 592. ISBN: 9780521493451.

[22] Robert P J Nieuwenhuizen et al. "Measuring image resolution in optical nanoscopy". In: *Nature methods* 10.6 (2013), pp. 557–62. ISSN: 1548-7105. DOI: 10.1038/nmeth.2448. arXiv: NIHMS150003. URL: http://www.ncbi.nlm.nih.gov/pubmed/23624665.

[23] Rainer Heintzmann et al. "Resolution enhancement by subtraction of confocal signals taken at different pinhole sizes". In: *Micron* 34.6-7 (2003), pp. 293–300. ISSN: 09684328. DOI: 10.1016/S0968-4328(03)00054-4.

[24] Rainer Heintzmann. "Band-limit and appropriate sampling in microscopy". In: *Cell Biology: A Laboratory Handbook*. 2006. Chap. 3. ISBN: 978-0121647261.

[25] Andrew G. York et al. "Resolution doubling in live, multicellular organisms via multifocal structured illumination microscopy". In: *Nature Methods* 9.7 (May 2012), pp. 749–754. ISSN: 1548-7091. DOI: 10.1038/nmeth.2025. arXiv: 1705.11170. URL: http://www.nature.com/doifinder/10.1038/nmeth.2025.

[26] Peter W Winter et al. "Two-photon instant structured illumination microscopy improves the depth penetration of super-resolution imaging in thick scattering samples." In: *Optica* 1.3 (2014), pp. 181–191. ISSN: 2334-2536. DOI: 10.1364/OPTICA.1.000181. arXiv: NIHMS150003. URL: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4256096%7B%5C&%7Dtool=pmcentrez%7B%5C&%7Drendertype=abstract.

[27] Jürgen Popp et al. *Handbook of Biophotonics*. 1st. Wiley-VCH, 2012. ISBN: 9783527407286.

[28] Peter Atkins and Julio de Paula. *Atkin's Physical Chemistry*. 8th. Oxford University Press, 2006. ISBN: 0-7167-8759-8.

[29] Hernando Sosa, Ana B. Asenjo, and Erwin J.G. Peterman. "Structure and Dynamics of the Kinesin–Microtubule Interaction Revealed by Fluorescence Polarization Microscopy". In: 2010, pp. 505–519. DOI: 10.1016/S0091-679X(10)95025-5. URL: https://linkinghub.elsevier.com/retrieve/pii/S0091679X10950255.

[30] Thermofisher. *Alexa-488 Properties*. 2022. URL: https://www.thermofisher.com/de/de/home/references/molecular-probes-the-handbook/fluorophores-and-their-amine-reactive-derivatives/alexa-fluor-dyes-spanning-the-visible-and-infrared-spectrum.html.

[31] E. A. Ingerman et al. "Signal, noise and resolution in linear and nonlinear structured-illumination microscopy". In: *Journal of Microscopy* 273.1 (2019), pp. 3–25. ISSN: 13652818. DOI: 10.1111/jmi.12753.

[32] Karl Pearson. "Notes on the History of Correlation". In: *Biometrika* 13.1 (Oct. 1920), p. 25. ISSN: 00063444. DOI: 10.2307/2331722. URL: https://academic.oup.com/nq/article/s13-I/21/406/4236555%20https://www.jstor.org/stable/2331722?origin=crossref.

[33] W. M. Gorman and Stefan Valavanis. "Econometrics: An Introduction to Maximum Likelihood Methods." In: *Journal of the Royal Statistical Society. Series A (General)* 123.4 (1960), p. 494. ISSN: 00359238. DOI: 10.2307/2343291. URL: https://www.jstor.org/stable/10.2307/2343291?origin=crossref.

[34] Norbert Wiener. *Extrapolation, Interpolation, and Smoothing of Stationary Time Series*. The MIT Press, 1949. ISBN: 9780262257190. DOI: 10.7551/mitpress/2946.001.0001. URL: https://direct.mit.edu/books/book/4361/Extrapolation-Interpolation-and-Smoothing-of.

[35] Kai Wicker. "Increasing resolution and light efficiency in fluorescence microscopy". PhD thesis. 2010.

[36] Leonid I Rudin, Stanley Osher, and Emad Fatemi. "Nonlinear total variation based noise removal algorithms". In: *Physica D: Nonlinear Phenomena* 60.1-4 (Nov. 1992), pp. 259–268. ISSN: 01672789. DOI: 10.1016/0167-2789(92)90242-F. URL: http://linkinghub.elsevier.com/retrieve/pii/016727899290242F%7B%5C%%7D0Apapers3://publication/doi/10.1016/0167-2789(92)90242-F%20https://linkinghub.elsevier.com/retrieve/pii/016727899290242F.

[37] I. J. Good. "Non-parametric Roughness Penalty for Probability Densities". In: *Nature Physical Science* 229.1 (Jan. 1971), pp. 29–30. ISSN: 0300-8746. DOI: 10.1038/physci229029a0. URL: http://www.nature.com/articles/physci229029a0.

[38] Bionanoimaging. *NanoImagingPack (GIT)*. URL: https://gitlab.com/bionanoimaging/nanoimagingpack.

[39] Ciyou Zhu et al. "Algorithm 778: L-BFGS-B". In: *ACM Transactions on Mathematical Software* 23.4 (Dec. 1997), pp. 550–560. ISSN: 0098-3500. DOI: 10.1145/279232.279236.

[40] Scipy. *Scipy*. URL: https://github.com/scipy/scipy/tree/v1.3.3.

[41] William H. Press et al. *Numerical Recipes 3rd Edition: The Art of Scientific Computing*. 3rd. Cambridge University Press, 2007. ISBN: 0521880688. DOI: 10.5555/1403886.

[42] Ilja Gubins. *Tiler*. URL: https://github.com/the-lay/tiler.

[43] Python Software Foundation. *Python*. URL: https://www.python.org/.

[44] Rene Lachmann. *MicroPy (GIT)*. URL: https://github.com/renerichter/micropy/releases/tag/v0.7.1a.

[45] Numpy. *Numpy*. URL: https://github.com/numpy/numpy/tree/v1.16.6.

[46] Google. *Tensorflow (GIT)*. 2019. URL: https://github.com/tensorflow/tensorflow/tree/v1.15.0.

[47] Benedict Diederich et al. *UC2 Github Hardware Repository*. 2019. URL: https://github.com/bionanoimaging/UC2-GIT (visited on 09/23/2019).

[48] Hongki Yoo et al. "Error analysis and tolerance allocation for confocal scanning microscopy using the Monte Carlo method". In: *Three-Dimensional and Multi-dimensional Microscopy: Image Acquisition and Processing XI*. Ed. by Jose-Angel Conchello, Carol J. Cogswell, and Tony Wilson. Vol. 5324. July 2004, p. 242. DOI: 10.1117/12.529942. URL: http://proceedings.spiedigitallibrary.org/proceeding.aspx?doi=10.1117/12.529942.

[49] Zichao Bian et al. *Autofocusing technologies for whole slide imaging and automated microscopy*. 2020. DOI: 10.1002/jbio.202000227.

[50] von F. Zernike. "Beugungstheorie des schneidenver-fahrens und seiner verbesserten form, der phasenkontrastmethode". In: *Physica* 1.7-12 (May 1934), pp. 689–704. ISSN: 00318914. DOI: 10.1016/S0031-8914(34)80259-5. URL: https://linkinghub.elsevier.com/retrieve/pii/S0031891434802595.

[51] Lawrence Firestone et al. "Comparison of autofocus methods for automated microscopy". In: *Cytometry* 12.3 (1991), pp. 195–206. ISSN: 0196-4763. DOI: 10.1002/cyto.990120302. URL: http://doi.wiley.com/10.1002/cyto.990120302.

[52] Yu Sun, Stefan Duthaler, and Bradley J. Nelson. "Autofocusing in computer microscopy: Selecting the optimal focus algorithm". In: *Microscopy Research and Technique* 65.3 (Oct. 2004), pp. 139–149. ISSN: 1059-910X. DOI: 10.1002/jemt.20118. URL: http://doi.wiley.com/10.1002/jemt.20118.

[53] A. SANTOS et al. "Evaluation of autofocus functions in molecular cytogenetic analysis". In: *Journal of Microscopy* 188.3 (Dec. 1997), pp. 264–272. ISSN: 0022-2720. DOI: 10.1046/j.1365-2818.1997.2630819.x. URL: https://onlinelibrary.wiley.com/doi/abs/10.1046/j.1365-2818.1997.2630819.x.

[54] Abraham. Savitzky and M. J. E. Golay. "Smoothing and Differentiation of Data by Simplified Least Squares Procedures." In: *Analytical Chemistry* 36.8 (July 1964), pp. 1627–1639. ISSN: 0003-2700. DOI: 10.1021/ac60214a047. URL: https://pubs.acs.org/doi/abs/10.1021/ac60214a047.

[55] Ronald W. Schafer. "What is a savitzky-golay filter?" In: *IEEE Signal Processing Magazine* 28.4 (2011), pp. 111–117. ISSN: 10535888. DOI: 10.1109/MSP.2011.941097.

[56] Python3. *STD-library – Timeit*. URL: https://docs.python.org/3/library/timeit.html.

[57] Benedict Diederich et al. "A versatile and customizable low-cost 3D-printed open standard for microscopic imaging". In: *Nature Communications* 11.1 (Dec. 2020), p. 5979. ISSN: 2041-1723. DOI: 10.1038/s41467-020-19447-9. URL: http://dx.doi.org/10.1038/s41467-020-19447-9%20http://www.nature.com/articles/s41467-020-19447-9.

[58] Ronny Förster et al. "Automated distinction of shearing and distortion artefacts in structured illumination microscopy". In: *Optics Express* 26.16 (Aug. 2018), p. 20680. ISSN: 1094-4087. DOI: 10.1364/oe.26.020680. URL: https://www.osapublishing.org/abstract.cfm?URI=oe-26-16-20680.

[59] G. P. Downey et al. "Retention of leukocytes in capillaries: Role of cell size and deformability". In: *Journal of Applied Physiology* 69.5 (1990), pp. 1767–1778. ISSN: 01617567. DOI: 10.1152/jappl.1990.69.5.1767.

[60] Mathieu P Rodero et al. "In vivo imaging reveals a pioneer wave of monocyte recruitment into mouse skin wounds". In: *PLoS ONE* 9.10 (Oct. 2014). Ed. by Christophe Egles, e108212. ISSN: 19326203. DOI: 10.1371/journal.pone.0108212. URL: https://dx.plos.org/10.1371/journal.pone.0108212.

[61] Daniel Bremer et al. "Longitudinal intravital imaging of the retina reveals long-term dynamics of immune infiltration and its effects on the glial network in experimental autoimmune uveoretinitis, without evident signs of neuronal dysfunction in the ganglion cell layer". In: *Frontiers in Immunology* 7.DEC (2016), pp. 1–15. ISSN: 16643224. DOI: 10.3389/fimmu.2016.00642.

[62] M. Bertero and P. Boccacci. *Introduction to Inverse Problems in Imaging*. 1st ed. CRC Press, Aug. 1998. ISBN: 9780367806941. DOI: 10.1201/9780367806941. URL: https://www.taylorfrancis.com/books/9781439822067.

[63] Gene H Golub, Per Christian Hansen, and Dianne P. O'Leary. "Tikhonov Regularization and Total Least Squares". In: *SIAM Journal on Matrix Analysis and Applications* 21.1 (Jan. 1999), pp. 185–194. ISSN: 0895-4798. DOI: 10.1137/S0895479897326432. URL: http://epubs.siam.org/doi/10.1137/S0895479897326432.

[64] Bo Huang et al. "Three-dimensional super-resolution imaging by stochastic optical reconstruction microscopy". In: *Science* 319.5864 (2008), pp. 810–813. ISSN: 00368075. DOI: 10.1126/science.1153529.

[65] Klaus Weisshart. "The Basic Principle of Airyscanning". In: *Zeiss Microscopy* (2014), p. 22.

[66] HDF5. *H5PY*. URL: https://github.com/h5py/h5py.

[67] Bionanoimaging. *InverseModelling (GIT)*. URL: https://gitlab.com/bionanoimaging/inversemodelling.

[68] René Lachmann, Benedict Diederich, and Xavier Uwurukundu. *UC2 Github Software Repository*. URL: https://github.com/bionanoimaging/UC2-Software-GIT/ (visited on 09/23/2019).

[69] I.N. Bronstein et al. *Handbuch der Mathematik I*. Ed. by Zeidler. Springer, 2013. ISBN: 978-3658002848.

[70] Herbert Gross, Wolfgang Singer, and Michael Totzeck. *Handbook of Optical Systems: Physical Image Formation*. Vol. 2. Wiley-VCH Verlag GmbH & Co. KGaA, 2006, pp. 1–690. ISBN: 9783527403783. DOI: 10.1002/3527606688.

[71] I.N. Bronstein and K.A. Semendjaew. *Springer-Handbuch der Mathematik III*. Ed. by Eberhard Zeidler. Wiesbaden: Springer Fachmedien Wiesbaden, 2013. ISBN: 978-3-658-00274-9. DOI: 10.1007/978-3-658-00275-6. URL: http://link.springer.com/10.1007/978-3-658-00275-6.

[72] J. Parsons et al. "A comparison of techniques used to simulate the scattering of electromagnetic radiation by metallic nanostructures". In: *Journal of Modern Optics* 57.5 (Mar. 2010), pp. 356–365. ISSN: 0950-0340. DOI: 10.1080/09500341003628702. URL: http://www.tandfonline.com/doi/abs/10.1080/09500341003628702.

[73] Maxim A. Yurkin et al. "Systematic comparison of the discrete dipole approximation and the finite difference time domain method". In: *Conference on Electromagnetic and Light Scattering*. Vol. 15. 26. Connecticut: Begellhouse, 2007, pp. 249–252. ISBN: 978-1-56700-244-7. DOI: 10.1615/ICHMT.2007.ConfElectromagLigScat.650. URL: http://www.dl.begellhouse.com/references/1bb331655c289a0a,2a1f17e2753400d3,3e2d41593af820b4.html.

[74] Thanh-an Pham et al. "Three-Dimensional Optical Diffraction Tomography With Lippmann-Schwinger Model". In: *IEEE Transactions on Computational Imaging* 6 (2020), pp. 727–738. ISSN: 2573-0436. DOI: 10.1109/tci.2020.2969070. arXiv: 1910.13844.

[75] L.D. Landa and E.M. Lifshitz. *Course of Theoretical Physics - Quantum Mechanics*. 2nd. Pergamon Press, 1965.

[76] Hamamatsu. *Photomultiplier Tubes - Basics and Applications*. 2007. URL: https://www.hamamatsu.com/resources/pdf/etd/PMT%7B%5C_%7Dhandbook%7B%5C_%7Dv3aE.pdf.

[77] Lucas J Van Vliet, Damir Sudar, and Ian T Young. "Digital Fluorescence Imaging Using Cooled CCD Array Cameras invisible". In: *Cell Biology* III.Castleman 1996 (1998), pp. 109–120. URL: www:%7B%5C%%7D0Ahttp://www.ph.tn.tudelft.nl/%7B~%7Dlucas.

[78] Hamamatsu. *ImagEM X-2*. 2020. URL: https://camera.hamamatsu.com/resources/pdf/sys/SCAS0092E%7B%5C_%7DIMAGEMX2s.pdf.

[79] Hamamatsu. *ORCA-Flash4.0 V3*. 2018. URL: https://www.hamamatsu.com/content/dam/hamamatsu-photonics/sites/documents/99%7B%5C_%7DSALES%7B%5C_%7DLIBRARY/sys/SCAS0134E%7B%5C_%7DC13440-20CU%7B%5C_%7Dtec.pdf.

[80] Robin van den Eynde et al. "Quantitative comparison of camera technologies for cost-effective super-resolution optical fluctuation imaging (SOFI)". In: *JPhys Photonics* 1.4 (2019). ISSN: 25157647. DOI: 10.1088/2515-7647/ab36ae.

[81] Rainer Heintzmann. "Estimating missing information by maximum likelihood deconvolution". In: *Micron* 38.2 (2007), pp. 136–144. ISSN: 09684328. DOI: 10.1016/j.micron.2006.07.009.

[82] Jacques Cohen-Tannoudji, Claude Dupont-Roc and Gilbert Grynberg. "Index". In: *Photons and Atoms*. Weinheim, Germany: Wiley-VCH Verlag GmbH, Dec. 2007, pp. 459–468. ISBN: 0-471-18433-0. DOI: 10.1002/9783527618422.index. URL: https://onlinelibrary.wiley.com/doi/10.1002/9783527618422.index.

[83] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. 7th. Cambridge University Press, 2009.

[84] D. A. Fish et al. "Blind deconvolution by means of the Richardson–Lucy algorithm". In: *Journal of the Optical Society of America A* 12.1 (1995), p. 58. ISSN: 1084-7529. DOI: 10.1364/josaa.12.000058.

[85] L. B. Lucy. "An iterative technique for the rectification of observed distributions". In: *The Astronomical Journal* 79.6 (1974), p. 745. ISSN: 00046256. DOI: 10.1086/111605. arXiv: arXiv:1011.1669v3. URL: http://adsabs.harvard.edu/cgi-bin/bib%7B%5C_%7Dquery?1974AJ.....79..745L.

[86] L. A. Shepp and Y. Vardi. "Maximum Likelihood Reconstruction for Emission Tomography". In: *IEEE Transactions on Medical Imaging* 1.2 (1982), pp. 113–122. ISSN: 1558254X. DOI: 10.1109/TMI.1982.4307558.

[87] Loïc A. Royer et al. "Adaptive light-sheet microscopy for long-term, high-resolution imaging in living organisms". In: *Nature Biotechnology* 34.12 (2016), pp. 1267–1278. ISSN: 15461696. DOI: 10.1038/nbt.3708. arXiv: arXiv:1705.04300v1.

[88] J F Brenner et al. "An automated microscope for cytologic research a preliminary evaluation." In: *Journal of Histochemistry & Cytochemistry* 24.1 (Jan. 1976), pp. 100–111. ISSN: 0022-1554. DOI: 10.1177/24.1.1254907. URL: http://journals.sagepub.com/doi/10.1177/24.1.1254907.

[89] J. F Schlag et al. "Implementation of automatic focusing algorithms for a computer vision system with camera control". In: *The Robotics Institute Technical Report CMU-RI-TR-83-14, Camegie-Mellon University, Pittsburgh, PA* August (1983). URL: https://www.ri.cmu.edu/pub%7B%5C_%7Dfiles/pub3/schlag%7B%5C_%7Djohn%7B%5C_%7Df%7B%5C_%7D1983%7B%5C_%7D1/schlag%7B%5C_%7Djohn%7B%5C_%7Df%7B%5C_%7D1983%7B%5C_%7D1.pdf.

[90] Irwin Sobel. "An Isotropic 3x3 image gradient operator". In: *Machine Vision for Three-Dimention Scenes,In H. Freeman editor* June (1990), pp. 376–379.

[91] D. Vollath. "The influence of the scene parameters and of noise on the behaviour of automatic focusing algorithms". In: *Journal of Microscopy* 151.2 (1988), pp. 133–146. ISSN: 13652818. DOI: 10.1111/j.1365-2818.1988.tb04620.x.

[92] Dieter Vollath. *Verfahren und Einrichtung zur automatischen Scharfeinstellung eines Bildes*. 1980. URL: https://patents.google.com/patent/EP0017726A1/de.

[93] Matej Kristan et al. "A Bayes-spectral-entropy-based measure of camera focus using a discrete cosine transform". In: *Pattern Recognition Letters* 27.13 (Oct. 2006), pp. 1431–1439. ISSN: 01678655. DOI: 10.1016/j.patrec.2006.01.016. URL: https://linkinghub.elsevier.com/retrieve/pii/S0167865506000596.

[94] Herbert Gross et al. *Handbook of Optical Systems: Aberration Theory and Correction of Optical Systems*. 1st ed. Wiley-VCH, 2006, p. 780. ISBN: 978-3527403790.

[95] Jörg Liesen and Volker Mehrmann. *Linear Algebra*. Springer Undergraduate Mathematics Series. Cham: Springer International Publishing, 2015. ISBN: 978-3-319-24344-3. DOI: 10.1007/978-3-319-24346-7. URL: http://link.springer.com/10.1007/978-3-319-24346-7.

[96] K. VISSCHER, G. J. BRAKENHOFF, and T. D. VISSER. "Fluorescence saturation in confocal microscopy". In: *Journal of Microscopy* 175.2 (1994), pp. 162–165. ISSN: 13652818. DOI: 10.1111/j.1365-2818.1994.tb03479.x.

[97] Rainer Heintzmann, Thomas M Jovin, and Christoph Cremer. "Saturated patterned excitation microscopy–a concept for optical resolution improvement." In: *Journal of the Optical Society of America. A, Optics, image science, and vision* 19.8 (2002), pp. 1599–1609. ISSN: 1084-7529. DOI: 10.1364/JOSAA.19.001599.

[98] Mats G L Gustafsson. "Nonlinear structured-illumination microscopy: wide-field fluorescence imaging with theoretically unlimited resolution." In: *Proceedings of the National Academy of Sciences of the United States of America* 102.37 (2005), pp. 13081–13086. ISSN: 0027-8424. DOI: 10.1073/pnas.0406877102.

[99] Gianguido C Cianci, Jianrong Wu, and Keith M Berland. "Saturation modified point spread functions in two-photon microscopy". In: *Microscopy Research and Technique* 64.2 (June 2004), pp. 135–141. ISSN: 1059910X. DOI: 10.1002/jemt.20071. URL: http://doi.wiley.com/10.1002/jemt.20071.

[100] Katsumasa Fujita et al. "High-resolution confocal microscopy by saturated excitation of fluorescence". In: *Physical Review Letters* 99.22 (Nov. 2007), p. 228105. ISSN: 00319007. DOI: 10.1103/PhysRevLett.99.228105. URL: http://link.aps.org/doi/10.1103/PhysRevLett.99.228105.

[101] Jörg Enderlein. "Breaking the diffraction limit with dynamic saturation optical microscopy". In: *Applied Physics Letters* 87.9 (2005), pp. 2003–2006. ISSN: 00036951. DOI: 10.1063/1.2034116.

[102] Jana Humpolíčková, Aleš Benda, and Jörg Enderlein. "Optical saturation as a versatile tool to enhance resolution in confocal microscopy". In: *Biophysical Journal* 97.9 (2009), pp. 2623–2629. ISSN: 15420086. DOI: 10.1016/j.bpj.2009.08.002.

[103] Yasunori Nawa et al. "Saturated excitation microscopy using differential excitation for efficient detection of nonlinear fluorescence signals". In: *APL Photonics* 3.8 (Aug. 2018), p. 080805. ISSN: 23780967. DOI: 10.1063/1.5039567. URL: http://aip.scitation.org/doi/10.1063/1.5039567.

[104] Kenta Temma et al. "Saturated-excitation image scanning microscopy". In: *Optics Express* 30.8 (2022), p. 13825. ISSN: 10944087. DOI: 10.1364/oe.455621.

[105] James P. Pawley, ed. *Handbook of Biological Confocal Microscopy*. 3rd ed. Springer, 2006. ISBN: 978-0387259215.

[106] Walter Hauswald et al. "Thermal illumination limits in 3D Raman microscopy: A comparison of different sample illumination strategies to obtain maximum imaging speed". In: *PLOS ONE* 14.8 (Aug. 2019). Ed. by Ireneusz Grulkowski, e0220824. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0220824. URL: http://dx.plos.org/10.1371/journal.pone.0220824.

G

PUBLICATIONS

G.1 ARTICLES IN PEER-REVIEWED JOURNALS

1. R. Förster, W. Müller, **R. Richter**, R. Heintzmann, "*Automated distinction of shearing and distortion artefacts in structured illumination microscopy*". In: Optics Express 26.16 (Aug. **2018**), p. 20680. issn: 1094-4087. doi: 10.1364/oe.26.020680

2. B. Diederich & **R. Lachmann**, S. Carlstedt, B. Marsikova, H. Wang, X. Uwurukundo, A. Mosig, R. Heintzmann, "*A versatile and customizable low-cost 3D-printed open standard for microscopic imaging*". In: Nature Communications 11.1 (Dec. **2020**), p. 5979. issn: 2041-1723. doi: 10.1038/s41467-020-19447-9

3. H. Wang, **R. Lachmann**, B. Marsikova, R. Heintzmann, B. Diederich, "*UCsim2: two-dimensionally structured illumination microscopy using UC2*". In: Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 380.2220 (Apr. **2022**), pp. 1–20. issn: 1364-503X. doi:10.1098/rsta.2020.0148.

4. K. Temma, R. Oketani, **R. Lachmann**, T. Kubo, N. Smith, R. Heintzmann, K. Fujita, "*Saturated-excitation image scanning microscopy*". In: Optics Express 30.8 (**2022**), p. 13825. issn: 10944087. doi: 10.1364/oe.455621.

G.2 CONFERENCE CONTRIBUTIONS

1. R. Meinel and **R. Richter**. "*Constructive proof of the kerr-newman blackhole uniqueness: derivation of the full solution from scratch*". In: Proceedings of the MG13 Meeting on General Relativity. 13. Stockholm University, **2012**, pp. 1828–1830. doi: 10.1142/9789814623995_0290.

2. B. Diederich, **R. Lachmann**, S. Carlstedt, X. Uwurukundo, H. Wang, A. Mosig, R. Heintzmann, "*UC2 – A 3D-printed General-Purpose Optical Toolbox for Microscopic Imaging*". In: Imaging and Applied Optics **2019** (COSI, IS, MATH,pcAOP). Vol. 2019. Washington, D.C.: OSA, 2019, ITh3B.5. isbn: 978-1-943580-63-7.doi: 10.1364/ISA.2019.ITh3B.5

3. F. Löffler, V. Wesp, S. Babalou, P. Kahn, **R. Lachmann**, B. Sateli, R. Witte, B. Konig-Ries, "*ScholarLensViz: A visualization framework for transparency in semantic user profiles*". In: CEUR Workshop Proceedings. Vol. 2721. **2020**, pp. 20–24.

4. B. Marsikova, **R. Lachmann**, E. Bingoel, R. Heintzmann, B. Diederich, "*Modular Open-Source Toolbox for Optics Education(openUC2)*". In: Education and Training in Optics & Photonics Conference **2021**. Washington, D.C.: OSA, 2021, Th1A.5. isbn: 978-1-943580-98-9. doi: 10.1364/ETOP.2021.

G.3 TALKS AND POSTERS

1. Poster | Focus on Microscopy 2017 (FOM), **03/17**, Bordeaux, France
2. Poster | Winter-Seminar, **01/15**, Klosters (Switzerland)
3. Poster | Winter-Seminar, **01/16**, Klosters (Switzerland)
4. Poster | Focus on Microscopy 2016 (FOM), **03/16**, Taipei, Taiwan
5. Poster | Winter-Seminar, **01/17**, Klosters (Switzerland)
6. Talk | Winter-Seminar, **01/19**,Klosters, Switzerland
7. Talk (invited) | OpenPlant Forum/Synthetic Biology, **07/19**,Cambridge, UK
8. Talk (invited) | Group Seminar, Bowman-Lab, Bath University, **07/19**,Bath, UK
9. Talk and Workshop (invited) | Janelia Research Campus (HHMI), **09/19**,Loundoun County, VA, USA
10. Talk and Workshop (invited) | UC2 Workshop for the NORMIC Association Oslo, **01/2020**,Oslo, Norway

G.4 CONDUCTED WORKSHOPS, CHALLENGES AND HACKATHONS

1. Challenge | *UC2-Edu goes AR*, Innovation Methods in Physics, **04/21**, Jena, Germany
2. Workshop | *Build your own microscopy using the open modular optical microscope toolbox UC2*, 5th NorMIC Workshop on Microscopy Image Processing, **01/20**, Oslo, Norway

3. Workshop | *Build your own microscopy using the open modular optical microscope toolbox UC2* | Junior Scientist Workshop on Biological Optical Microscopy, **09/19**, Loundoun County, VA, USA

4. Challenge | VDI Photonik Akademie, **07/19**, Jena, Germany

5. Hackathon | *Build your own light-sheet microscopy-based on the open modular optical microscope toolbox UC*, The Day of Light, **05/19**, Lichtwerkstatt, Jena Germany

G.5 AWARDS AND SCHOLARSHIPS

1. Scholarship | **2010**, Studienstiftung des deutschen Volkes

2. Scholarship Tokubetsu Choukousei | **2013**, Reitaku University, Japan

3. Scholarship Joint Japan Research | **2018**, Studienstiftung des deutschen Volkes and RIKEN Japan

4. Innovation-Price | **2018**, Leibniz Institute IPHT Jena

5. 1. price of Hot Stuff Award | **2019**, Digital Innovation Hub Photonics (DIHP), Jena

6. Special honor for employee commitment | **2019**, Leibniz Institute for Photonic Technology e.V., Jena, Germany

7. 8th Leibniz Founder's Award | **2021**, Leibniz Association, Berlin, Germany