Large-Scale Discovery of Experimental Designs in Super-Resolution Microscopy with XLUMINA

Carla Rodríguez [1] **Sören Arlt** [1] **Leonhard Möckl** [1] **Mario Krenn** [1]

Abstract

Driven by human ingenuity and creativity, the discovery of super-resolution techniques, which circumvent the classical diffraction limit of light, represent a leap in optical microscopy. However, the vast space encompassing all possible experimental configurations suggests that some powerful concepts and techniques might have not been discovered yet, and might never be with a human-driven direct design approach. Thus, AI-based exploration techniques could provide enormous benefit, by exploring this space in a fast, unbiased way. We introduce XLUMINA, an open-source computational framework developed using JAX, which offers enhanced computational speed enabled by its accelerated linear algebra compiler (XLA), just-in-time compilation, and its seamlessly integrated automatic vectorization, auto-differentiation capabilities and GPU compatibility. Remarkably, XLUMINA demonstrates a speed-up of 4 orders of magnitude compared to well-established numerical optimization methods. We showcase XLUMINA's potential by rediscovering two foundational techniques in advanced microscopy, together with new superior experimental layouts. Ultimately, XLUMINA identified a novel experimental blueprint featuring sub-diffraction imaging capabilities. This work constitutes and important step in AI-driven scientific discovery of new concepts in optics and advanced microscopy.

1. Introduction

The space of all possible experimental optical configurations is enormous. For example, if we consider experiments that consist of just 10 optical elements, chosen from 5 different components (such as lasers, lenses, phase shifters, beam splitters and cameras), we already get 10 million possible discrete arrangements. The experimental topology (i.e., how the elements are arranged) will further increase this number greatly. Finally, each of these optical components can have tunable parameters (such as lenses' focal lengths, laser power or splitting ratios of beam splitters) which lead to additional high-dimensional continuous parameter space for each of the previously mentioned discrete possibilities. This vast search space contains all experimental designs possible, including those with exceptional properties. So far, researchers have been exploring this space of possibilities guided by experience, intuition and creativity – and have uncovered countless exciting experimental configurations and technologies. But due to the complexity of this space, it might be that some powerful concepts and techniques have not been discovered so far, and might never be with a human-driven direct design approach. This is where AI-based exploration techniques could provide enormous benefit, by exploring the space in a fast, unbiased way (Wang et al., 2023; Krenn et al., 2022).

Optical microscopes in today's sense were invented 300 years ago by Antonj van Leeuwenhoek (Wollman et al., 2015). Since then, few techniques used in the sciences have seen a similarly rapid development and impact on diverse fields, ranging from material sciences all the way to medicine (Reigoto et al., 2021; Weisenburger & Sandoghdar, 2015; Bullen, 2008; Antony et al., 2013). Arguably, optical microscopy is currently most widely used in biological sciences, where precise labeling of imaging targets enables fluorescence microscopy with exquisite sensitivity and specificity (Grimm & Lavis, 2022; M. & Palmer, 2014). In the past two decades, several breakthroughs have broadened the scope of optical microscopy in this area even further. Among them, through the ingenuity and creativity of human researchers, the discovery of super-resolution (SR) methods, which circumvent the classical diffraction limit of light, stand out in particular. Examples for versatile and powerful SR techniques are STED (Hell & Wichmann, 1994), PALM/F-PALM (Betzig et al., 2006; Hess et al., 2006), (d)STORM (Rust et al., 2006; van de Linde et al., 2011), SIM (Gustafsson, 2005a), and MINFLUX (Balzarotti

[1]Max Planck Institute for the Science of Light, Erlangen, Germany. Correspondence to: Carla Rodríguez <carla.rodriguez@mpl.mpg.de>.

AI for Science Workshop of the 41st International Conference on Machine Learning, Vienna, Austria. PMLR 235, 2024. Copyright 2024 by the author(s).

et al., 2017), with considerable impact in biology (Möckl et al., 2019; Xu et al., 2013; Yildiz et al., 2003), chemistry (Zhang et al., 2015) and material sciences (Müller et al., 2019) for example. Crucially, the motivation of our work goes far beyond small-scale optimization of already known optical techniques. Rather, this work sets out to discover novel, experimentally viable concepts for advanced optical microscopy that are at-present entirely untapped.

We introduce XLUMINA, an efficient open-source framework developed using JAX (Bradbury et al., 2018), for the ultimate goal of discovering new optical design principles. XLUMINA offers enhanced computational speed enabled by its accelerated linear algebra compiler (XLA), just-in-time (jit) compilation, and its seamlessly integrated automatic vectorization or batching, auto-differentiation capabilities (Baydin et al., 2018) and GPU compatibility. We leverage its scope with a specific focus on the area of SR microscopy, which is a set of techniques that has revolutionized biological and biomedical research over the past decade, highlighted by the 2014 Chemistry Nobel Prize (Möckl et al., 2014). The software's workflow is depicted in Fig. 1a. Fundamentally, the simulator is the heart of digital discovery efforts. It translates an experimental design (one point in the vast space of possible designs) to a physical output. The physical output, such as a detector or camera output, can then be used in an objective function to describe the desired design goal. The simulator can either be called directly by gradient-based optimization techniques, or it can be used for generating the training data for deep-learning-based surrogate models. A simulator that can be used for automated design and discovery of new experimental strategies must be (1) fast, (2) reliable, and (3) general. XLUMINA's optical simulator fulfills precisely the aforementioned requirements for advanced microscopy.

The work is structured as follows. Upon reviewing previous work, we describe XLUMINA and highlight its efficiency and computational speed advantage over conventional approaches. We define a quasi-universal computational *ansatz* to translate the hybrid discrete-continuous optimization problem, arising from configuring both the optical topology and the optical parameters, into a purely continuous optimization framework which can be solved with efficient gradient-based methods. Then, we demonstrate the applicability of our approach by rediscovering, together with new superior experimental topologies, two foundational techniques in SR microscopy following pure AI-exploratory strategies within our fully continuous framework: a beam-shaping technique as employed in STED (stimulated emission depletion) microscopy (Hell & Wichmann, 1994) and the SR technique exploiting optical vortices (Dorn et al., 2003). Ultimately, we showcase XLUMINA's capability for genuine discovery identifying a novel solution that integrates the underlying physical principles present in the two aforementioned SR techniques into a single experimental blueprint, the performance of which exceeds the capabilities of each individual setup. We conclude with final remarks and future perspectives.

1.1. Previous work

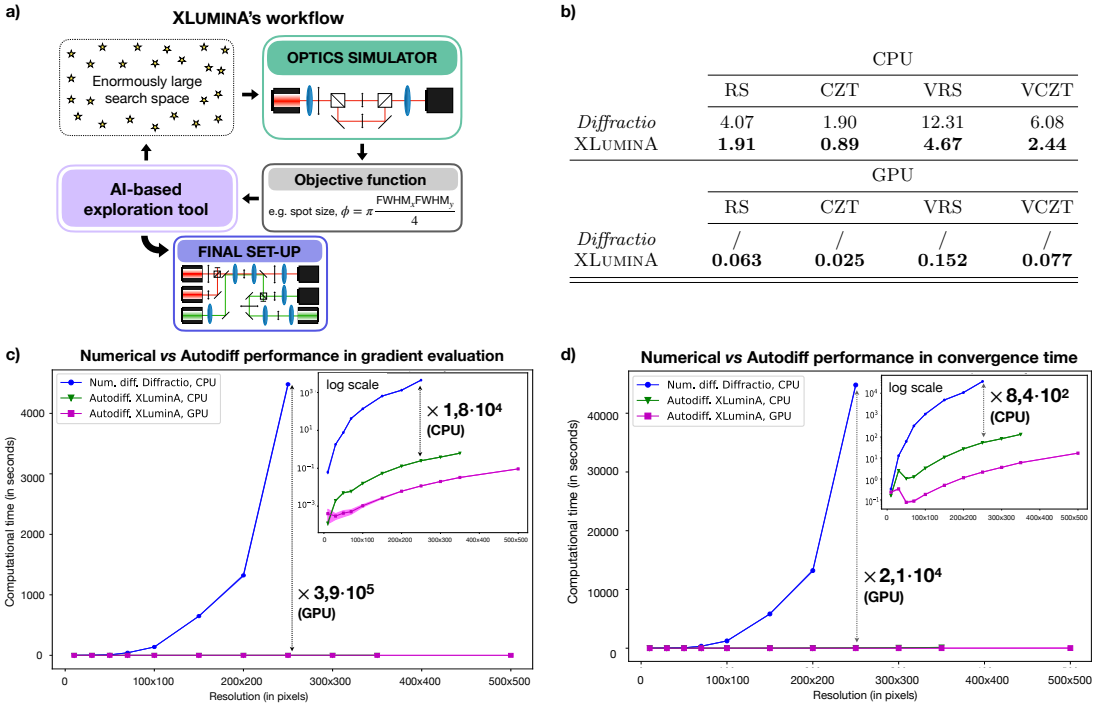
Our approach is radically different from previous strategies that employ AI for data-driven design of single optical elements (Herath et al., 2023; Yanny et al., 2020) or data analysis in microscopy, e.g. denoising, contrast enhancement or point-spread-function (PSF) engineering (Nehme et al., 2020). While these techniques are influential, they are not meant to change the principle of the experimental approach or the optical layout itself. In contrast, XLUMINA is equipped with tools to simulate, optimize and automatically design new optical setups and concepts from scratch.

Several open-source software tools facilitate classical optics phenomena simulations. Some examples are *Diffractio* for light diffraction and interference simulations (Brea, 2019), *Finesse* for simulating gravitational wave detectors (Freise et al., 2013), which do not support auto-differentiation nor GPU compatibility; and *POPPY*, developed as a part of the simulation package of the James Webb Telescope (Perrin et al., 2012), with GPU compatibility but lacking autodiff capabilities. There are also specialized resources like those focusing on the design of Laguerre-Gaussian mode sorters utilizing multi-plane light conversion (MPLC) methods (Fontaine et al., 2019; Labroille et al., 2014), which also do not support GPU computations and autodiff. While these software solutions offer optics simulation capabilities, XLUMINA uniquely integrates simulation with AI-driven automated design powered with JAX's autodiff, just-in-time compilation and automatic GPU compatibility.

2. Software workflow and performance

XLUMINA (Link to GitHub repo) allows for the simulation of classical optics hardware configurations and enables the optimization and automated discovery of new setup designs. The software is developed using JAX (Bradbury et al., 2018), which provides an advantage of enhanced computational speed (enabled by accelerated linear algebra compiler, XLA, with just-in-time compilation, jit) while seamlessly integrating the auto-differentiation framework (Baydin et al., 2018) and automatic GPU compatibility. It is important to remark that our approach is not restricted to run on CPU (as NumPy-based softwares do): due to JAX-integrated functionalities, by default runs on GPU if available, otherwise automatically falls back to CPU.

The ultimate goal is to discover new concepts and experimental blueprints in optics. Importantly, the most computational expense of an optimization loop comes from running



	CPU			
	RS	CZT	VRS	VCZT
Diffractio	4.07	1.90	12.31	6.08
XLUMINA	**1.91**	**0.89**	**4.67**	**2.44**
	GPU			
	RS	CZT	VRS	VCZT
Diffractio	/	/	/	/
XLUMINA	**0.063**	**0.025**	**0.152**	**0.077**

Figure 1. Overview and performance of XLUMINA. (a) Software's workflow, demonstrating the integrated feedback between the AI discovery tool and the optics simulator. Stars depict experimental blueprints with exceptional and useful properties. We start by feeding the system an initial random set of optical parameters, which shape the hardware design on a virtual optical table. The performance of the virtual experiment is computed by the simulator, which leads to detected light (e.g., captured images at the camera). From those simulated outputs, the objective function (e.g., the spot size), is computed. To improve the metric of the cost function, the optimizer adjusts the optical parameters in the initial virtual setup and the cycle is repeated. The whole process is a back-and-forth between the simulator and the optimizer, refining the setup until a convergence is observed. (b) Average execution time (in seconds) over 100 runs, within a resolution of 2048×2048 pixels, for scalar and vectorial field propagation using Rayleigh-Sommerfeld (RS, VRS) and Chirped z-transform (CZT, VCZT) algorithms in *Diffractio* and XLUMINA. Times for XLUMINA correspond to runs with pre-compiled jitted functions. Our methods demonstrate enhanced computational speeds for simulating light diffraction and propagation: a factor of $\times 2$ for RS and CZT and about $\times 2.5$ for VRS and VCZT using the CPU. With GPU utilization, the speed increases up to two orders of magnitude with factors of $\times 64$ for RS, $\times 76$ for CZT, $\times 80$ for VRS and $\times 78$ for VCZT. (c) Average time (in seconds) over 5 runs for a single gradient evaluation using numerical differentiation with *Diffractio*'s optical simulator (blue dots) and autodiff methods (green triangles for CPU and magenta squares for GPU) with XLUMINA's optical simulator for different resolutions. The use of XLUMINA with autodiff methods improves the gradient evaluation time by a factor of $\times 3.9 \cdot 10^5$ in the GPU and a factor of $\times 1.8 \cdot 10^4$ on the CPU for resolutions of 250×250 pixels. The superior efficiency of autodiff over traditional numerical methods allows for highly efficient optimizations particularly employing the large high resolutions we use (up to 2048×2048 pixels). (d) Average time (in seconds) over 5 runs for convergence time, using numerical differentiation with *Diffractio*'s optical simulator (blue dots) and autodiff methods (green triangles for CPU and magenta squares for GPU) with XLUMINA's optical simulator for different resolutions. Autodiff methods on XLUMINA improves the convergence time with respect to numerical methods by a factor of $\times 2.1 \cdot 10^4$ in the GPU and a factor of $\times 8.4 \cdot 10^2$ in the CPU for a resolution of 250×250 pixels. Shaded regions correspond to standard deviation values. The numerical and autodiff methods are computed using BFGS and Adam optimizers, respectively. Further comparison across different optimizers is presented in the Appendix section A. All the experiments were run on an Intel CPU Xeon Gold 6130 and Nvidia GPU Quadro RTX 6000.

individual optical simulations in each iteration. Thus, it is essential to reduce the computation time by maximizing the speed of optical simulation functions. XLUMINA is equipped with an optics simulator which contains a diverse set of optical manipulation, interaction and measurement technologies. Some specific optical propagation implementations of XLUMINA are inspired by the optics framework *Diffractio* (Brea, 2019). *Diffractio* is a high-quality, open-source NumPy-based Python module for optics simulation with an active developer community, and is employed in numerous studies in optics and physics in general. We have rewritten and optimized these optical propagation implementations leveraging JAX's jit functionality, which allows for highly efficient code execution, although it imposes some restrictions such as specifying all data structures' dimensions and ensuring their immutability at compile time. On

top of that, we developed completely new functions which significantly expand the software capabilities, such as high-resolution propagation methods, and numerous new optical devices which made the current study possible. Further details on the optics simulator can be found in the Appendix section A. We evaluate the performance of our optimized functions against their counterparts in *Diffractio*. The acquired run-times are shown in Fig. 1b. Clearly, our methods significantly enhance computational speeds for simulating light diffraction and propagation. For instance, we observe a speedup of a factor of $\times 2$ for RS (Rayleigh-Sommerfeld, a general Fast Fourier Transform-based light propagation algorithm) and CZT (Chirped z-transform, a speed-up version of RS) and about $\times 2.5$ for VRS and VCZT (the vectorized versions of RS and CZT, respectively) using the CPU. With GPU utilization, the speedup factors are of $\times 64$ for RS, $\times 76$ for CZT, $\times 80$ for VRS and $\times 78$ for VCZT.

To include the automated discovery feature, XLUMINA's optical simulator and optimizer are tied together by the loss function, as depicted in Fig. 1a. The automated discovery tool is designed to explore the vast parameter space encompassing all possible optical designs. When it comes to the nature of the optimizer, it can be either direct (gradient-based) or deep learning-based (surrogate models or deep generative models, e.g., variational autoencoders (Flam-Shepherd et al., 2022)). In this work, we adopt a gradient-based strategy, where the experimental setup's parameters are adjusted iteratively in the steepest descent direction. We first evaluate the time it takes for numerical and analytical (auto-differentiation) methods to compute one gradient evaluation and their convergence times over different resolutions and devices. For this purpose, we use two gradient-descent techniques: the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm (Nocedal & Wright), which numerically computes the gradients and higher-order derivative approximations and the Adaptive moment estimation (Adam) (Kingma & Ba, 2017), an instance of the stochastic-gradient-descent (SGD) method. While BFGS is part of the open-source SciPy Python library (Virtanen et al., 2020) and operates on the CPU, Adam is integrated within the JAX framework and runs in both CPU and GPU. For this last, we take advantage of JAX's built-in autodiff framework and compute analytically the gradients of the loss function. Combined with the jit (just-in-time) compilation functionality, this approach enables the optimizer to efficiently construct an internal gradient function, considerably reducing the computational time per iteration. The acquired results are depicted in Figs. 1a and 1b. The detailed description of both evaluations is provided in the Appendix section A. Clearly, autodiff consistently outperforms numerical methods on the gradient evaluation time by up to 4 orders of magnitude on CPU and 5 orders on GPU. In convergence time, autodiff demonstrates superior efficiency up to almost 3 orders of magnitude on CPU and 4 orders on GPU. Given that certain optical elements, such as phase masks, may operate at resolutions as high as 2048×2048 pixels, the resulting search space can easily expand to around 8.4 million parameters. This makes the use of autodiff within GPU-accelerated frameworks more appropriate for efficient experimentation. Overall, the computational performance of XLUMINA highlights its suitability for running complex simulations and optimizations with a high level of efficiency.

3. Large-scale discovery framework

We aim to use XLUMINA to discover new microscopy concepts. In essence, discovering new experimental configurations entails an hybrid discrete-continuous search problem. The discrete aspect originates from configuring the optical network topology, whereas the continuous part is linked to the settings of optical elements, such as laser power and beam splitter reflectivity. Discrete-continuous optimization is extremely difficult computationally, therefore we invent a way to translate this hybrid discrete-continuous optimization problem into a purely continuous optimization problem which can be solved with efficient gradient-based methods. We design the quasi-universal computational *ansatz* illustrated in Fig. 2, which is designed in a way that setting different (continuous) parameters leads to different optical setup topologies.

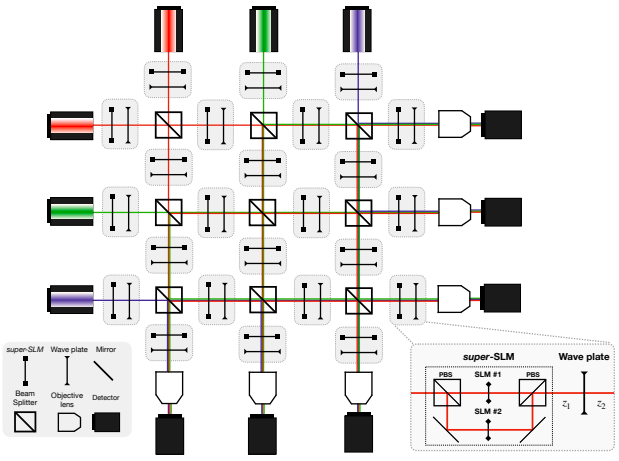


Figure 2. General virtual optical setup for large-scale discovery schemes. Gray boxes represent fundamental building units, each containing a *super-SLM* and a wave plate positioned a distance z_1 apart. These units are inter-connected through free propagation distances z_2, and beam splitters. The *super*-SLM is a hardware-box-type which consists of two spatial light modulators (SLMs), each one independently imprinting a phase pattern on the horizontal and vertical polarization components of the field. The setup's complexity and size can be arbitrarily extended by incorporating additional connections, building units, light sources, detectors, etc.

Remarkably, for a very discrete approach of available parameters, the number of possible discrete arrangements within this general framework scales up to $\sim 10^{20}$. Details on this derivation can be found in the Appendix section B.1.

Now, the task of XLUMINA is to automatically discover new superior topologies together with their parameter setting, using purely continuous optimization. To achieve this, we initialize the setups with a large and complex optical topology, inspired by other fields that start with highly expressive initial circuits (Sim et al., 2019; Krenn et al., 2021). From here, XLUMINA should be able to extract much more complex solutions which humans might not have thought about yet (Krenn et al., 2022).

3.1. Loss function

The loss function, $\mathcal{L}$, is calculated as the inverse of the density of the total detected intensity over a certain threshold, I_ε. Thus, minimizing $\mathcal{L}$ aims to maximize the generation of small, high intensity beams. In particular,

$$\mathcal{L} = \frac{1}{\text{Density}} = \frac{\text{Area}}{I_\varepsilon} \tag{1}$$

where I_ε is the sum of pixel intensity values greater than the threshold value $\varepsilon \cdot i_{\text{max}}$, where $0 \leq \varepsilon \leq 1$ and i_{max} corresponds to the maximum detected intensity. The Area corresponds to the total number of camera pixels fulfilling the same condition. The loss function $\mathcal{L}$ is common to all the optical setups henceforth described. Importantly, light gets detected across various devices. Thus, we compute the loss function at each detector and the parameter update is driven by the device demonstrating the minimum loss value. This selection is performed in a fully differentiable manner. Details on the derivation of the loss function and camera selection are provided in the Appendix section B.2.

4. Results

In this section, we showcase the virtual optical designs generated by XLUMINA. As benchmarks, we aim to rediscover the concepts of two foundational experiments: the beam shaping as employed in STED microscopy (Hell & Wichmann, 1994), and the super-resolution technique by Dorn, Quabis and Leuchs in Ref. (Dorn et al., 2003). Finally, we demonstrate the discovery of a new experimental blueprint. For all the experiments hereby conducted we set-up a discovery scheme where no training data is involved. The showcased solutions are the result from running multiple optimizations.

4.1. Rediscovery through exploration

In this section we target XLUMINA to rediscover the aforementioned SR techniques. The goal here is to discover both the optical topology and the phase patterns to imprint onto the light beams using the available optimizable optical parameters (i.e., SLMs, distances, beam splitter ratios and wave plate's angles). The detailed description of the optimization processes hereby conducted are provided in the Appendix section C.

STED microscopy (Hell & Wichmann, 1994; Hofmann et al., 2005) is one of the first discovered techniques that circumvent the classical diffraction limit of light. The key idea of this technique is the use of two diffraction-limited laser beams, one probe to activate (excite) the light emitters of the sample and one, doughnut-shaped beam to deactivate its excitation in a controlled way (depletion). Thus, the ultimately detected light is that of the emitters laying in the central region of the doughnut-shaped beam. This effectively reduces the area of normal fluorescence, which leads to super-resolution imaging. To simulate one of the fundamental concepts of STED without having to rely on time-dependent processes, such as the energy level relaxation times of the excited emitters, we perform a nonlinear modulation of the focused light based on the Beer-Lambert law (Mayerhöfer et al., 2020), commonly used to describe the optical attenuation in light-matter interaction. The details of our model are provided in the Appendix section B.3.

We initialize XLUMINA with the virtual optical table in Supplementary Fig. 2. The loss function corresponds to equation (1) considering the radial component of the effective beam resulting from the STED process, $|E_x|^2 + |E_y|^2$. The discovered topology and phase patterns for STED microscopy are depicted in Figs. 3a and 3b, respectively. As demonstrated for the original STED microscopy, the system imprints a phase singularity onto the depletion beam to produce a doughnut shape. In this case, however, it also modulates the excitation beam. The radial intensity profiles of the discovered solution and the reference experiment are depicted in Fig. 3c.

The generation of an ultra-sharp focus is a feature that breaks the diffraction limit in the longitudinal direction as demonstrated by Dorn, Quabis and Leuchs in Ref. (Dorn et al., 2003). This super-resolution is achieved when a radially polarized beam is tightly focused (Quabis et al., 2000; Quinteiro et al., 2017). We initialize XLUMINA with the virtual optical table in Supplementary Fig. 2. The loss function corresponds to equation (1) considering the measured intensity as the field's longitudinal component, $|E_z|^2$. The discovered topology and phase patterns are depicted in Figs. 3d and 3e, respectively. These produce a $LG_{2,1}$ Laguerre-Gaussian mode (Rubinsztein-Dunlop et al., 2016), which demonstrates an intensity pattern of concentric rings with a phase singularity in its center. Surprisingly, XLUMINA found an alternative way to imprint a phase singularity onto

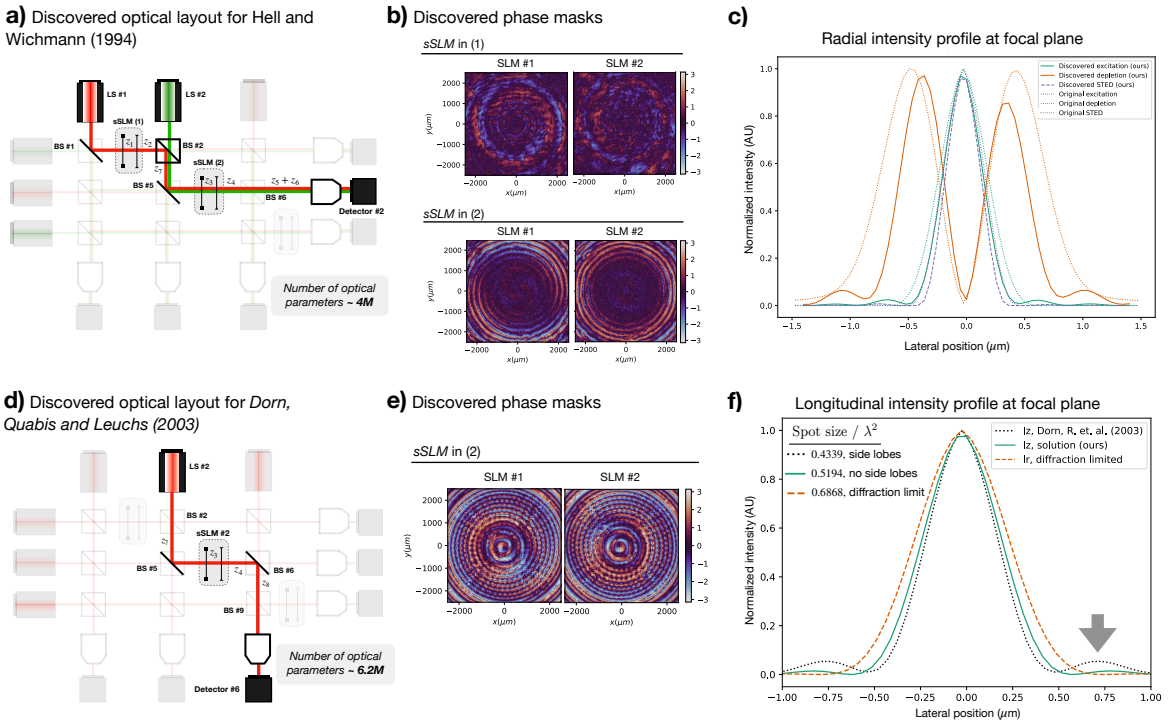


Figure 3. Rediscovery of optical solutions within the large-scale framework. The parameter space is defined by 3 *super*-SLMs (i.e., 6 SLMs), 9 beam splitter ratios, 8 distances and 3 wave plates (with variable phase retardance η and orientation angle θ). The values of the identified optical parameters are specified in the Appendix section C. (a) Discovered optical topology for STED microscopy (Hell and Wichmann, 1994). The parameter space corresponds to $\sim$ 4 million optical parameters. (b) Discovered phase patterns for sSLM in (1) and sSLM in (2). (c) Radial intensity profile in horizontal beam section: excitation (green), depletion (orange), and super-resolution effective STED beam (dashed blue line). The data corresponding to the original STED experiment - i.e., computed using a spiral phase mask - are indicated with dotted lines. Lateral position indicates lateral distance from the optical axis. (d) Discovered virtual optical setup topology for Dorn, Quabis and Leuchs (2003). The parameter space corresponds to $\sim$ 6.2 million optical parameters. (c) Discovered phase patterns for sSLM in (2). (d) Normalized longitudinal intensity profile for Dorn, Quabis, and Leuchs (2003) and the identified solution (black dotted, and green lines, respectively) and radial intensity profile of the diffraction-limited linearly polarized beam (orange dotted line). Lateral position indicates lateral distance from the optical axis. The spot size is computed as $\phi = (\pi/4)\text{FWHM}_x\text{FWHM}_y$, where FWHM denotes for Full Width Half Maximum. The discovered approach breaks the diffraction limit with a spot size close to the reference. Remarkably, it does not feature side lobes (indicated with a gray arrow), which can limit practical imaging techniques.

the beam and produce pronounced longitudinal components on the focal plane. The longitudinal intensity profiles of the discovered solution and the reference experiment are depicted in Fig. 3f. Remarkably, the identified solution demonstrates a spot size close to the reference and does not feature side lobes, which can limit practical imaging techniques.

Importantly, we are not restricted to the use of 3×3 optical grids. We further test our framework by conducting a pure topological optimization for Dorn, Quabis and Leuchs (2003), this time within a 6×6 optical system containing four SLMs displaying pre-defined phase masks. The goal is to discover the optical topology using the available optimizable optical parameters: beam splitter ratios, distances and wave plate's angles. The initial virtual setup and the displayed phase masks are depicted in Supplementary Figs. 3a and 3b, respectively. The loss function corresponds to equation (1) considering the measured intensity as the field's longitudinal component, $|E_z|^2$. The discovered topology is depicted in Fig. 4a. The longitudinal intensity profiles of the discovered solution and the reference experiment are depicted in Fig. 4b. Clearly, the system successfully rediscovered the reference experiment conducting a pure topological discovery within a large-scale framework.

In all cases we successfully demonstrate how XLUMINA can explore different experimental topologies and optical solutions in a fully continuous manner: by adjusting the optical parameters, (e.g., beam splitter ratios), the optimizer can "*turn off*" the optical paths.

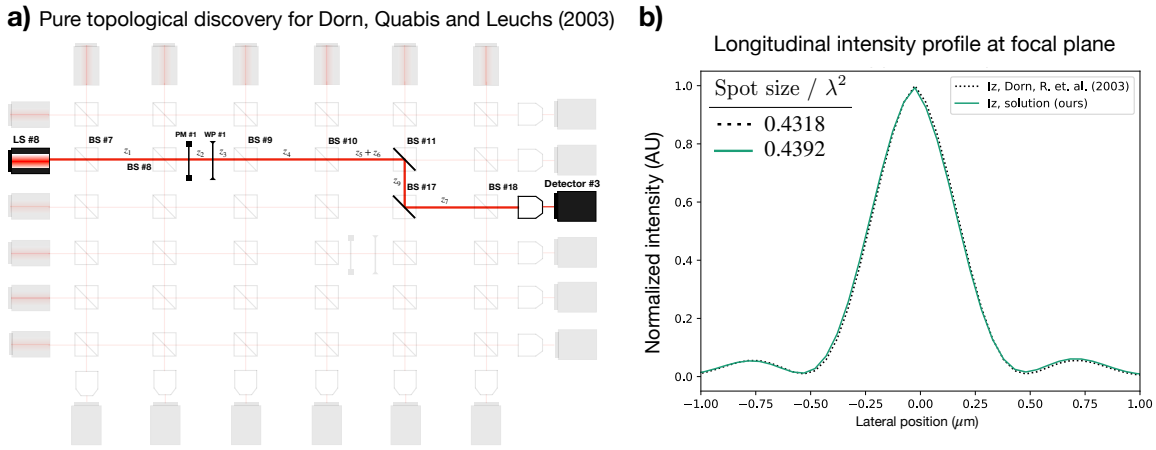


Figure 4. Pure topological discovery within a 6×6 large-scale optical setup. The optical elements corresponding to phase masks (PM#1 and PM#2) remain fixed during the optimization. The values of the identified optical parameters are specified in the Appendix section C. (a) Discovered topology for Dorn, Quabis and Leuchs (2003). The phase mask #1 corresponds to the polarization converter originally demonstrated in Dorn, Quabis, and Leuchs (2003). The minimum value of the loss is demonstrated in detector #3. (b) Normalized longitudinal intensity profile for the reference and the identified solution (black dotted, and green lines, respectively). Lateral position indicates lateral distance from the optical axis. The spot size is computed as $\phi = (\pi/4)\text{FWHM}_x\text{FWHM}_y$, where FWHM denotes for Full Width Half Maximum. The system successfully rediscovered the reference experiment.

4.2. Discovery of a new experimental blueprint

Finally, we demonstrate the capability of XLUMINA for genuine discovery. We initialize the system in the virtual setup depicted in Supplementary Fig. 2. The details of the optimization are provided in the Appendix section D. The loss function corresponds to equation (1) considering the total intensity of the effective beam resulting from the STED process, $|E_x|^2 + |E_y|^2 + |E_z|^2$. The discovered topology and identified phase patterns are depicted in Figs. 5a and 5b, respectively. The detected intensity topologies reveal the system generates a doughnut-shaped and a Gaussian-like beams. We compute the vertical cross-section of the focused intensity patterns for both beams and the resulting effective beam (green, orange and dotted blue lines in Fig. 5c, respectively). The horizontal cross-section exhibits analogous features. We further compare the effective beam intensity with the simulated STED reference (Hell & Wichmann, 1994) and the the discovered Gaussian-like beam with the simulated sharp focus reference (Dorn et al., 2003). The obtained results are showcased in Fig. 5d. Strikingly, the discovered solution exploits the underlying physical concepts of two aforementioned optical systems. In one hand, it generates a doughnut-shaped "depletion" beam as demonstrated in Ref. (Hell & Wichmann, 1994). On the other hand, it generates a Gaussian-like "excitation" signal with a sharper focus, achieving smaller effective intensity spots resulting from the STED process. The discovered solution showcases an effective beam profile which is sharper than the simulated STED reference. This occurs due to the enhanced sharpening of the longitudinal component of the excitation beam, which demonstrates similar profile as the simulated sharp focus reference (Dorn et al., 2003). To the best of our knowledge, this technique has never been discussed in the scientific literature before. Regardless of its physical realizability, this solution demonstrates the ability of XLUMINA to uncover interesting solutions within highly complex systems.

5. Discussion and outlook

In this work we present XLUMINA, a highly efficient computational framework with seamlessly integrated auto-differentiation capabilities, just-in-time compilation, automatic vectorization and GPU compatibility, for the discovery of novel optical setups in super-resolution microscopy. We demonstrate the high-performance and efficiency of XLUMINA with a computational speed-up of $\times\, 2.1 \cdot 10^4$ on GPU, and $\times\, 8.4 \cdot 10^2$ on CPU, compared to standard numerical optimization methods. We further prove the accuracy and reliability of our methods by successfully rediscovering two foundational super-resolution techniques within a large-scale discovery framework. More significantly, XLUMINA identified a novel experimental blueprint that breaks the diffraction limit by integrating the physical principles of two well-known super-resolution techniques.

Having laid the groundwork with XLUMINA for an efficient, versatile optics simulator, many other microscopy and imaging techniques follow as a natural extension. For example, by enabling time information using algorithms iterating over the simulator, light scattering of probe samples can

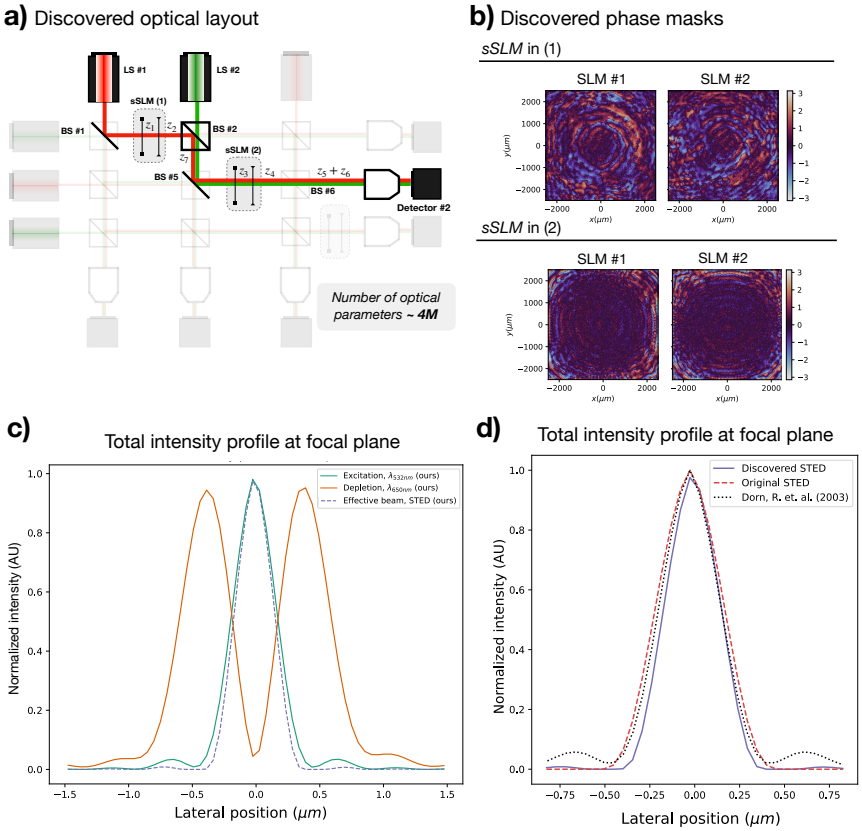


Figure 5. Discovery of a new experimental blueprint within a highly parameterized optical setup. The parameter space (of ~ 4 million optical parameters) is defined by 3 *super*-SLMs (i.e., 6 SLMs), 9 beam splitter ratios, 8 distances and 3 wave plates (with variable phase retardance η and orientation angle θ). The values of the identified optical parameters are specified in the Appendix section D. (a) Discovered optical topology. The minimum value of the loss is demonstrated in detector #2. (b) Discovered phase masks corresponding to the *super*-SLMs (*sSLM*) in (1) and (2). (c) Total intensity ($|E_x|^2 + |E_y|^2 + |E_z|^2$) horizontal cross-section of the detected light beams of 650 nm (orange), 532 nm (green), and effective beam emulating stimulated emission (dashed blue). (d) Horizontal cross-section of the normalized total intensity of the effective beam from the discovered solution (blue), the simulated STED reference (dashed red), and the simulated reference (dotted black) using 532 nm wavelength. The discovered solution outperforms both simulated references for STED microscopy and the sharp focus from Dorn, Quabis and Leuchs (2003).

be easily implemented, enabling systems such as iSCAT (Taylor & Sandoghdar, 2019), structured illumination microscopy (Gustafsson, 2005b), and localization microscopy (Lelek et al., 2021). Also, inspired by the work in (Nehme et al., 2020), we could leverage similar approaches to optimize hardware-software discovery within our extensive framework. Additionally, one could formulate interesting experiments (e.g., in low light conditions) where the use of noise or loss sources (e.g., absorption, vibration, etc.) become relevant.

Additionally, XLUMINA provides already the basis for an expansion to complex quantum optics microscopy techniques (Taylor et al., 2013) or other quantum imaging techniques (Moreau et al., 2019), as a quantum of light (i.e., a photon) is nothing else than an excitation of the modes of the electromagnetic field. Looking further into the future, one can expect that matter-wave beams (governed by Schrödinger's equation, which is closely related to the paraxial wave equation, a special case of the electromagnetic field) can be simulated in the same framework. This might allow for the AI-based design of microscopy techniques which could harness entirely new ideas combining light and complex matter wave beams such as electron-beams (Chirita Mihaila et al., 2022; Kalinin et al., 2022; 2023) or coherent beams of high-mass particles (Kiałka et al., 2022). Ultimately, bringing so far unexplored concepts from diverse areas of physics to microscopy applications is at the heart of AI-driven discovery in this area, and we hope that this work constitutes a first step in this direction.

Impact Statement

This paper presents work whose goal is to advance the field of AI-driven discovery of new optics solutions with a specific focus on super-resolution microscopy. There are many potential societal consequences of our work, none which we feel must be specifically highlighted here.

References

Antony, P., Trefois, C., Stojanovic, A., Baumuratov, A., and Kozak, K. Light microscopy applications in systems biology: opportunities and challenges. *Cell Communication and Signaling*, 11(1), 2013. URL https://doi.org/10.1186/1478-811X-11-24.

Balzarotti, F., Eilers, Y., Gwosch, K. C., Gynnå, A. H., Westphal, V., Stefani, F. D., Elf, J., and Hell, S. W. Nanometer resolution imaging and tracking of fluorescent molecules with minimal photon fluxes. *Science*, 355 (6325):606–612, 2017. URL https://doi.org/10.1126/science.aak9913.

Baydin, A. G., Pearlmutter, B. A., Radul, A. A., and Siskind, J. M. Automatic differentiation in machine learning: a survey. *Journal of Machine Learning Research*, 18(153):1–43, 2018. URL http://jmlr.org/papers/v18/17-468.html.

Betzig, E., Patterson, G. H., Sougrat, R., Lindwasser, O. W., Olenych, S., Bonifacino, J. S., Davidson, M. W., Lippincott-Schwartz, J., and Hess, H. F. Imaging intracellular fluorescent proteins at nanometer resolution. *Science*, 313(5793):1642–1645, 2006. URL https://doi.org/10.1126/science.1127344.

Bradbury, J., Frostig, R., Hawkins, P., Johnson, M. J., Leary, C., Maclaurin, D., Necula, G., Paszke, A., VanderPlas, J., Wanderman-Milne, S., and Zhang, Q. JAX: composable transformations of Python+NumPy programs, 2018. URL http://github.com/google/jax.

Brea, L. M. S. Diffractio, python module for diffraction and interference optics, 2019. URL https://pypi.org/project/diffractio/.

Bullen, A. Microscopic imaging techniques for drug discovery. *Nature Reviews Drug Discovery*, 7(1):54–67, 2008. URL https://doi.org/10.1038/nrd2446.

Chirita Mihaila, M. C., Weber, P., Schneller, M., Grandits, L., Nimmrichter, S., and Juffmann, T. Transverse electron-beam shaping with light. *Phys. Rev. X*, 12:031043, Sep 2022. doi: 10.1103/PhysRevX.12.031043. URL https://link.aps.org/doi/10.1103/PhysRevX.12.031043.

Dorn, R., Quabis, S., and Leuchs, G. Sharper focus for a radially polarized light beam. *Physical Review Letters*, 91:233901, 2003. doi: 10.1103/PhysRevLett.91.233901. URL https://link.aps.org/doi/10.1103/PhysRevLett.91.233901.

Flam-Shepherd, D., Wu, T. C., Gu, X., Cervera-Lierta, A., Krenn, M., and Aspuru-Guzik, A. Learning interpretable representations of entanglement in quantum optics experiments using deep generative models. *Nature Machine Intelligence*, 4(6):544–554, 2022. URL https://doi.org/10.1038%2Fs42256-022-00493-5.

Fontaine, N. K., Ryf, R., Chen, H., Neilson, D. T., Kim, K., and Carpenter, J. Laguerre-gaussian mode sorter. *Nature Communications*, 10(1), 2019. URL https://doi.org/10.1038%2Fs41467-019-09840-4.

Freise, A., Brown, D., and Bond, C. Finesse, frequency domain interferometer simulation software, 2013.

Grimm, J. B. and Lavis, L. D. Caveat fluorophore: an insiders' guide to small-molecule fluorescent labels. *Nature Methods*, 19(2), 2022. URL https://doi.org/10.1038/s41592-021-01338-6.

Gustafsson, M. G. L. Nonlinear structured-illumination microscopy: Wide-field fluorescence imaging with theoretically unlimited resolution. *Proceedings of the National Academy of Sciences*, 102(37):13081–13086, 2005a. URL https://doi.org/10.1073/pnas.0406877102.

Gustafsson, M. G. L. Nonlinear structured-illumination microscopy: Wide-field fluorescence imaging with theoretically unlimited resolution. *Proceedings of the National Academy of Sciences*, 102(37):13081–13086, 2005b. URL https://doi.org/10.1073/pnas.0406877102.

Hell, S. W. and Wichmann, J. Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy. *Optics Letters*, 19(11):780–782, 1994. doi: 10.1364/OL.19.000780. URL https://opg.optica.org/ol/abstract.cfm?URI=ol-19-11-780.

Herath, K., Haputhanthri, U., Hettiarachchi, R., Kariyawasam, H., Ahmad, R. N., Ahmad, A., Ahluwalia, B. S., Edussooriya, C. U. S., and Wadduwage, D. N. Differentiable microscopy designs an all optical phase retrieval microscope, 2023.

Hess, S. T., Girirajan, T. P., and Mason, M. D. Ultra-high resolution imaging by fluorescence photoactivation localization microscopy. *Biophysical*

Journal, 91(11):4258–4272, 2006. ISSN 0006-3495. URL https://doi.org/https://doi.org/10.1529/biophysj.106.091116.

Hofmann, M., Eggeling, C., Jakobs, S., and Hell, S. W. Breaking the diffraction barrier in fluorescence microscopy at low light intensities by using reversibly photoswitchable proteins. *Proceedings of the National Academy of Sciences*, 102(49):17565–17569, 2005. doi: 10.1073/pnas.0506010102. URL https://www.pnas.org/doi/abs/10.1073/pnas.0506010102.

Hu, Y., Wang, Z., Wang, X., et al. Efficient full-path optical calculation of scalar and vector diffraction using the bluestein method. *Light: Science & Applications*, 9:119, 2020. URL https://doi.org/10.1038/s41377-020-00362-z.

Kalinin, S. V., Ophus, C., Voyles, P. M., Erni, R., Kepaptsoglou, D., Grillo, V., Lupini, A. R., Oxley, M. P., Schwenker, E., Chan, M. K., et al. Machine learning in scanning transmission electron microscopy. *Nature Reviews Methods Primers*, 2(1):11, 2022. doi: 10.1038/s43586-022-00095-w. URL https://doi.org/10.1038/s43586-022-00095-w.

Kalinin, S. V., Liu, Y., Biswas, A., Duscher, G., Pratiush, U., Roccapriore, K., Ziatdinov, M., and Vasudevan, R. Human-in-the-loop: The future of machine learning in automated electron microscopy. *arXiv:2310.05018*, 2023. URL https://arxiv.org/abs/2310.05018.

Kiałka, F., Fein, Y. Y., Pedalino, S., Gerlich, S., and Arndt, M. A roadmap for universal high-mass matter-wave interferometry. *AVS Quantum Science*, 4(2), 2022. URL https://doi.org/10.1116/5.0080940.

Kingma, D. P. and Ba, J. Adam: A method for stochastic optimization, 2017.

Krenn, M., Kottmann, J. S., Tischler, N., and Aspuru-Guzik, A. Conceptual understanding through efficient automated design of quantum optical experiments. *Phys. Rev. X*, 11:031044, Aug 2021. doi: 10.1103/PhysRevX.11.031044. URL https://link.aps.org/doi/10.1103/PhysRevX.11.031044.

Krenn, M., Pollice, R., Guo, S. Y., Aldeghi, M., Cervera-Lierta, A., Friederich, P., dos Passos Gomes, G., Häse, F., Jinich, A., Nigam, A., et al. On scientific understanding with artificial intelligence. *Nature Reviews Physics*, 4(12):761–769, 2022. URL https://doi.org/10.1038/s42254-022-00518-3.

Labroille, G., Denolle, B., Jian, P., Genevaux, P., Treps, N., and Morizur, J.-F. Efficient and mode selective spatial mode multiplexer based on multiplane light conversion. *Opt. Express*, 22(13):15599–15607, Jun 2014. doi: 10.1364/OE.22.015599. URL https://opg.optica.org/oe/abstract.cfm?URI=oe-22-13-15599.

Lelek, M., Gyparaki, M. T., Beliu, G., Schueder, F., Griffié, J., Manley, S., Jungmann, R., Sauer, M., Lakadamyali, M., and Zimmer, C. Single-molecule localization microscopy. *Nature Reviews Methods Primers*, 1(39), 2021. URL https://doi.org/10.1038/s43586-021-00038-x.

Li, J., Fan, Z., and Fu, Y. FFT calculation for Fresnel diffraction and energy conservation criterion of sampling quality. In *Lasers in Material Processing and Manufacturing*, volume 4915, pp. 180 – 186. SPIE, 2002. doi: 10.1117/12.482883. URL https://doi.org/10.1117/12.482883.

M., D. K. and Palmer, A. E. Advances in fluorescence labeling strategies for dynamic cellular imaging. *Nature Chemical Biology*, 10(7), 2014. URL https://doi.org/10.1038/nchembio.1556.

Mayerhöfer, T. G., Pahlow, S., and Popp, J. The bouguer-beer-lambert law: Shining light on the obscure. *Chemphyschem*, 21:2029 – 2046, 2020. URL https://api.semanticscholar.org/CorpusID:220520649.

Moreau, P.-A., Toninelli, E., Gregory, T., and Padgett, M. J. Imaging with quantum states of light. *Nature Reviews Physics*, 1(6):367–380, 2019. URL https://doi.org/10.1038/s42254-019-0056-0.

Möckl, L., Lamb, D. C., and Bräuchle, C. Super-resolved fluorescence microscopy: Nobel prize in chemistry 2014 for eric betzig, stefan hell, and william e. moerner. *Angewandte Chemie International Edition*, 53(51):13972–13977, 2014. URL https://doi.org/10.1002/anie.201410265.

Möckl, L., Pedram, K., Roy, A. R., Krishnan, V., Gustavsson, A.-K., Dorigo, O., Bertozzi, C. R., and Moerner, W. Quantitative super-resolution microscopy of the mammalian glycocalyx. *Developmental Cell*, 50(1):57–72.e6, 2019. ISSN 1534-5807. doi: https://doi.org/10.1016/j.devcel.2019.04.035. URL https://www.sciencedirect.com/science/article/pii/S1534580719303326.

Müller, P., Müller, R., Hammer, L., Barner-Kowollik, C., Wegener, M., and Blasco, E. Sted-inspired laser lithography based on photoswitchable spirothiopyran moieties. *Chemistry of Materials*, 31(6):1966–1972, 2019. doi: 10.1021/acs.chemmater.8b04696. URL https://doi.org/10.1021/acs.chemmater.8b04696.

Nehme, E., Freedman, D., Gordon, R., Ferdman, B., Weiss, L. E., Alalouf, O., Naor, T., Orange, R., Michaeli, T., and Shechtman, Y. Deepstorm3d: dense 3d localization microscopy and psf design by deep learning. *Nature Methods*, 17:734–740, 2020. URL https://doi.org/10.1038/s41592-020-0853-5.

Nocedal, J. and Wright, S. J. *Numerical optimization*. Springer.

Perrin, M. D., Soummer, R., Elliott, E. M., Lallo, M. D., and Sivaramakrishnan, A. Simulating point spread functions for the james webb space telescope with webbpsf. *Space Telescopes and Instrumentation 2012: Optical, Infrared, and Millimeter Wave*, 8442, 2012. URL https://doi.org/10.1117/12.925230.

Quabis, S., Dorn, R., Eberler, M., Glöckl, O., and Leuchs, G. Focusing light to a tighter spot. *Optics Communications*, 179(1):1–7, 2000. ISSN 0030-4018. doi: https://doi.org/10.1016/S0030-4018(99)00729-4. URL https://www.sciencedirect.com/science/article/pii/S0030401899007294.

Quinteiro, G. F., Schmidt-Kaler, F., and Schmiegelow, C. T. Twisted-light–ion interaction: The role of longitudinal fields. *Phys. Rev. Lett.*, 119:253203, Dec 2017. doi: 10.1103/PhysRevLett.119.253203. URL https://link.aps.org/doi/10.1103/PhysRevLett.119.253203.

Reigoto, A. M., Andrade, S. A., Seixas, M. C. R. R., Costa, M. L., and Mermelstein, C. A comparative study on the use of microscopy in pharmacology and cell biology research. *PLOS ONE*, 16(1):1–13, 2021. URL https://doi.org/10.1371/journal.pone.0245795.

Rubinsztein-Dunlop, H., Forbes, A., Berry, M. V., Dennis, M. R., Andrews, D. L., Mansuripur, M., Denz, C., Alpmann, C., Banzer, P., Bauer, T., Karimi, E., Marrucci, L., Padgett, M., Ritsch-Marte, M., Litchinitser, N. M., Bigelow, N. P., Rosales-Guzmán, C., Belmonte, A., Torres, J. P., Neely, T. W., Baker, M., Gordon, R., Stilgoe, A. B., Romero, J., White, A. G., Fickler, R., Willner, A. E., Xie, G., McMorran, B., and Weiner, A. M. Roadmap on structured light. *Journal of Optics*, 19 (1):013001, nov 2016. doi: 10.1088/2040-8978/19/1/013001. URL https://dx.doi.org/10.1088/2040-8978/19/1/013001.

Rust, M., Bates, M., and Zhuang, X. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (storm). *Nature Methods*, 3:793–796, 2006. URL https://doi.org/10.1038/nmeth929.

Shen, F. and Wang, A. Fast-fourier-transform based numerical integration method for the rayleigh-sommerfeld diffraction formula. *Applied Optics*, 45(6):1102–1110, 2006. doi: 10.1364/AO.45.001102. URL https://opg.optica.org/ao/abstract.cfm?URI=ao-45-6-1102.

Sim, S., Johnson, P. D., and Aspuru-Guzik, A. Expressibility and entangling capability of parameterized quantum circuits for hybrid quantum-classical algorithms. *Advanced Quantum Technologies*, 2(12):1900070, 2019. URL https://doi.org/10.1002/qute.201900070.

Taylor, M. A., Janousek, J., Daria, V., Knittel, J., Hage, B., Bachor, H.-A., and Bowen, W. P. Biological measurement beyond the quantum limit. *Nature Photonics*, 7:229–233, 2013. URL https://doi.org/10.1038/nphoton.2012.346.

Taylor, R. W. and Sandoghdar, V. Interferometric scattering microscopy: Seeing single nanoparticles and molecules via rayleigh scattering. *Nano Letters*, 19(8):4827–4835, 2019. URL https://doi.org/10.1021/acs.nanolett.9b01822.

van de Linde, S., Löschberger, A., Klein, T., Heidbreder, M., Wolter, S., Heilemann, M., and Sauer, M. Direct stochastic optical reconstruction microscopy with standard fluorescent probes. *Nature Protocols*, 6:991–1009, 2011. URL https://doi.org/10.1038/nprot.2011.336.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C. J., Polat, İ., Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., van Mulbregt, P., and SciPy 1.0 Contributors. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272, 2020. doi: 10.1038/s41592-019-0686-2.

Wang, H., Fu, T., Du, Y., Gao, W., Huang, K., Liu, Z., Chandak, P., Liu, S., Van Katwyk, P., Deac, A., et al. Scientific discovery in the age of artificial intelligence. *Nature*, 620(7972):47–60, 2023. doi: 10.1038/s41586-023-06221-2. URL https://doi.org/10.1038/s41586-023-06221-2.

Weisenburger, S. and Sandoghdar, V. Light microscopy: an ongoing contemporary revolution. *Contemporary Physics*, 56(2):123–143, 2015. URL https://doi.org/10.1080/00107514.2015.1026557.

Wollman, A. J. M., Nudd, R., Hedlund, E. G., and Leake, M. C. From animaculum to single molecules: 300 years

of the light microscope. *Open Biology*, 5(4), 2015. URL https://doi.org/10.1098/rsob.150019.

Xu, K., Zhong, G., and Zhuang, X. Actin, spectrin, and associated proteins form a periodic cytoskeletal structure in axons. *Science*, 339(6118):452–456, 2013. doi: 10.1126/science.1232251. URL https://www.science.org/doi/abs/10.1126/science.1232251.

Yanny, K., Antipa, N., Liberti, W., Dehaeck, S., Monakhova, K., Liu, F. L., Shen, K., Ng, R., and Waller, L. Miniscope3d: optimized single-shot miniature 3d fluorescence microscopy. *Light: Science & Applications*, 9(171), 2020. URL https://doi.org/10.1038/s41377-020-00403-7.

Ye, H., Qiu, C.-W., Huang, K., Teng, J., Luk'Yanchuk, B., and Yeo, S. Creation of a longitudinally polarized subwavelength hotspot with an ultra-thin planar lens: Vectorial rayleigh-sommerfeld method. *Laser Physics Letters*, 10(6), 2013. URL https://doi.org/10.1088/1612-2011/10/6/065004.

Yildiz, A., Forkey, J. N., McKinney, S. A., Ha, T., Goldman, Y. E., and Selvin, P. R. Myosin v walks hand-over-hand: Single fluorophore imaging with 1.5-nm localization. *Science*, 300(5628):2061–2065, 2003. doi: 10.1126/science.1084398. URL https://www.science.org/doi/abs/10.1126/science.1084398.

Zhang, Y., Lucas, J. M., Song, P., Beberwyck, B., Fu, Q., Xu, W., and Alivisatos, A. P. Superresolution fluorescence mapping of single-nanoparticle catalysts reveals spatiotemporal variations in surface reactivity. *Proceedings of the National Academy of Sciences*, 112(29):8959–8964, 2015. doi: 10.1073/pnas.1502005112. URL https://www.pnas.org/doi/abs/10.1073/pnas.1502005112.

A. Features and performance of XLUMINA

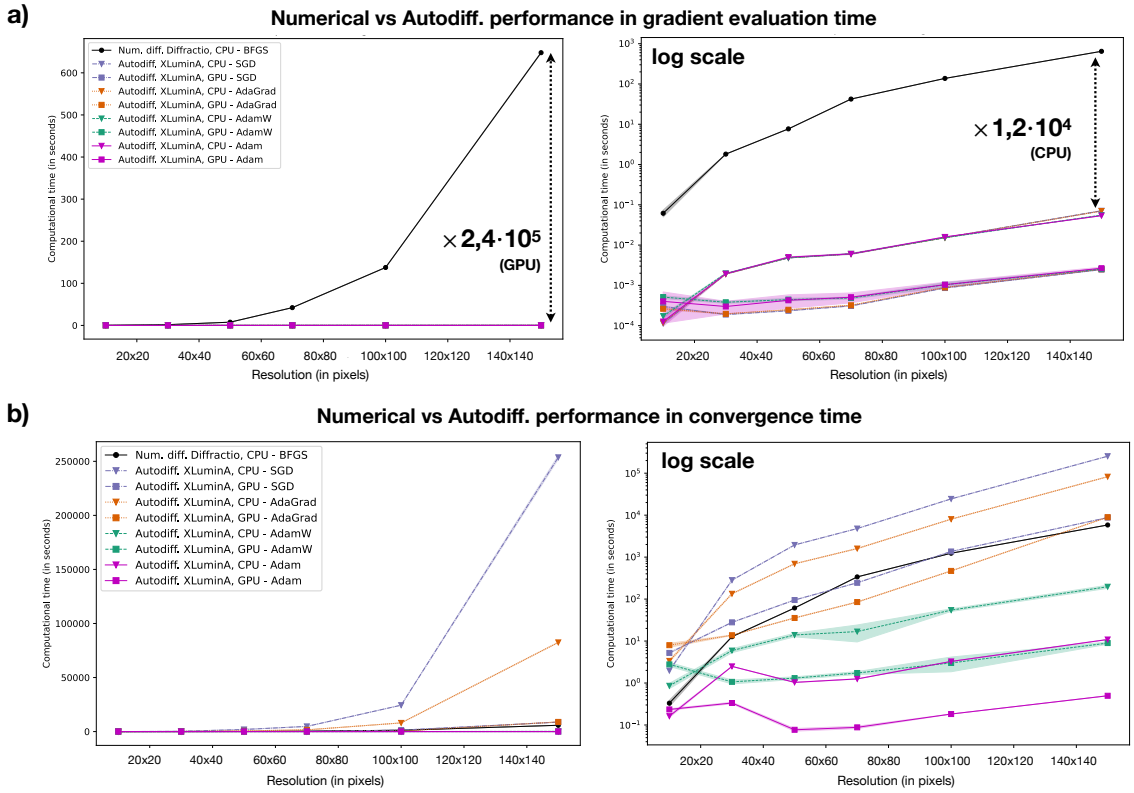
In this section we provide the detailed description of XLUMINA's simulation features and performance. The simulator enables, among many other features, to define light sources (of any wavelength and power), phase masks (i.e., spatial light modulators, SLMs), polarizers, variable retarders (e.g., wave plates, WPs), diffraction gratings, and high numerical aperture (NA) lenses to replicate strong focusing conditions. Light propagation and diffraction is simulated by two methods, each available for both scalar and vectorial regimes: the fast-Fourier-transform (FFT) based numerical integration of the Rayleigh-Sommerfeld (RS) diffraction integral (Shen & Wang, 2006; Ye et al., 2013) and the Chirped z-transform (CZT) (Hu et al., 2020). The CZT is an accelerated version of the RS algorithm, which allows for arbitrary selection and sampling of the region of interest. These algorithms are based on the FFT and require a reasonable sampling for the calculation to be accurate (Li et al., 2002). In our simulations we consider light sources emitting Gaussian beams of 1.2 mm beam waist. To avoid possible boundary-generated artifacts during the simulation, we define these beams in larger computational spaces of 4 mm or 5 mm. Thus, the pixel resolutions often span 1024×1024, or 2048×2048.

Some functionalities of XLUMINA's optics simulator (e.g., optical propagation algorithms, planar lens or amplitude masks) are inspired in an open-source NumPy-based Python module for diffraction and interferometry simulation, *Diffractio* (Brea, 2019), although we have rewritten and modified these approaches to combine them with JAX just-in-time (jit) functionality. In essence, *jit* compilation optimizes sequences of operations together and runs them at once. For this purpose, the first run of a jitted function builds an abstract representation of the sequence of operations specified by the function. This representation encodes the shape and the dtype of the arrays - but is agnostic to the values of such arrays. If the input shapes and dtypes are not modified, the abstract structure of the function can be then re-used for subsequent runs, without re-compilation, which allows to execute the subsequent calls faster. However, if the input shape or dtype is modified, the function automatically gets re-compiled. This will cause an extra overhead time due to the extraction of a new abstract structure of the function for the new shapes/dtypes. On top of that, we developed completely new functions (e.g., beam splitters, WPs or propagation through high NA objective lens with CZT methods, to name a few) which significantly expand the software capabilities. The most important hardware addition on the optical simulator are the SLMs, each pixel of which possesses an independent (and variable) phase value. They serve as a universal approximation for phase masks, including lenses, and offer a computational advantage: given a specific pixel resolution, they allow for unrestricted phase design selection. Such flexibility is crucial during the parameter space exploration, as it allows the software to autonomously probe all potential solutions. In addition, we defined under the name of *super-SLM* (*sSLM*) a hardware-box-type which consists of two SLMs, each one independently imprinting a phase mask on the horizontal and vertical polarization components of the field.

To evaluate the performance of numerical and auto-differentiation methods we chose to use BFGS (from SciPy's Python library) and Adam (included in the JAX library) as optimizers. Further comparison including SGD (Stochastic-Gradient-Descent), AdaGrad (Adaptive Gradient) and AdamW (Adam with weight decay) is presented in Supplementary Fig. 1.

As the optical system, we set-up a Gaussian beam propagating over a distance z and interacting with a phase mask. The objective function is the mean squared error between the detected light and the ground truth, characterized by a Gaussian beam with a spiral phase imprinted on its wavefront. We initialize the system with an arbitrary phase mask configuration. We first evaluate the computational time for a single gradient evaluation for numerical and autodiff methods across different computational window sizes (from 10×10 up to 500×500 pixels) and devices (CPU and GPU). We keep the default settings for BFGS. For Adam, the step size is set to 0.1. The optimization process is terminated if there is no improvement in the loss value (meaning it has not decreased below the best value recorded), over 50 consecutive iteration steps. For each resolution window, we collect the convergence time of both optimizers and divide it by the total number of gradient evaluations for BFGS and the total number of steps for Adam. The acquired gradient evaluation times correspond to the mean value over 5 runs. Obtained results are depicted in Fig. 1a. It is clear how autodiff outperforms numerical methods by up to 4 orders of magnitude on CPU and 5 orders of magnitude when running in the GPU. The advantage over larger sizes is clear given that we run simulations with resolutions of 1024×1024 and 2024×2048 pixels.

We then conduct the evaluation of the convergence time for both methods. We keep the aforementioned settings for the optimizers. We initialize the systems 5 times and compute their mean value. The acquired results are depicted in Fig. 1b. On the CPU, numerical methods exhibit exponential scaling in convergence time, reaching about $4.5 \cdot 10^4$ seconds (roughly 12 hours) for 250×250 pixel resolution. In contrast, autodiff demonstrates superior efficiency, reducing it to roughly 53 seconds. GPU optimization performance is even more pronounced, reaching convergence in 0.24 seconds for 250×250 pixels, and 16 seconds for a resolution of 500×500.



Supplementary Figure 1. Performance of XLUMINA (auto-differentiation) compared to *Diffractio* (numerical methods) across different resolutions and optimizers in (a) single gradient evaluation and (b) convergence time. Data corresponds to the average time over 5 runs. Numerical differentiation is computed using *Diffractio*'s optical simulator and the Broyden-Fletcher-Goldfarb-Shanno (BFGS) optimizer (black dots) and auto-differentiation (triangles for CPU and squares for GPU) on XLUMINA. The Stochastic-Gradient-Descent (SGD), Adaptive Gradient (AdaGrad), Adaptive moment estimation with weight decay (AdamW) and Adaptive moment estimation (Adam) correspond to blue (dash-dot line), orange (dotted line), green (dash line) and magenta (continuous line), respectively. Shaded regions correspond to standard deviation values. The learning rate is set to 0.1 and is common to all the optimizers. For AdamW, the weight decay is set to 10^{-4}. The stopping condition is common to all the frameworks: the optimization is terminated if there is no improvement in the loss value (i.e., it has not decreased below the best value recorded), over 500 consecutive iteration steps. This condition is checked every 100 steps. The use of XLUMINA with autodiff methods improves the gradient evaluation time by a factor of $\times 2.4 \cdot 10^5$ in the GPU and a factor of $\times 1.2 \cdot 10^4$ on the CPU for resolutions of 150×150 pixels. This behavior is common to all the tested optimizers (Adam, AdamW, SGD and AdaGrad). When evaluating the convergence time, the use of Autodiff methods on XLUMINA using the Adam and AdamW optimizers improve the performance with respect to numerical methods by a factor of $\times 1.1 \cdot 10^4$ and $\times 6.5 \cdot 10^2$ in the GPU, respectively, for a resolution of 150×150 pixels. The performance of Adam and AdamW in the CPU demonstrates factors of $\times 5.8 \cdot 10^2$ and $\times 2.9 \cdot 10^1$, respectively, for the same resolution. Remarkably, the use of *Diffractio* with numerical methods (BFGS) outperforms both AdaGrad and SGD in convergence time. In particular, numerical methods outperform AdaGrad by a factor of $\times 1.53$ in the GPU and $\times 14.14$ in the CPU, for a resolution of 150×150 pixels. Similar behavior is demonstrated for SGD: numerical methods outperform it by a factor of $\times 1.50$ in the GPU and $\times 43.50$ in the CPU, for a resolution of 150×150 pixels. Overall, the use of Autodiff methods (in particular, using Adam or AdamW) within GPU-accelerated frameworks is a more appropriate choice to conduct efficient optimization.

B. Large-scale discovery framework

In this section we detail the methodology for the optimizations conducted using our quasi-universal computational *ansatz*, a purely continuous framework. We first discuss the enormous search space corresponding to large-scale optical setups. Then, we provide the derivation of the loss function in equation (1) and our simulated emission depletion model.

B.1. The large-scale space

The large-scale optical setup depicted in Fig. 2 consists 6 light sources that emit linearly polarized Gaussian beams with different wavelengths (e.g., 625 nm, 530 nm and 470 nm). Through 82 vectorial propagation (vectorial Rayleigh-Sommerfeld,

VRS), these beams interact with a total of 9 beam splitters, 24 *sSLM*s (i.e., 48 SLMs), 24 wave plates, and get ultimately detected by 6 high NA objective lenses focusing on light detectors.

We analyze the number of possible discrete arrangements within this general optical setup. For a very discrete approach of beam splitter ratios (either transmit, reflect or have light in both arms) and only allowing the SLMs and wave plates (WP) to be switched ON/OFF (i.e., displaying a constant zero phase or adding zero retardance to the incoming light), the number of possible discrete layouts is of

$$N_{\text{Discrete layouts}} = 3^{9_{\text{BS}}} \cdot 2^{48_{\text{SLM}}} \cdot 2^{24_{\text{WP}}} = 2 \cdot 10^{20}. \tag{2}$$

All these are considering that the available beam splitter ratios are restricted to 3 values and the SLMs and wave plates to turn ON/OFF, respectively. In practice, beam splitter ratios and phase values are continuous variables and can take any value (from 0 to 1 and $-\pi$ to π, respectively) which increases even more the dimension of our search space.

In the following Table 1 we present a summary detailing the main properties of the optimizations conducted within our large-scale ansatz: the number of tunable elements, the dimension of the parameter space and the available number of topologies (for a very discrete approach).

Table 1. Outline of the main properties of the five digital experiments conducted within our large-scale ansatz. Displays the total number of tunable elements, the dimension of the parameter space and the available topologies (for a very discrete approach).

Experiment (Fig. #)	**# tunable elements**	**Parameter space**	**# available topologies**
Fig. 3a	26	$\sim$ 4 million	10^7
Fig. 3d		$\sim$ 6.3 million	
Fig. 4	50	52	10^{18}
Fig. 5	26	$\sim$ 4 million	10^7

B.2. Loss function derivation

The loss function, $\mathcal{L}$, is inversely proportional to the total detected intensity density that is above a specified intensity threshold, I_ε. Thus, minimizing $\mathcal{L}$ aims to maximize the generation of small, high intensity beams. In particular, it reads

$$\mathcal{L} = \frac{1}{\text{Density}} = \frac{\text{Area}}{I_\varepsilon} \ . \tag{3}$$

The total intensity I_ε above the threshold is computed as

$$I_\varepsilon = \sum_{k,l}^{N} i_\varepsilon(k,l) \ , \tag{4}$$

where N is the total number of pixels in the camera's sensor and $i_\varepsilon(k,l)$ represents the intensity value at each pixel once the threshold condition is applied. This condition is defined as follows:

$$i_\varepsilon(k,l) = \begin{cases} i_{\text{det}}(k,l) & \text{if } i_{\text{det}}(k,l) > \varepsilon i_{\text{max}} \ , \\ 0 & \text{otherwise} \ , \end{cases} \tag{5}$$

where $i_{\text{det}}(k,l)$ is the intensity value at the *i*-th row and *j*-th column in the detected 2D intensity pattern, $\varepsilon i_{\text{max}}$ (with $0 \leq \varepsilon \leq 1$) is the threshold value, with i_{max} being the maximum intensity value in the entire 2D detector array.

The Area is determined using a variation of the Heaviside function Θ applied to i_ε, quantifying the area where the intensity is above the threshold:

$$\text{Area} = \sum_{k,l}^{N} \Theta(i_\varepsilon(k,l)) \ , \tag{6}$$

where N is the total number of pixels in the camera's sensor and $\Theta(i_\varepsilon(k,l))$ is defined as:

$$\Theta(i_\varepsilon(k,l)) = \begin{cases} 1 & \text{if } i_\varepsilon(k,l) > 0 \ , \\ 0 & \text{otherwise} \ . \end{cases} \tag{7}$$

Therefore, the loss function can be read as follows:

$$\mathcal{L} = \frac{1}{\text{Density}} = \frac{\text{Area}}{I_\varepsilon} = \frac{\sum_{k,l}^{N} \Theta(i_\varepsilon(k,l))}{\sum_{k,l}^{N} i_\varepsilon(k,l)} \,. \tag{8}$$

Importantly, the camera pixel selection in equation (7) is a discrete operation. However, JAX offers some interesting capabilities due to its integrated autodiff framework. In particular, control flow operations in JAX are supported and differentiable. Therefore, we compute the loss function in a fully differentiable manner using `jax.numpy.where()`.

Crucially, light is detected across six different devices. Therefore, we compute the loss function at each detector and the parameter update is driven by the detector that shows the minimum value of the loss. We conduct this selection by using a differentiable, smooth approximation using `jax.nn.logsumexp()` as:

```
def softmin(l_det, beta):
    return - logsumexp(-beta * l_det)/ beta,
```

where `l_det` is the array of the loss values corresponding to each detector and `beta` is the strength of the modulation.

B.3. Stimulated emission depletion model

STED microscopy (Hell & Wichmann, 1994; Hofmann et al., 2005) is based on excitation and spatially targeted depletion of fluorophores. In order to achieve this, a Gaussian-shaped excitation beam and a doughnut-shaped depletion beam (generated by imprinting a spiral phase into its wavefront) are concentrically overlapped. The depletion beam has zero intensity in the center, where the excitation beam has its maximum. Fluorophores that are not in the center of the beams are forced to emit at the wavelength of the depletion beam. Their emission is spectrally filtered out. Only fluorophores in the center of the beams are allowed to fluoresce normally, and only their emission is ultimately detected. This effectively reduces the area of normal fluorescence, which leads to super-resolution imaging.

We simulate one of the fundamental concepts of STED microscopy without having to rely on time-dependent processes related to absorption and fluorescence. To do so, we perform a nonlinear modulation of the intensity of the excitation and depletion beams based on the Beer-Lambert law (Mayerhöfer et al., 2020). We define the effective fluorescence that would ultimately be detected as:

$$I_{\text{eff}} = I_{\text{ex}} \left[1 - \beta \left(1 - e^{-(I_{\text{dep}}/I_{\text{ex}})} \right) \right] \,, \tag{9}$$

where I_{ex} and I_{dep} correspond to the excitation and depletion intensities, respectively, and $0 \leq \beta \leq 1$ captures the quenching efficiency of the depletion beam. This expression bounds the effect of the depletion beam such that scenarios with negative effective intensity or unrealistically high values are avoided. In particular, assuming a perfect efficiency of the depletion beam in suppressing the excitation (i.e., $\beta = 1$), we obtain an expression resembling the Beer-Lambert law:

$$I_{\text{eff}} = I_{\text{ex}} \cdot e^{-(I_{\text{dep}}/I_{\text{ex}})} \,. \tag{10}$$

Thus, the effective detected light falls off exponentially with the intensity ratio $I_{\text{dep}}/I_{\text{ex}}$. In the limit case where there is no excitation intensity, $I_{\text{ex}} = 0$, the detected light is zero as well, $I_{\text{eff}} = 0$. If there is no depletion intensity, $I_{\text{dep}} = 0$, the detected light corresponds to the excitation beam $I_{\text{eff}} = I_{\text{ex}}$. The trivial case of null efficiency in the quenching, $\beta = 0$, leads to the same result.

To evaluate the nonlinear effect we consider $\beta = 1$ and $I_{\text{dep}} = \frac{1}{2} I_{\text{ex}}$. From equation (9) we obtain

$$I_{\text{eff}} = I_{\text{ex}} e^{-1/2} \approx 0.6 I_{\text{ex}} \,. \tag{11}$$

Now, by slightly increasing the depletion energy, e.g., $I_{\text{dep}} = \frac{3}{2} I_{\text{ex}}$, it reads

$$I_{\text{eff}} = I_{\text{ex}} e^{-3/2} \approx 0.2 I_{\text{ex}} \,. \tag{12}$$

Therefore, a small change in the depletion energy causes a large effect in the effective intensity. As a further example, if we set an intermediate efficiency of $\beta = 0.5$ and $I_{\text{dep}} = \frac{1}{2} I_{\text{ex}}$ we obtain

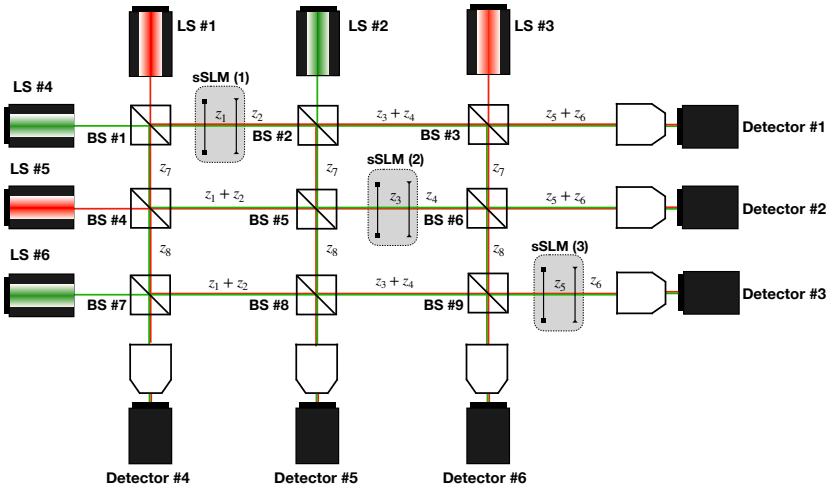
$$I_{\text{eff}} = I_{\text{ex}} \left[\frac{1 + e^{-1/2}}{2} \right] \approx 0.8 I_{\text{ex}} \,. \tag{13}$$

which clearly demonstrates the effect of diminishing the efficiency of the suppression. Overall, we successfully imprinted the nonlinear behavior of the quenching for different range of effectiveness, achieving a realistic, bounded physical model for STED.

C. Rediscovery through exploration

We task XLUMINA to rediscover the super-resolution techniques of STED microscopy (Hell & Wichmann, 1994) and Dorn, Quabis and Leuchs (Dorn et al., 2003). For this purpose we build the 3×3 optical setup depicted in Supplementary Fig. 2.

It consists of six light sources emitting linearly polarized Gaussian beams of wavelengths 650 nm and 532 nm. Three building blocks, which contain one *super*-SLM (i.e., two SLMs imprinting independent phase masks to orthogonal polarization states) and a wave plate separated a distance z, are placed within the diagonal of the grid (grey boxes in Supplementary Fig. 2). Light gets ultimately detected across six detectors. As discussed in the previous Appendix section B.2, the loss function is computed at each detector, the parameter update is driven by the device demonstrating the minimum value. This selection is conducted in a fully differentiable manner using `jax.nn.logsumexp()`.



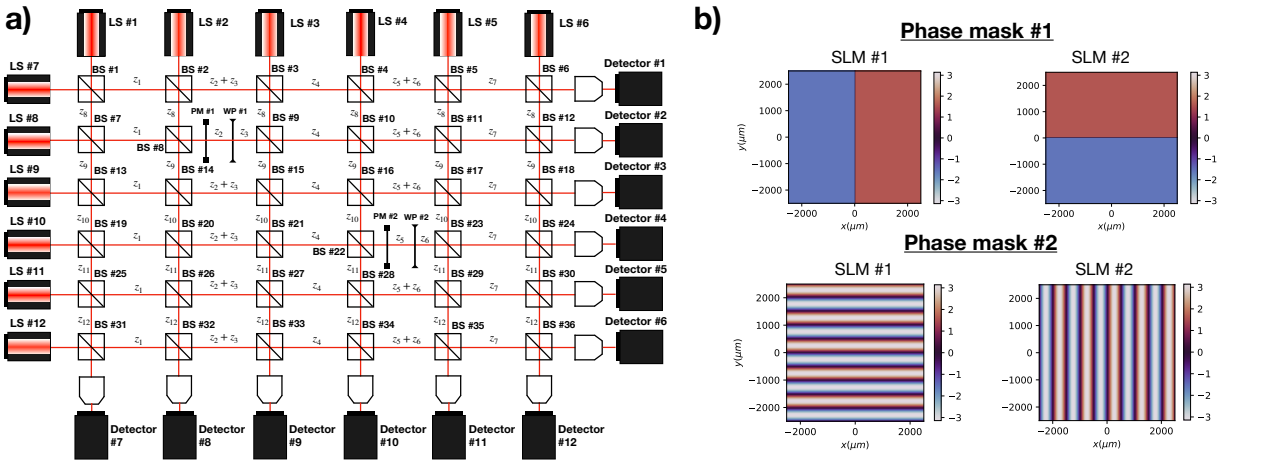
Supplementary Figure 2. Initial virtual optical setup utilized for rediscover SR-techniques and discover a new experimental blueprint. It features light sources that emit Gaussian beams with wavelengths of 650 nm and 532 nm linearly polarized at 45^o. Gray boxes, numbered from (1) to (3), represent the building units, each comprising one *super-SLM* (sSLM) and one wave plate (WP). Distances are denoted as z_i where $i = 1, ..., 8$. After interacting with a high NA objective lens (of NA$= 0.9$), light gets detected across six detectors ($\#1$ - $\#6$) with $0.05\mu m$ pixel size screen. The parameter space (~ 4 million optical parameters) contains three *sSLM*, four wave plates (WPs) with variable phase retardance η and orientation angle θ, eight distances and nine beam splitter ratios. For Dorn, Quabis and Leuchs (2003), all the light sources are set to emit 635 nm wavelength and the resolution is set to 1024×1024 pixels with a computational pixel size of 4.8 μm.

We first target XLUMINA to rediscover the concept of STED microscopy within the general setup in Supplementary Fig. 2. The parameter space (~ 4 million parameters) corresponds to three *super*-SLMs (i.e., 6 SLMs) with a resolution of 824×824 pixels with a pixel size of $6.06\mu m$, three wave plates, eight distances and nine beam splitter ratios. The loss function corresponds to equation (1), in this instance considering the radial intensity of the effective light emerging from the STED process, $|E_x|^2 + |E_y|^2$, and $\varepsilon = 0.5$. We simulate the stimulated emission depletion effect using equation (9) with the efficiency set to $\beta = 1$. We set-up the AdamW optimizer with a learning rate of 10^{-3} and a weight decay of 10^{-2}. The system is initialized with random optical parameters with values between 0 and 1. The optimization is terminated if there is no improvement in the loss value over 500 consecutive iteration steps. This condition is checked every 100 steps. The system converged (in roughly 1.3 hours using a GPU) into the topology highlighted in Fig. 3a, demonstrating the smallest loss value in the detector $\#2$. The identified optical parameters correspond to: the wave plates, in radians (1): $\eta = -1.39$, $\theta = -1.64$, and (2): $\eta = -1.61$, $\theta = -0.86$. The propagation distances (in cm) are $z_1 = 59.52$, $z_2 = 10.14$, $z_3 = 76.36$, $z_4 = 17.93$, $z_5 = 37.07$, $z_6 = 65.95$, and $z_7 = 38.68$. The beam splitter ratios, in [Transmittance, Reflectance] pairs: BS$\#1$: [0.000, 0.999], BS$\#2$: [0.201, 0.799], BS$\#5$: [0.000, 0.999], and BS$\#6$: [0.999, 0.000]. The setup topology is retrieved from detector $\#2$ following the identified beam splitter ratios across the system: $BS\#6$ has 99.9% transmittance,

which implies that light is directly is coming from $BS\#5$. In turn, $BS\#5$ has a reflectance of 99.9% which redirects light from $BS\#2$, which has a transmittance of 79.9%. This already defines light source $\#2$ as the emitter of the green (excitation) beam. Then, the remaining light comes from $BS\#1$ which shows a 99.9% reflectance, meaning that the red light comes from source $\#1$.

To rediscover the concept used in Dorn, Quabis and Leuchs (Dorn et al., 2003), we initialize the system using the general setup in Supplementary Fig. 2. This time, however, the light sources emitting 532 nm wavelength are switched to emit 650 nm. The parameter space ($\sim$ 6.2 million parameters) corresponds to three *super*-SLMs (i.e., 6 SLMs) with a resolution of 1024×1024 pixels with a pixel size of $4.8\mu m$, three wave plates, eight distances and nine beam splitter ratios. The loss function corresponds to equation (1), in this instance considering the intensity from the longitudinal component of the electric field, $|E_z|^2$, and $\varepsilon = 0.7$. We set-up the AdamW optimizer with a learning rate of 0.05 and a weight decay of 10^{-5}. The system is initialized with random optical parameters with values between 0 and 1. The optimization is terminated if there is no improvement in the loss value over 500 consecutive iteration steps. This condition is checked every 100 steps. The system converged (in roughly 35 minutes using a GPU) into the topology highlighted in Supplementary Fig. 3d, demonstrating the smallest loss value in the detector $\#6$. The identified optical parameters correspond to: the wave plate's $\eta = 1.51, \theta = 3.95$; propagation distances (in cm): $z_3 = 20.49$, $z_4 = 63.26$, $z_7 = 47.92$ and $z_8 = 31.33$. The beam splitter ratios, in [Transmittance, Reflectance] pairs: BS$\#$2: [0.999, 0.000], BS$\#$5: [0.000, 0.999], BS$\#$6: [0.000, 0.999], and BS$\#$9: [0.999, 0.000]. As for the previous example, the discovered topology can be identified from the beam splitter ratios: starting from detector $\#6$, $BS\#9$ has 99.9% transmittance, which implies that light is directly is coming from $BS\#6$. In turn, $BS\#6$ has a reflectance of 99.9%. This means that light has interacted with $BS\#5$ and gone through the sSLM and WP. Then, $BS\#5$ has has a reflectance of 99.9%, which defines the incoming light at $BS\#2$. This, in turn, shows a 99.9% transmittance, meaning that light comes from the light source $\#2$.

Importantly, we are not restricted to the use of 3×3 optical grids. Thus, we conduct the same optimization procedure for Dorn, Quabis and Leuchs (2003) this time within a 6×6 optical system. We initialize the optical system depicted in Supplementary Fig. 3a. It consists of 12 light sources emitting a 635 nm wavelength Gaussian beam that are linearly polarized at 45^o interacting with 36 beam splitters and four phase masks displaying the fixed phase patterns in Fig. 3b.



Supplementary Figure 3. Large-scale system utilized for pure topological discovery. (a) Initial virtual setup for Dorn, Quabis and Leuchs (2003). It consists of 12 light sources emitting a 635 nm wavelength Gaussian beam that are linearly polarized at 45^o. The parameter space is defined by 36 beam splitter ratios (BS$\#$1 - BS$\#$36), 12 distances (z_1 - z_{12}) and 2 wave plates with variable applied retardance and orientation). After interacting with a high NA objective lens, light gets detected across 12 detectors ($\#1$ - $\#6$) with $0.05\mu m$ pixel size screen. The parameter update is driven by the camera demonstrating the minimum loss value. (b) Fixed phase patterns corresponding to each phase mask placed on *super*-SLMs. The phase mask $\#1$ corresponds to the polarization converter demonstrated in Dorn, Quabis, and Leuchs (2003). Phase mask $\#2$ correspond to horizontal and vertical gratings.

The loss function corresponds to equation (1) considering the intensity of the electromagnetic field's longitudinal component, $|E_z|^2$, and $\epsilon = 0.7$. We set-up the AdamW optimizer with a step size of 0.05 and a weight decay of 10^{-4}. The system is initialized with random optical parameters with values between 0 and 1. The optimization is terminated if there is no improvement in the loss value over 500 consecutive iteration steps. This condition is checked every 100 steps. The system

converged (in roughly 1 hour using a GPU) into the topology highlighted in Fig. 4a., demonstrating the smallest loss value in the detector $\#3$. The identified optical parameters are: wave plate's (in radians) $\eta = 3.13$ and $\theta = 3.13$; distances (in cm): $z_1 = 167.15$, $z_2 = 133.78$, $z_3 = 102.45$, $z_4 = 65.06$, $z_5 = 81.97$, $z_6 = 71.99$, $z_7 = 104.96$ and $z_9 = 169.96$: and beam splitter ratios, in [Transmittance, Reflectance] pairs: BS$\#7$ [0.999, 0.001], BS$\#8$ [0.996, 0.004], BS$\#9$ [0.841, 0.159], BS$\#10$ [0.791, 0.209], BS$\#11$ [0.015, 0.984], BS$\#17$ [0.280, 0.720], and BS$\#18$ [0.962, 0.038]. As for the previous examples, the discovered topology in Fig. 4a can be retrieved from the beam splitter ratios.

D. Discovery of a new experimental blueprint

Finally, we demonstrate the capabilities of XLUMINA for genuine discovery. We use the initial optical setup in Supplementary Fig. 2. The parameter space (~ 4 million parameters) corresponds to three *super*-SLMs (i.e., 6 SLMs) with a resolution of 824×824 pixels with a pixel size of $6.06\mu m$, three wave plates, eight distances and nine beam splitter ratios. The loss function corresponds to equation (1), in this instance considering the total intensity of the effective light emerging from the STED process, $|E_x|^2 + |E_y|^2 + |E_z|^2$, and $\varepsilon = 0.5$. We simulate the stimulated emission depletion effect using equation (9) with the efficiency set to $\beta = 1$. We set-up the AdamW optimizer with a step size of 10^{-3} and a weight decay of 10^{-3} and initialize the system with random optical parameters with values between 0 and 1. The optimization is terminated if there is no improvement in the loss value over 500 consecutive iteration steps. This condition is checked every 100 steps. The system converged (in roughly 3.8 hours using a GPU) into the topology highlighted in Fig. 5a, demonstrating the smallest loss value in the detector $\#2$. The identified optical parameters correspond to: the beam splitter ratios, in [Transmittance, Reflectance] pairs: BS$\#1$: [0.000, 0.999], BS$\#2$: [0.338, 0.662], BS$\#5$: [0.000, 0.999], and BS$\#6$: [0.999, 0.000]. The wave plates, in radians (1): $\eta = 1.09$, $\theta = 0.28$, and (2): $\eta = 0.19$, $\theta = -3.16$. The propagation distances (in cm) are $z_1 = 29.98$, $z_2 = 56.91$, $z_3 = 58.28$, $z_4 = 99.91$, $z_5 = 42.89$, $z_6 = 53.96$, and $z_7 = 50.96$. The discovered topology can be identified from the beam splitter ratios starting from detector $\#2$: $BS\#6$ has 99.9% transmittance, which implies that light is directly is coming from $BS\#5$. In turn, $BS\#5$ has a reflectance of 99.9% which redirects light from $BS\#2$, which has a transmittance of 66.2%. This already defines light source $\#2$ as the emitter of the green (excitation) beam. Then, the remaining light comes from $BS\#1$ which shows a 99.9% reflectance, meaning that the red light (depletion) comes from source $\#1$.