Compressive Imaging through Optical Fiber with Partial Speckle Scanning*

Stéphanie Guérit[†], Siddharth Sivankutty[‡], John A. Lee[‖§], Hervé Rigneault[‡¶], and Laurent Jacques[†‖]

Abstract. Fluorescence imaging through ultrathin fibers is a promising approach to obtain high resolution imaging with molecular specificity at depths much larger than the scattering mean-free paths of biological tissues. Such imaging techniques, generally termed *lensless endoscopy*, rely upon the wavefront control at the distal end of a fiber to coherently combine multiple spatial modes of a multicore (MCF) or multimode fiber (MMF). Typically, a spatial light modulator (SLM) is employed to combine hundreds of modes by phase-matching to generate a high intensity focal spot. This spot is subsequently scanned across the sample to obtain an image. We propose here a novel scanning scheme, partial speckle scanning (PSS), inspired by compressive sensing theory, that avoids the use of an SLM to perform fluorescent imaging with optical fibers with reduced acquisition time. Such a strategy avoids photo-bleaching while keeping high reconstruction quality. We develop our approach on two key properties of the MCF: (*i*) the ability to easily generate speckles, and (*ii*) the memory effect that allows to use fast scan mirrors to shift light patterns. First, we show that speckles are sub-exponential random fields. Despite their granular structure, an appropriate choice of the reconstruction parameters makes them good candidates to build efficient sensing matrices. Then, we numerically validate our approach and apply it on experimental data. The proposed sensing technique outperforms conventional raster scanning: higher reconstruction quality is achieved with far fewer observations. For a fixed reconstruction quality, our speckle scanning approach is faster than compressive sensing schemes which require to change the speckle pattern for each observation.

Key words. Compressive sensing, lensless endoscope, lensless imaging, speckle imaging, speckle scanning

AMS subject classifications. 78A46, 78A70

1. Introduction.

Nowadays, imaging and exploring the human body for clinical purposes is quite common. Every newborn already experienced at least one imaging modality during its fetal development through ultrasound scans. During their life, most people will encounter medical issues and will possibly undergo physical examinations based on imaging techniques. *In vivo* techniques are traditionally divided into structural and functional imaging.

Although structural imaging has been around since the early days, functional imaging has been the field of tremendous research over the last decades both in the fields of nuclear medecine (*e.g.*, single photon emission computed tomography and positron emission tomography) and optical microscopy. Photonic approaches have the advantage to be simpler to implement. They avoid radioactive labels and can achieve better spatial resolution (sub-micron).

*Submitted to the editors October 18, 2021.

[†]ISPGroup, ELEN and INMA Departments, ICTEAM Institute, UCLouvain, Belgium (st.guerit@gmail.com, laurent.jacques@uclouvain.be).

[‡]Aix Marseille Univ, CNRS, Centrale Marseille, Institut Fresnel, Marseille, France (siddharth@gmail.com, herve.rigneault@fresnel.fr).

[§]Center of Molecular Imaging, Radiotherapy and Oncology (MIRO), IREC Institute, UCLouvain, Belgium (john.lee@uclouvain.be)

[¶]Lightcore Tehnologies, 37-39 Rue d'Antibes, Cannes, France.

[‖]JAL and LJ are Senior Research Associates with the Belgian F.R.S.-FNRS.

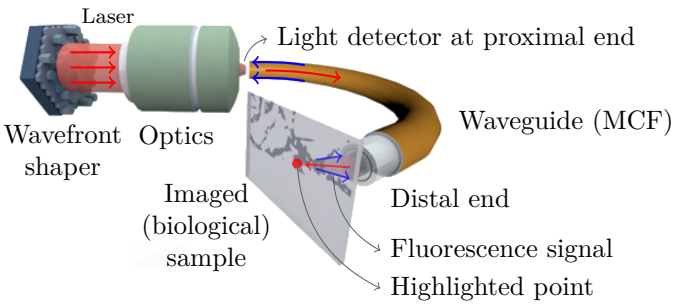


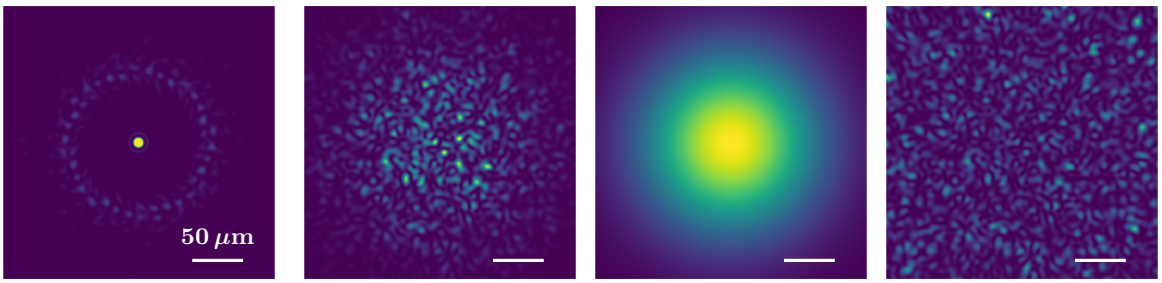
Figure 1: Lensless endoscope principle in raster scanning mode (*i.e.*, with a focused light beam). Excitation signal is red and backscattered fluorescence signal is blue. Source: Institut Fresnel[‡].

However, they usually provide limited field-of-views (FOV) and limited penetration depth due to tissues scattering and absorption as compared to non-optical techniques [45, 5]. To circumvent this latter limitation, optical endoscopy and endomicroscopy are being developed with the goal to provide both structural and functional information.

1.1. Lensless endoscopes as ultrathin devices. For instance, endomicroscopes, already used in clinical applications, use a fiber bundle as waveguide [65, 48]. Recent developments in adaptive optics, optical fibers and computational imaging techniques have opened a new class of imaging systems called *lensless endoscopes* (LE) [20, 5, 10, 50]. In these implementations, a *single* fiber in combination with wavefront shaping devices is employed as an ultrathin imaging system. The extreme miniaturization of the imaging probe (diameter $\leq 200\,\mu$m) offers a minimally invasive route to image at depths unreachable in microscopy. In this paper, we focus our work on LE that use multicore fiber (MCF) [5] and are made with hundreds of individual single core fibers arranged in a single and monolithic silica waveguide [3]. Although the diameter of multicore LE fibers is larger than the diameter of multimode LE fibers [49], MCF can be made resilient to fiber bending [67]. Most importantly, they exhibit a memory effect that allows the output light pattern to be scanned by simply adding a phase tilt at the MCF entrance [65].

As illustrated in Figure 1, an MCF LE consists of four main parts: a wavefront shaper, an optics part, an MCF and an optical detector. The role of the wavefront shaper is to appropriately shape the phase of the light that is injected into the individual cores. This results in the formation of specific illumination patterns at the distal end of the MCF. The optics part is made of mirrors and lenses. It is used to focus the light from the wavefront shaper into the individual cores. To collect the generated fluorescence, the MCF features a double cladding that collects and brings back the fluorescent light towards a high sensitive detector [4]. The sample image is finally reconstructed from the scanning of the focused beam across the sample (by applying a phase tilt at the MCF entrance) and the simultaneous signal collection with the high sensitive detector.

This imaging scheme is known as raster scanning (RS) [50, 59]. It consists in scanning at constant rate each (discretized) position in the high sensitive FOV. For each position, the fluorescence signal is measured by the single pixel detector. The image is then readily



(a) (b) (c) (d)

Figure 2: (a) Focused light beam (brightened for visualization purpose) obtained with the LE with Fermat's golden spiral arranged MCF, (b) speckle pattern (realization of S), (c) mean speckle field $\bar{S}$, and (d) residual field (realization of $\tilde{R}$). Images were simulated with the following parameters (see section 2): $\lambda = 1\,\mu$m, $z = 500\,\mu$m, $3\sigma_c = 3.2\,\mu$m and $\varpi = 0.5\,\mu$m.

reconstructed, pixel by pixel, without any extra post-processing step. As mentioned before, this scanning technique requires preliminary calibration of the SLM to generate a focus beam at the MCF output. A focused illumination pattern is displayed in Figure 2a. This specific focused beam can be obtained when the fiber cores in the MCF are arranged in a Fermat's golden spiral shape. It provides a larger FOV for the LE [60] than periodically arranged cores.

1.2. Current challenges and related works. RS acquisition provides fast image reconstruction but at the cost of some drawbacks, for instance: (*i*) the SLM needs to be calibrated to get a focused beam, (*ii*) imperfections of the focused beam are not corrected, and (*iii*) it is necessary to collect as many observations as the number of pixels in the discrete representation of the image. These considerations highlight constraints on the usage of LE in realistic situations: (*i*) the device is sensitive to perturbations during the calibration and imaging, and (*ii*) the slow update rates of conventional SLMs limit the imaging speed [3].

Hence, the core question of this paper is the following: can we achieve accurate image reconstruction compared to RS using less measurements while keeping a short acquisition time? Without SLM calibration, the observed illumination pattern is a *speckle*, a pattern with multiple bright grains of light and dark regions with no well-defined global peak (see Figure 2b). Such a pattern results from the interference of the multiple coherent light beams produced by each single mode cores when they are configured with arbitrary phases. In some applications, the presence of a speckle pattern is considered as an issue and methods were developed to suppress it or at least to minimize its effects [33]. But in some other imaging modalities, its random structure is appealing as it makes LE closer to recent compressive sensing and computational imaging procedures, *e.g.*, indirectly observing images with random sensing strategies [15, 23, 10, 50].

Back to 1996, Bolshtyansky *et al.* studied an acquisition scheme very similar to ours [9] but involving a single multimode fiber (MMF). Pseudothermal ghost imaging also relies on a similar approach but uses a rotating diffuser to generate speckle patterns [38]. Speckles have also been extensively used to perform super-resolution fluorescence microscopy, most often in

a blind context [44, 43, 39]. To compensate the fact that the patterns may be unknown, some works exploit the statistical properties of the speckles [44, 39] or consider weak assumptions on the original fluorophores distribution [43]. Bertolotti *et al.* studied a blind framework where the diffuser is completely opaque and prevents access to both the sample and the generated speckle [7]. To cover up this inaccessibility, they exploit the *memory effect* summarized as follows: a small change in the incident angle of the light leads to a translation of the speckle pattern. In speckle scanning microscopy, Stasio *et al.* also took advantage of this effect with an MCF fiber [62]. Except in super-resolution, the fluorescence signal in the aforementioned applications is measured with a single-pixel detector (or a similar device). The image of interest is estimated via a weighted sum of the speckle patterns [9], via the correlations of the observation vector with the illumination patterns [38, 56] or via the autocorrelation of the measurement vector [7, 62].

While the above-mentioned works consider a number of measurements at least equal to the number of pixels in the final image, several authors reported the development of compressive optical imaging techniques requiring far fewer observations. These techniques are based on the seminal work of Candès *et al.* and Donoho on compressive sensing (CS) [15, 22]. The pseudo-randomness of the speckles and their easy generation make them interesting to build an efficient sensing matrix. Katz *et al.* proposed a proof of concept for CS approach to pseudothermal ghost imaging including information about the image structure in the reconstruction algorithm [38]. A compressive approach was used in super-resolution microscopy [47] where the authors reach higher resolution by using saturated illumination with speckle patterns. This type of super-resolution compressive imaging was studied theoretically, in a noiseless framework, by Lochocki *et al.* in 2021 [42]. They provide insights on the required number of measurements and the ultimate resolution limits. CS was also exploited in experimental setups involving an MCF and/or an MMF, *e.g.*, in fluorescence microscopy [57, 16], optical photoacoustic [16] or microendoscopy [19]. Authors assume small total variation [57] or sparsity in some wavelet basis [19] as prior information on the image structure. In these three works, speckles patterns are recorded. This recording can be seen as a calibration step. However, as mentioned by [16], this step is simpler than the conventional calibration via wavefront shaping because it only requires to measure the speckles intensities while beamforming requires speckle field measurements before the experiment.

For the wide deployment of CS based endoscopes outside a lab environment, we need to address key issues such as acquisition rate, robustness, and a low barrier to device development (technical complexity and cost). In the literature, fast acquisition is typically satisfied by using a digital-micromirror device as a light-shaping device capable of reaching up to 22 kHz [2, 57, 16]. Nevertheless, these devices are inherently lossy and require a specialized skill-set to realize high-fidelity shaping in conjuction with optical fiber [68]. Since MMFs are highly sensitive to bending, the resilience to minor perturbation is pursued by the use of fiber-bundles or MCFs. The common strategy is to illuminate the sample by launching light into one core at a time at the proximal end and using either a scattering layer [57] or a photonic lantern [19] at the distal end to generate an unique speckle pattern. Addressing only one single mode fiber core at a time ensures a high degree of robustness to perturbation. However, it imposes limitations on the total number of patterns that can be generated (the number of cores) and, importantly, the illumination power that can be delivered to the sample [27].

1.3. Proposed approach and contributions. The literature covered in the previous section shows that speckle illumination seems to be a promising way to face one of the main constraints preventing the use of LE: it removes the need for interferometric stability and costly calibration step before each acquisition, and it can possibly be used in a blind context. Experimental studies show that sensing matrices based on speckles are valid candidates to acquire measurements in a compressive framework while still providing good reconstruction results [69, 57, 42]. In fluorescence imaging, more than a low number of measurements, a short acquisition time is highly desirable to preserve the biological sample of photo-bleaching. Photo-bleaching is an irreversible damage resulting from light exposure and leading to a complete loss of fluorescence. If the MCF has a low coupling between the cores, this time can be substantially reduced by exploiting the memory effect of the fiber thanks to scan mirrors (*e.g.*, galvanometric mirrors) [4]. According to Andresen *et al.*, only strategies exploiting this effect will be able to satisfy the speed and resolution requirements for *in vivo* imaging with LE, as moving such mirrors is much faster than changing an SLM configuration [5].

In this paper, we study a novel acquisition strategy named *partial speckle scanning* (PSS). It combines elements of compressive sensing with the specific properties of the considered LE, namely its *robustness to spatial and temporal distortion*, its extremely *low coupling* between the cores of the fiber, and its *memory effect*. In particular, we generate only $P \leq M$ distinct speckles and we shift each of them $M_P = M/P \leq M$ times, thus achieving a *partial* scanning of the object, to acquire a total of $M = PM_P$ observations. We thus achieved a versatile parametrization where, for $P = M$, we observe the sample with M different speckle illuminations, while if $P = 1$, a unique speckle pattern is shifted M times.

Using this shifting (that is for $M_P > 1$) has two interesting effects. First, by adjusting P and M_P, a tradeoff can be found between the refreshing rate of the SLM and the faster scan mirror rate to reduce the global acquisition time. Second, scanning the speckle also allows for faster imaging procedures—a critical aspect for future high-resolution LE applications. Since the sensing model is computed numerous times inside most iterative image reconstruction methods, this scanning reduces both the complexity and the storage of the sensing model, with direct impact on the global computational time (including the reconstruction).

The first contribution of this paper consists in characterizing an accurate linear sensing model of PSS. In particular, we deduce its equivalent sensing matrix, explaining the relation between the fluorescent sample and the collected observations. We then justify the possibility to estimate an image from PSS observations by establishing links with CS theory. This requires us to first analyze the non-asymptotical distribution of a speckle random field as well as its autocorrelation. The second contribution consists in optimizing the sampling parameters (M_P, P, and the induced speckle shift between two consecutive scanning mirror orientations in a line scanning mode) to get the best reconstruction quality for a given acquisition time or a prescribed number of measurements M. These demonstrations indicate the potential of MCFs in conjunction with computational reconstruction to be viable building blocks for robust and fast lensless endoscopy.

1.4. Outline. The rest of the paper is structured as follows. In section 2, we first determine a model for the illumination produced by the MCF in the object plane. We provide a far-field approximation of this illumination pattern in subsection 2.1 in function of the elec-

tromagnetic field emitted by each core. We explain in subsection 2.2 how we can shift this pattern by using scan mirrors. The formation of a focused illumination (used in RS, see Figure 1) is explained in subsection 2.3 before to introduce and (statistically) characterize in subsections 2.4 and 2.5 the speckle patterns produced by randomly setting the core fields. In section 3, we focus on fluorescent imaging with LE. We detail how we can, in general, image a sample by recording the light emitted under illumination. We first develop a continuous forward model in subsections 3.1 and 3.2 relating the illuminated sample—*i.e.*, the original fluorophore density—to the recorded observations. In subsection 3.3, we detail the discretization of this model, a mandatory step for any computational imaging method. In subsection 3.4, we then show how this discrete forward model is classically inverted in RS imaging, before to describe in subsection 3.4.2 a general image estimation algorithm. It inverts the (possibly ill-conditioned) discrete forward model by regularizing the produced image estimate (*e.g.*, using total variation or more advanced priors). This algorithm can be solved with iterative proximal methods (subsection 3.4.3), and we describe in subsection 3.4.4 a cross-validation strategy to automatically balance the fidelity to the observations with this regularization. In section 4, we develop two compressive imaging strategies for LE, namely speckle imaging (SI) in subsection 4.1, where a distinct random speckle pattern is generated for each observation, and partial speckle scanning (PSS) in subsection 4.2, where the sample is observed from partial scanning of a few randomly generated speckles. These strategies, their respective formulations as inverse problems and the results of the simulations are also presented in subsections 4.1 and 4.2, respectively. Finally, section 5 presents the experimental setup designed to compare different acquisition strategies on real fluorescent samples. We compare our compressive approaches, SI and PSS, to RS and explain how PSS allows us to reduce the acquisition time compared to SI while providing a similar reconstruction quality. Equivalently, we show that PSS permits to image the sample under limited acquisition time budget when SI fails. We conclude in section 6 and provide there possible future improvements for compressive LE imaging.

1.5. Conventions and notations. We find useful to introduce here the conventions and notations used throughout this work. Light symbols are used for scalars and functions, while bold symbols are used for vectors and matrices. The "dot" notation $f(\cdot;\eta)$ refers to the variability of the function f relatively to the pointed parameter with other parameters (in η) fixed. The uniform distribution on a set $\mathcal{S}$ (*e.g.*, $\mathcal{S}=[a,b]\subset\mathbb{R}$ or the unit ball $\mathcal{S}=\mathbb{B}_2\subset\mathbb{R}^2$) is $\mathcal{U}(\mathcal{S})$. $\mathcal{P}(\mu)$ and $\mathcal{N}(\mu,\sigma^2)$ are the Poisson distribution with mean $\mu>0$ and Gaussian distribution with mean μ and variance σ^2, respectively. We also use the index set $[J]\coloneqq\{1,\cdots,J\}$; the Kronecker symbol $\delta_{j,k}$; the vectors of ones and zeros $\mathbf{1}_d$ and $\mathbf{0}_d$, respectively (the subscript d is omitted when clear from the context); the scalar product $\langle\boldsymbol{a},\boldsymbol{b}\rangle=\boldsymbol{a}^\top\boldsymbol{b}$ between two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$; the ℓ_2-norm $\|\boldsymbol{a}\|=\sqrt{\langle\boldsymbol{a},\boldsymbol{a}\rangle}$ of a vector $\boldsymbol{a}$; the relation $\boldsymbol{a}\succeq\mathbf{0}$ as a shorthand for $a_i\geq 0$ for all components i; the identity matrix $\boldsymbol{I}_d$ in $\mathbb{R}^d$; Frobenius norm $\|\boldsymbol{M}\|_F$ and spectral norm $\|\boldsymbol{M}\|$ of a matrix $\boldsymbol{M}$; the function $\mathrm{disk}(\boldsymbol{x})$ equal to 1 if $\|\boldsymbol{x}\|\leq 1$ and 0 otherwise; the symbols "*" and "*" for complex conjugation and Legendre-Fenchel conjugate, respectively; the notation $A\lesssim B$ for meaning that $A\leq cB$ for some constant $c>0$ independent of A and B; and finally, the Fourier transform of f defined as $\hat{f}(\boldsymbol{k})\coloneqq\mathcal{F}[f](\boldsymbol{k})=\int_{\mathbb{R}^2}f(\boldsymbol{x})\,e^{-\mathrm{i}\boldsymbol{k}^\top\boldsymbol{x}}\,\mathrm{d}\boldsymbol{x}$, with inverse $\mathcal{F}^{-1}[g](\boldsymbol{x})=\frac{1}{(2\pi)^2}\int_{\mathbb{R}^2}g(\boldsymbol{k})e^{\mathrm{i}\boldsymbol{k}^\top\boldsymbol{x}}\,\mathrm{d}\boldsymbol{k}$.

2. Illumination modeling. Before going into the details of the acquisition modeling and subsequent sensing strategies, let us have a close look at the models and principles explaining the illumination produced by the considered MCF. We first study how we can set the electromagnetic field of each MCF core to focus the illumination on a small spot on the object plane, before to consider the *speckle illumination* formed by randomly configuring the core field. We also provide a statistical analysis of speckle illumination by studying the first- and second-order statistics of this random pattern as well as an characterization of its distribution. Our conclusions will serve the sensing models developped in section 3.

2.1. Principles and approximations. Given a laser beam of wavelength λ injected into a MCF with J cores (as illustrated in Figure 1), we consider an illumination pattern S in a plane $\mathcal{Z}$, parallel to the planar MCF endface $\mathcal{Z}_0$, and at a distance $z > 0$ from it.

This pattern results from the interferences of the electromagnetic fields radiated by the J cores, each centered on a location $\boldsymbol{q}_j \in \mathbb{R}^2$ ($j \in [J]$) in $\mathcal{Z}_0$. The field of the j^{th} core in fiber endface is well described by a complex amplitude α_j with $|\alpha_j| = 1$ multiplied by a narrow Gaussian envelope [59]. These amplitudes are collectively represented by the vector $\boldsymbol{\alpha} \coloneqq (\alpha_1, \cdots, \alpha_J)^\top$.

In this context, the pattern S is the intensity of the electromagnetic field E_z radiated from $\mathcal{Z}_0$ to the plane $\mathcal{Z}$. Fourier optics tells us that, at each location $\boldsymbol{x} \in \mathcal{Z} \subset \mathbb{R}^2$, E_z is obtained via the angular spectrum representation[1] [32]:

$$E_z(\boldsymbol{x}; \boldsymbol{\alpha}) = (E_0(\cdot; \boldsymbol{\alpha}) * H_z)(\boldsymbol{x}), \tag{2.1}$$

where $E_0(\cdot; \boldsymbol{\alpha})$ is the electromagnetic field emanating from the MCF endface,

$$E_0(\boldsymbol{x}; \boldsymbol{\alpha}) = \left[U_0 * \textstyle\sum_{j=1}^J \alpha_j \delta(\cdot - \boldsymbol{q}_j)\right](\boldsymbol{x}), \tag{2.2}$$

and H_z is the inverse Fourier transform of the angular spectrum propagator defined as $\hat{H}_z(\boldsymbol{k}) = \exp(-\mathrm{i}z(k^2 - \|\boldsymbol{k}\|^2)^{1/2})$ with $k = 2\pi/\lambda$ [32]. The vector $\boldsymbol{k} = (k_x, k_y)^\top$ is the 2-D coordinate in the reciprocal space of $\boldsymbol{x}$ and U_0 is the field emitted by a single core. The field U_0 is approximately Gaussian, *i.e.*, $U_0(\boldsymbol{x}) \approx \exp(-\|\boldsymbol{x}\|^2/(2\sigma_c^2))$, with standard deviation σ_c depending on wavelength λ and related experimentally to mode-field diameter d through $d = 2.35\sigma_c$ [59].

The illumination pattern on a plane located at a distance z of the fiber distal end is therefore

$$S(\boldsymbol{x}; \boldsymbol{\alpha}) \coloneqq |E_z(\boldsymbol{x}; \boldsymbol{\alpha})|^2 = \left|\textstyle\sum_{j=1}^J \alpha_j U_z(\boldsymbol{x} - \boldsymbol{q}_j)\right|^2, \tag{2.3}$$

with $U_z(\boldsymbol{x}) \coloneqq (U_0 * H_z)(\boldsymbol{x})$. This last expression can be simplified with the *far-field* assumption. This assumption considers that the illumination pattern is far from the MCF endface, *i.e.*, $z \gg \frac{2D\mu}{\lambda}$ with D the diameter of the MCF distal end and μ the complex coherence factor of the fiber [34], which is approximately the mode-field waist of a single fiber core, d. In this case, the Fraunhofer approximation holds [32]. Combined with the paraxial approximation

[1] Assuming optical field in homogeneous, isotropic, linear and source-free medium [32].

which supposes a narrow FOV compared to z, *i.e.*, $\|\boldsymbol{x}\| \ll z$, we express E_z as a modulated scaling of the Fourier transform $\hat{E}_0$ (computed relatively to $\boldsymbol{x}$) [21],

$$E_z(\boldsymbol{x};\boldsymbol{\alpha}) \approx -\tfrac{e^{-\mathrm{i}kz}}{\mathrm{i}\lambda z} e^{-\frac{\mathrm{i}k}{2z}\|\boldsymbol{x}\|^2} \hat{E}_0\big(-\tfrac{k}{z}\boldsymbol{x};\boldsymbol{\alpha}\big). \tag{2.4}$$

Therefore, using (2.2) and the convolution theorem, the far-field approximation of the illumination S reads

$$S(\boldsymbol{x};\boldsymbol{\alpha}) \approx \tfrac{1}{(\lambda z)^2} \big|\hat{U}_0\big(-\tfrac{2\pi}{\lambda z}\boldsymbol{x}\big)\big|^2 \Big|\textstyle\sum_{j=1}^{J} \alpha_j e^{\frac{2\pi\mathrm{i}}{\lambda z}\boldsymbol{q}_j^\top \boldsymbol{x}}\Big|^2. \tag{2.5}$$

We introduce now a convenient rewriting of (2.3). We first define the *mean intensity* $\bar{S}$ corresponding to non-interacting fields, *i.e.*, to the sum of all intensity fields produced by each core in $\mathcal{Z}$:

$$\bar{S}(\boldsymbol{x}) \coloneqq \textstyle\sum_{j=1}^{J} |U_z(\boldsymbol{x}-\boldsymbol{q}_j)|^2 = \big[|U_z|^2 * \mathrm{AF}\big](\boldsymbol{x}), \tag{2.6}$$

where $\mathrm{AF}(\boldsymbol{x}) \coloneqq \sum_{j=1}^{J} \delta(\boldsymbol{x}-\boldsymbol{q}_j)$ is the array factor (AF) related to the spatial arrangement of the cores [59]. Quantity $|U_z|^2$ only depends on wavelength λ, the diameter of each core, and distance z. Figure 2c represents $\bar{S}$ for a Fermat's spiral core arrangement. Regarding (2.5), an approximation for the mean intensity is

$$\bar{S}(\boldsymbol{x}) \approx \tfrac{J}{(\lambda z)^2} \big|\hat{U}_0\big(-\tfrac{2\pi}{\lambda z}\boldsymbol{x}\big)\big|^2. \tag{2.7}$$

From (2.3), since the illumination S amounts to summing $\bar{S}$ and all the cross-terms $\sum_{j,k=1;j\neq k}^{J} \alpha_j \alpha_k^* U_z(\boldsymbol{x}-\boldsymbol{q}_j) U_z^*(\boldsymbol{x}-\boldsymbol{q}_k)$, we can study the variations of S around $\bar{S}$ by defining the residual field

$$\tilde{R}(\boldsymbol{x};\boldsymbol{\alpha}) \coloneqq \bar{S}(\boldsymbol{x})^{-1}\big(S(\boldsymbol{x};\boldsymbol{\alpha}) - \bar{S}(\boldsymbol{x})\big), \text{ such that } S(\boldsymbol{x};\boldsymbol{\alpha}) = \bar{S}(\boldsymbol{x})\big(1+\tilde{R}(\boldsymbol{x};\boldsymbol{\alpha})\big). \tag{2.8}$$

This field accounts for the (constructive and destructive) interferences between the individual fields emitted by the cores. Under the Fraunhofer approximation, (2.5) and (2.7) provide

$$\tilde{R}(\boldsymbol{x};\boldsymbol{\alpha}) \approx \tfrac{1}{J} \textstyle\sum_{j,k=1}^{J} \alpha_j \alpha_k^* e^{\frac{2\pi\mathrm{i}}{\lambda z}(\boldsymbol{q}_j-\boldsymbol{q}_k)^\top \boldsymbol{x}} - 1. \tag{2.9}$$

As expressed in (2.8), $\bar{S}$ acts like a *vignetting* window, an envelope on the intensity variations encoded in $\tilde{R}$ for any configurations of $\boldsymbol{\alpha}$ (see Figures 2c and 2d). Moreover, (2.9) shows that $|\tilde{R}(\boldsymbol{x};\boldsymbol{\alpha})| = \mathcal{O}(J)$ and we prove in subsection 2.5 that, for random complex amplitudes $\boldsymbol{\alpha}$, *i.e.*, when S is a speckle pattern, the probability that $|\tilde{R}(\boldsymbol{x};\boldsymbol{\alpha})|$ strongly deviates from a threshold $t > 0$ decays exponentially fast when t increases.

2.2. Translating the illumination. We can translate any illumination pattern produced by the MCF in the object plane $\mathcal{Z}$ by leveraging the low coupling between the J MCF cores, *i.e.*, the MCF *memory effect* [3]. This expresses the fact that, up to an additive constant term depending on intrinsic properties of the core, the phase of the light emitted from j^{th} core is the same as the phase of the input light [5]. Therefore, by modulating the light incident to the MCF—by using galvanometric scan mirrors, as in Figure 1—we act on the complex amplitude

of each core field at the distal end of the fiber, and we can shift the illumination pattern at distance z of the fiber. In particular, we can translate the residual field $\tilde{R}$ defined in (2.8) within the intensity vignetting imposed by the mean speckle field $\bar{S}$.

This effect, which is valid for any configuration of the complex amplitudes, is easily explained as follows. A relative tilt $\boldsymbol{\theta} = (\theta_x, \theta_y)^\top$ of the scan mirrors modifies the light optical path on the different core locations and induces a phase ramp[2] $\frac{2\pi}{\lambda z}\boldsymbol{\theta}^\top \boldsymbol{x}$ on the complex amplitude. The vector $\boldsymbol{\alpha}$ is modified as

$$\boldsymbol{\alpha}' \coloneqq \operatorname{diag}(\boldsymbol{\gamma}(\boldsymbol{\theta}))\,\boldsymbol{\alpha}, \quad \text{with } \gamma_k(\boldsymbol{\theta}) \coloneqq e^{\frac{2\pi \mathrm{i}}{\lambda z}\boldsymbol{\theta}^\top \boldsymbol{q}_k},\ k \in [J]. \tag{2.10}$$

Therefore, considering the far-field approximation of the residual field (2.9), which is valid using the Fraunhofer and the paraxial approximations, this amplitude modulation changes the residual field into

$$\begin{aligned} \tilde{R}(\boldsymbol{x};\boldsymbol{\alpha}') &\approx \tfrac{1}{J}\textstyle\sum_{j,k=1}^{J} \alpha'_j {\alpha'_k}^{*} e^{\frac{2\pi \mathrm{i}}{\lambda z}(\boldsymbol{q}_j-\boldsymbol{q}_k)^\top \boldsymbol{x}} - 1 \\ &= \tfrac{1}{J}\textstyle\sum_{j,k=1}^{J} \alpha_j \alpha_k^{*} e^{\frac{2\pi \mathrm{i}}{\lambda z}(\boldsymbol{q}_j-\boldsymbol{q}_k)^\top (\boldsymbol{x}+\boldsymbol{\theta})} - 1 \approx \tilde{R}(\boldsymbol{x}+\boldsymbol{\theta};\boldsymbol{\alpha}), \end{aligned} \tag{2.11}$$

so that, from (2.8), the speckle field becomes

$$S(\boldsymbol{x};\boldsymbol{\alpha}') \approx \bar{S}(\boldsymbol{x})(1 + \tilde{R}(\boldsymbol{x}+\boldsymbol{\theta};\boldsymbol{\alpha})). \tag{2.12}$$

Thus, provided that the far-field conditions are respected, *i.e.*, if $\|\boldsymbol{x}\| \ll z$ and $z \gg D_0\mu/\lambda$, a non-zero tilt $\boldsymbol{\theta}$ induces a shift $-\boldsymbol{\theta}$ of the residual field, but the vignetting window remains unchanged.

2.3. Focused illumination. For some arrangements of the fiber cores, such as the Fermat's spiral configuration [59], we can focus the intensity pattern S on a narrow intensity spot (see Figure 2a). From a convenient calibration of the optical system [58], this focused beam is obtained by programming the SLM so that each core field has unit complex amplitude in the fiber endface. This induces constructive interferences of the propagated core fields in the origin of the object plane $\mathcal{Z}$. Mathematically, this can be seen by adapting (2.3), or (2.4) in the Fraunhofer approximation, to this particular unit amplitude configuration, which provides

$$\begin{aligned} S_{\mathrm{foc}}(\boldsymbol{x}) \coloneqq S(\boldsymbol{x};\mathbf{1}_J) = |\textstyle\sum_{j=1}^{J} U_z(\boldsymbol{x}-\boldsymbol{q}_j)|^2 &\approx \tfrac{1}{(\lambda z)^2}\big|\hat{U}_0(-\tfrac{2\pi}{\lambda z}\boldsymbol{x})\big|^2 \Big|\textstyle\sum_{j=1}^{J} e^{\frac{2\pi \mathrm{i}}{\lambda z}\boldsymbol{q}_j^\top \boldsymbol{x}}\Big|^2 \\ &= \tfrac{1}{(\lambda z)^2}\big|\hat{U}_0(-\tfrac{2\pi}{\lambda z}\boldsymbol{x})\big|^2 \big|\mathcal{F}[\mathrm{AF}](\tfrac{2\pi}{\lambda z}\boldsymbol{x})\big|^2 \\ &= J^{-1}\,\bar{S}(\boldsymbol{x})\ \big|\mathcal{F}[\mathrm{AF}](\tfrac{2\pi}{\lambda z}\boldsymbol{x})\big|^2. \end{aligned} \tag{2.13}$$

In the far-field, the focused beam $S_{\mathrm{foc}}(\boldsymbol{x})$ is thus also vignetted by $\bar{S}$; it is given by the energy spectral density of the array factor of the fiber cores, as represented in Figure 2a for the focused beam achieved by Fermat's spiral core arrangement [59].

[2]The phase ramp is normalized by $\frac{2\pi}{\lambda z}$ to ease our next developments; the actual tilt is $\boldsymbol{\theta}' = \frac{\lambda z}{2\pi}\boldsymbol{\theta}$.

2.4. Speckle illumination. Speckle patterns were observed for the first time at the end of the 19$^{\text{th}}$ century from astronomical observations [25, 33]. They result from the interferences between coherent light components with random relative delays, as induced by light reflection on a rough surface. In this work, we propose to generate speckle pattern with the MCF by randomly configuring the complex amplitudes $\boldsymbol{\alpha}$ of the intensity field $S(\boldsymbol{x},\boldsymbol{\alpha})$ in (2.3). We set these amplitudes so that their phases follow a uniform distribution on the complex circle, *i.e.*, $\alpha_j \sim_{\text{i.i.d.}} \exp\big(\mathrm{i}\mathcal{U}([0,2\pi])\big)$ for $j\in[J]$. This random configuration can be reached by properly shaping the laser beam thanks to the SLM (see Figure 1).

In this random configuration, we easily show that, from (2.3), the mean intensity $\bar{S}$ is also the expectation of S since $\mathbb{E}\alpha_j\alpha_k^* = \delta_{jk}$, *i.e.*,

$$\mathbb{E}S(\boldsymbol{x},\boldsymbol{\alpha}) = \mathbb{E}\big|\textstyle\sum_{j=1}^J \alpha_j U_z(\boldsymbol{x}-\boldsymbol{q}_j)\big|^2 = \sum_{j=1}^J \big|U_z(\boldsymbol{x}-\boldsymbol{q}_j)\big|^2 = \bar{S}(\boldsymbol{x}). \tag{2.14}$$

The residual field $\tilde{R}$ (defined in (2.8)) of such a speckle pattern exhibits a spatial granular structure indicating that, for two points $\boldsymbol{x},\boldsymbol{x}'$ close to each other, the two r.v.s $\tilde{R}(\boldsymbol{x};\boldsymbol{\alpha})$ and $\tilde{R}(\boldsymbol{x}';\boldsymbol{\alpha})$ are correlated. Given the diameter D of the MCF fiber, we show below that such a correlation exists as soon as $\|\boldsymbol{x}-\boldsymbol{x}'\| \lesssim \lambda z/D$, which shows that the size of a "speckle grain" scales like $\mathcal{O}(\lambda z/D)$. This is achieved by studying the second-order statistics of the random field $\tilde{R}$: the spatial autocorrelation defined as

$$\Gamma_{\tilde{R}}(\boldsymbol{x},\boldsymbol{x}+\boldsymbol{\tau}) \coloneqq \mathbb{E}[\tilde{R}(\boldsymbol{x};\boldsymbol{\alpha})\tilde{R}^*(\boldsymbol{x}+\boldsymbol{\tau};\boldsymbol{\alpha})], \tag{2.15}$$

with $\boldsymbol{\tau} \coloneqq (\tau_x,\tau_y)^\top$. Using the approximation (2.9), $\mathbb{E}[\alpha_j\alpha_k^*] = \delta_{jk}$, and $\mathbb{E}[\alpha_j\alpha_k^*\alpha_l^*\alpha_m] = \delta_{jk}\delta_{lm}+\delta_{jl}\delta_{km}-\delta_{jklm}$, we find

$$\begin{aligned}
\Gamma_{\tilde{R}}(\boldsymbol{x},\boldsymbol{x}+\boldsymbol{\tau}) &\approx \tfrac{1}{J^2}\textstyle\sum_{j,k,l,m}\mathbb{E}[\alpha_j\alpha_k^*\alpha_l^*\alpha_m]e^{\frac{2\pi\mathrm{i}}{\lambda z}(\boldsymbol{q}_j-\boldsymbol{q}_k)^\top\boldsymbol{x}}e^{-\frac{2\pi\mathrm{i}}{\lambda z}(\boldsymbol{q}_l-\boldsymbol{q}_m)^\top(\boldsymbol{x}+\boldsymbol{\tau})} - 1\\
&= \tfrac{1}{J^2}\textstyle\sum_{j,l=1}^J 1 + \tfrac{1}{J^2}\sum_{j,k=1}^J e^{-\frac{2\pi\mathrm{i}}{\lambda z}(\boldsymbol{q}_j-\boldsymbol{q}_k)^\top\boldsymbol{\tau}} - \tfrac{1}{J^2}\sum_{j=1}^J 1 - 1\\
&= \tfrac{1}{J^2}\big|\textstyle\sum_{j=1}^J e^{-\frac{2\pi\mathrm{i}}{\lambda z}\boldsymbol{q}_j^\top\boldsymbol{\tau}}\big|^2 - \tfrac{1}{J}\\
&= \tfrac{1}{J^2}\Big|\mathcal{F}[\mathrm{AF}]\big(\tfrac{2\pi}{\lambda z}\boldsymbol{\tau}\big)\Big|^2 - \tfrac{1}{J} \approx \Gamma_{\tilde{R}}(\boldsymbol{0},\boldsymbol{\tau}) =: \Gamma_{\tilde{R}}(\boldsymbol{\tau}).
\end{aligned} \tag{2.16}$$

Therefore, the Fraunhofer approximation shows us that the autocorrelation $\Gamma_{\tilde{R}}(\boldsymbol{x},\boldsymbol{x}+\boldsymbol{\tau}) \approx \Gamma_{\tilde{R}}(\boldsymbol{\tau})$ does not depend on $\boldsymbol{x}$ anymore. We also see that, in the far-field, $\Gamma_{\tilde{R}}$ is directly related to the Fourier transform of the array factor AF, and it displays a peak on the origin with $\Gamma_{\tilde{R}}(\boldsymbol{0}) = 1-J^{-1} \le 1$. Note that the Fermat's spiral configuration is designed for ensuring small variations of $\Gamma_{\tilde{R}}$ far from the origin, thus keeping only a dominant peak in $\boldsymbol{0}$ [59].

The decay of $\Gamma_{\tilde{R}}$ when $\tau = \|\boldsymbol{\tau}\|$ increases can be further estimated by assuming the core locations homogeneously distributed over 2-D fiber endface. Following [33], assuming these locations distributed uniformly at random over 2-D fiber endface of diameter D, *i.e.*, $\boldsymbol{q}_j \sim_{\text{i.i.d.}} \mathcal{U}(\frac{1}{2}D\mathbb{B}^2)$ with $\mathbb{B}^2$ the unit ball in $\mathbb{R}^2$, we find, for J large and $\boldsymbol{q},\boldsymbol{q}' \sim_{\text{i.i.d.}} \mathcal{U}(\frac{1}{2}D\mathbb{B}^2)$,

$$\begin{aligned}
\Gamma_{\tilde{R}}(\boldsymbol{\tau}) &\approx \tfrac{1}{J^2}\textstyle\sum_{j,k=1;j\neq k}^J e^{-\frac{2\pi\mathrm{i}}{\lambda z}(\boldsymbol{q}_j-\boldsymbol{q}_k)^\top\boldsymbol{\tau}}\\
&\approx \tfrac{J-1}{J}\mathbb{E}e^{-\frac{2\pi\mathrm{i}}{\lambda z}(\boldsymbol{q}-\boldsymbol{q}')^\top\boldsymbol{\tau}} = \tfrac{J-1}{J}\big|2(\tfrac{\pi D}{\lambda z}\tau)^{-1}J_1(\tfrac{\pi D}{\lambda z}\tau)\big|^2,
\end{aligned} \tag{2.17}$$

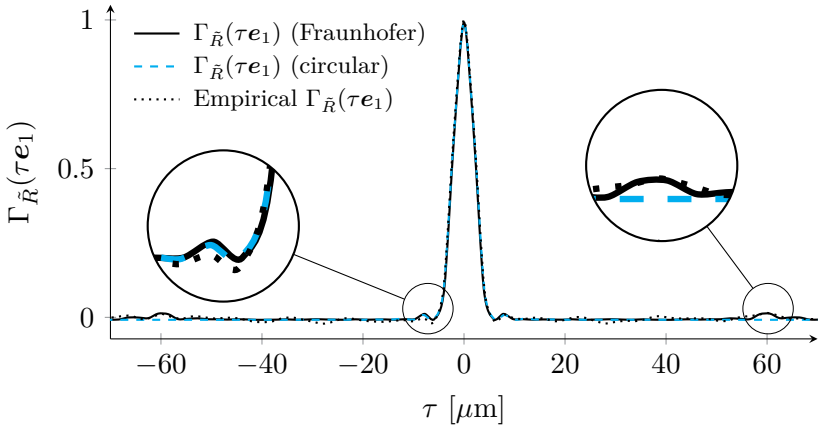


Figure 3: Autocorrelation of field $\tilde{R}$ for three cases: (*i*) using Fraunhofer approximation (solid line), (*ii*) using Fraunhofer approximation and assumption of a circular aperture (dashed cyan line), and (*iii*) by empirically computing mean (over 25 trials) autocorrelation based on 1,000 realizations of $\tilde{R}$ (dotted line). The fiber core arrangement is a Fermat's golden spiral with $J = 120$ cores, diameter $D = 113\,\mu$m and $3\sigma_c = 3.2\,\mu$m. Experiment parameters are $\lambda = 1\,\mu$m, $z = 500\,\mu$m, and $\varpi = 2\,\mu$m.

where J_1 is the first-order Bessel function of the first kind, and $\lim_{s\to 0} J_1(s)/s = 1/2$. Following [33], we define the average area of a speckle grain as $\mathcal{A} \coloneqq \int\!\!\int_{-\infty}^{\infty} (\Gamma_{\tilde{R}}(\boldsymbol{\tau})/\Gamma_{\tilde{R}}(\mathbf{0}))\mathrm{d}\boldsymbol{\tau}$. For a circular aperture as considered above, $\mathcal{A}$ is obtained using Parseval's identity and the average radius is defined as $r \coloneqq \sqrt{\mathcal{A}/\pi}$, which gives

$$\mathcal{A} = \frac{(\lambda z)^2}{\pi (D/2)^2} \quad \text{and} \quad r = \frac{\lambda z}{\pi (D/2)}. \tag{2.18}$$

We thus get $r \propto \lambda z/D$. For the parameters of the simulation in Figure 3, the radius of a speckle grain is $2.8\,\mu$m.

Remark 2.1. Quantities (2.17) and (2.18) are easily computable since they only depend on the experiment parameters J, D, λ and z. Figure 3 shows $\Gamma_{\tilde{R}}$ as a function of $\boldsymbol{\tau}$ (with $\boldsymbol{\tau} = \tau \boldsymbol{e}_1$ and $\boldsymbol{e}_1$ the unit vector aligned with the horizontal axis of the 2-D image) obtained with Fraunhofer approximation and the stronger assumption of a circular spot. Mean empirical estimations of $\Gamma_{\tilde{R}}$ based on 1 000 realizations of the residual field are also displayed for $\boldsymbol{x} = \mathbf{0}$. Both theoretical approximations explain well the main peak of the autocorrelation, that is similar to the peak observed when the light beam is focused (see Figure 2a). Even if distant side lobes are absent with circular aperture assumption, (2.17) still provides a reliable closed form expression for the radius of a speckle grain.

Remark 2.2. Interestingly, (2.16) shows us that the focused beam (2.13) is actually related to the autocorrelation $\Gamma_{\tilde{R}}$ of a random speckle, *i.e.*, $S_{\mathrm{foc}}(\boldsymbol{x})/\bar{S}(\boldsymbol{x}) \approx J\Gamma_{\tilde{R}}(\boldsymbol{x}) + 1$ since $J^2\Gamma_{\tilde{R}}(\boldsymbol{\tau}) + J \approx |\mathcal{F}[\mathrm{AF}](\frac{2\pi}{\lambda z}\boldsymbol{\tau})|^2$.

2.5. Speckle distribution. In [33], Goodman derives the probability density function of a speckle with a number of phasors that tends to infinity (corresponding, for instance, to

the use of a diffuser). In this *asymptotic* case, applying the central limit theorem is relevant and leads to an intensity $S(\boldsymbol{x};\boldsymbol{\alpha})$ distributed according to an exponential law [33, chapter 3]. However, assuming such an asymptotic behavior is not appropriate when we work with a small number of cores like in the MCF described in this work. In this section, we show that the (non-asymptotic) distribution of $S(\boldsymbol{x};\boldsymbol{\alpha})$ is sub-exponential when we choose i.i.d. complex amplitudes α_j, no matter their distribution. Recently, Bender *et al.* demonstrated methods to tailor more specifically the speckle intensity statistics (and not only the tail decay) by appropriately choosing the distribution of the complex amplitudes [6].

Following [70, 71], a *sub-exponential* (or *sub-Gaussian*) r.v. $X \in \mathbb{C}$ is characterized by a rapidly decaying tail bound $\mathbb{P}\{|X| \geq t\} \leq 2e^{-ct/K}$ (resp. $\mathbb{P}\{|X| \geq t\} \leq 2e^{-ct^2/K^2}$) when $t \geq 0$ increases. In this bound, the constant $c > 0$ is universal, and $K > 0$ depends only on the distribution of X. This characterization naturally includes Laplacian (resp. Gaussian) r.v.s, which explains the name of this distribution class.

Equivalently, a sub-exponential (or *sub-Gaussian*) r.v. X has a finite sub-exponential (resp. sub-Gaussian) norm $\|X\|_{\psi_1}$ (resp. $\|X\|_{\psi_2}$), with

$$\|X\|_{\psi_p} \coloneqq \sup_{q\geq 1} q^{-1/p}(\mathbb{E}|X|^q)^{1/q}, \quad p \in \{1,2\}. \tag{2.19}$$

In fact, for $p \in \{1,2\}$, the tail bound constant K is proportional to $\|X\|_{\psi_p}$, so that one is finite *if and only if* the other is.

We can show that for the complex amplitudes α_j both norms are finite, *i.e.*, $\|\alpha_j\|_{\psi_1} \leq 1$ and $\|\alpha_j\|_{\psi_2} \leq 1$ for $j \in [J]$; if a r.v. X is bounded ($|X| \leq C$), then $\|X\|_{\psi_p} \leq C$ for $p \in \{1,2\}$.

Proposition 2.3. *Given a location $\boldsymbol{x}$ and J i.i.d. complex random amplitudes $\{\alpha_j, j \in [J]\}$ with unit modulus, both the speckle intensity $S(\boldsymbol{x};\boldsymbol{\alpha})$ and the residual field $\tilde{R}(\boldsymbol{x};\boldsymbol{\alpha})$ are sub-exponential, and*

$$\|S(\boldsymbol{x};\boldsymbol{\alpha})\|_{\psi_1} \lesssim \bar{S}(\boldsymbol{x}) \quad \textit{and} \quad \|\tilde{R}(\boldsymbol{x};\boldsymbol{\alpha})\|_{\psi_1} \lesssim 1. \tag{2.20}$$

Proof. Defining $\boldsymbol{u}(\boldsymbol{x}) \coloneqq (U_z(\boldsymbol{x}-\boldsymbol{q}_1), \cdots, U_z(\boldsymbol{x}-\boldsymbol{q}_J))^* \in \mathbb{C}^J$, we can rewrite (2.3) as $S(\boldsymbol{x};\boldsymbol{\alpha}) = |\langle \boldsymbol{\alpha}, \boldsymbol{u}(\boldsymbol{x})\rangle|^2$. From [70, Lemma 5.14], $\|X\|_{\psi_2}^2 \leq \|X^2\|_{\psi_1} \leq 2\|X\|_{\psi_2}^2$, therefore, $\||\langle \boldsymbol{\alpha}, \boldsymbol{u}(\boldsymbol{x})\rangle|^2\|_{\psi_1} \leq 2\|\langle \boldsymbol{\alpha}, \boldsymbol{u}(\boldsymbol{x})\rangle\|_{\psi_2}^2$. Moreover, from [70, Lemma 5.9], for J i.i.d. sub-Gaussian r.v.s X_j, we have the approximate *rotation invariance* $\|\sum_{j=1}^J X_j\|_{\psi_2}^2 \lesssim \sum_{j=1}^J \|X_j\|_{\psi_2}^2$. Note that, for $X_j \coloneqq \alpha_j u_j^*(\boldsymbol{x})$, $\langle \boldsymbol{\alpha}, \boldsymbol{u}(\boldsymbol{x})\rangle = \sum_{j=1}^J X_j$, using $\|\alpha_j\|_{\psi_2} \leq 1$ and the fact that $\|\cdot\|_{\psi_2}$ is a norm, this involves that

$$\|\langle \boldsymbol{\alpha}, \boldsymbol{u}(\boldsymbol{x})\rangle\|_{\psi_2}^2 \lesssim \sum_{j=1}^J \|X_j\|_{\psi_2}^2 = \sum_{j=1}^J \|\alpha_j\|_{\psi_2}^2 |u_j(\boldsymbol{x})|^2 \leq \sum_{j=1}^J |u_j(\boldsymbol{x})|^2 = \bar{S}(\boldsymbol{x}),$$

which proves the first inequality in (2.20). Regarding the second inequality, since $\|\cdot\|_{\psi_1}$ is a norm, it respects the triangular inequality, and from (2.8), we get

$$\begin{aligned}\|\tilde{R}(\boldsymbol{x};\boldsymbol{\alpha})\|_{\psi_1} &= \|\bar{S}(\boldsymbol{x})^{-1}(S(\boldsymbol{x};\boldsymbol{\alpha}) - \bar{S}(\boldsymbol{x}))\|_{\psi_1} \\ &\leq \|\bar{S}(\boldsymbol{x})^{-1}S(\boldsymbol{x};\boldsymbol{\alpha})\|_{\psi_1} + 1 = \bar{S}(\boldsymbol{x})^{-1}\|S(\boldsymbol{x};\boldsymbol{\alpha})\|_{\psi_1} + 1 \lesssim 1.\end{aligned}$$

■

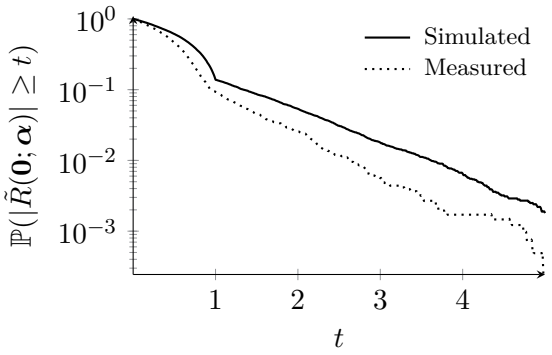


Figure 4: Estimation of $\mathbb{P}\big(|\tilde{R}(\mathbf{0};\boldsymbol{\alpha})| \geq t\big)$ as a function of $t \geq 0$ based on 10,000 simulated speckle patterns of size 128×128 (solid line) and 4,096 experimentally measured speckle patterns of size 256×256 (dotted line). The fiber core arrangement is a Fermat's golden spiral. In both cases, we do observe an exponential decay of the tail of the distribution ($t \geq 1$). Simulation parameters are $J = 120$, $D = 113\,\mu$, $3\sigma_c = 3.2\,\mu$m, $\lambda = 1\,\mu$m, $z = 500\,\mu$m, and $\varpi = 2\,\mu$m. The experimental setup used to acquire real observations is described in subsection 5.1.

Proposition 2.3 informs us that the tail of the distribution of $\tilde{R}(\boldsymbol{x};\boldsymbol{\alpha})$ decreases exponentially fast with a threshold $t \geq 0$, *i.e.*, for some universal constants $C, c > 0$,

$$\mathbb{P}\big(|\tilde{R}(\boldsymbol{x};\boldsymbol{\alpha})| \geq t\big) \leq Ce^{-ct}.$$

Figure 4 displays approximations of $\mathbb{P}(|\tilde{R}(\mathbf{0};\boldsymbol{\alpha})| \geq t)$ as a function of t based on both simulated and experimental discretized speckle patterns. For $t \geq 1$ (corresponding to the tail), we do observe an exponential decay of the distributions. The difference in the decay rates is due to slightly different simulation and experiment parameters (*e.g.*, λ and ϖ).

3. From illumination to fluorescent imaging. In this section, we first describe the fluorescence phenomenon, *i.e.*, the physical process underlying the generation of observations in lensless endoscopy. Second, we describe the model of photons collection by the fiber. Then, based on those previous models, we propose a forward model explaining how the observations are related to the original fluorophore density map. Finally, we formulate the imaging process as the solving of an inverse problem, *i.e.*, the minimization of an objective function made of a data fidelity term combined with regularizing priors on the expected density map, and explain how to solve it numerically.

More formally, we consider a sample—either a synthetic phantom or a very thin slice of an *ex vivo* biological sample—containing some molecules of interest that have been tagged with fluorescent dyes. For the sake of simplicity, we assume that this sample is restricted to a 2-D FOV, also called object space $\Omega \subset \mathbb{R}^2$. A density function $f_0 : \Omega \times \mathbb{R} \to \mathbb{R}_+$ associates a fluorophore density $f_0(\boldsymbol{x}, t)$ with each location $\boldsymbol{x} \in \Omega$ at time instant t.

3.1. Fluorescence model. Fluorescent molecules possibly absorb light and emit photons under illumination by a light pattern with static intensity $S(\boldsymbol{x})$. The observed and measured photon flux $\varphi : \Omega \times \mathbb{R} \to \mathbb{R}$ depends on the intensity of the incident light but also on the

fluorophore density f_0. In this section, we derive the expression of the photon flux $\varphi(\boldsymbol{x}, t)$ at time t emitted by the fluorophores located at $\boldsymbol{x} \in \Omega$ as a function of f_0, S and physical parameters. In the next section, we will relate this expression to the measurements obtained with the optical device.

The main phenomena controlling the number of emitted photons are singlet state and triplet state saturations, photo-bleaching, and background fluorescence [66]. We refer the reader to [61, 66] for detailed explanations about the physics of fluorescence emission.

In the context of LE, we formulate the following assumptions about the illumination duration and intensity, both tunable acquisition parameters: (i) illumination duration t_{acq} is short compared to the time needed to reach triplet excited steady state but long enough for singlet excited state to reach steady state, and (ii) intensity S at each $\boldsymbol{x}$ is low and, consequently, far from saturation level. The first assumption allows us to consider only singlet excited state saturation and its corresponding flux. The second assumption amounts to considering that $S(\boldsymbol{x}) \ll k_d/\sigma$, where k_d is the constant rate associated with the return to the ground state and σ is a parameter depending on laser wavelength λ and on the chosen dye [18, 66]. In this case, the photon flux emitted at $\boldsymbol{x}$ depends linearly on the product of $S(\boldsymbol{x})$ and $f_0(\boldsymbol{x})$,

$$\varphi(\boldsymbol{x}) \approx Q_e \sigma \, S(\boldsymbol{x}) f_0(\boldsymbol{x}). \tag{3.1}$$

where quantum yield Q_e characterizes the efficiency of the emission process. This time-invariant model also assumes that there is no molecules displacement during the acquisition—*i.e.*, the illumination duration is short—and we neglect photo-bleaching of fluorescent dye.

3.2. Photon collection model. For a static illumination pattern $S(\boldsymbol{x})$, the LE collects within the acquisition duration t_{acq} a non-negative number of photons Y, a fraction of those emitted by the fluorescent sample. Neglecting for now any other noise sources that could plague this photon collection such as thermal and readout noise in the optical detector, Y is a Poisson r.v. whose mean is

$$\mathbb{E}Y = \int_\Omega \int_0^{t_{\text{acq}}} c_0 \left(\varphi(\boldsymbol{x}) + \varphi_b(\boldsymbol{x})\right) \, \mathrm{d}t \, \mathrm{d}\boldsymbol{x}, \tag{3.2}$$

where φ and φ_b are the direct and the background photon fluxes generated by the object, and $0 < c_0 < 1$ is a constant accounting for the fraction of photons captured by the LE. As explained in subsection 5.1, the model (3.2) assume no sensitivity attenuation of the light collection (as induced by the fiber numerical aperture) closer to the periphery of the FOV.

Using (3.1) and (3.2), and assuming φ_b is induced by constant density over the FOV, the continuous forward model relating the photon count Y to f_0 is equivalent to

$$Y \sim \mathcal{P}\Big(\int_\Omega S(\boldsymbol{x}) [c_1 f_0(\boldsymbol{x}) + c_2] \, \mathrm{d}\boldsymbol{x} \Big), \tag{3.3}$$

with constants $c_1, c_2 > 0$ depending on c_0, Q_e, σ and t_{acq}. In this study, we are interested in the relative contrast between different regions of the object space Ω. Recovering f_0 up to a scaling factor and an offset is adequate. Defining $f \coloneqq c_1 f_0 + c_2$, model (3.3) becomes

$$Y \sim \mathcal{P}\Big(\int_\Omega S(\boldsymbol{x}) f(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} \Big). \tag{3.4}$$

3.3. Simplified discrete forward model. We now simplify (3.4) by turning it into a discrete representation. For this, we suppose that both f and S can be represented in a finite set of $N = n^2$ orthonormal functions $\{\psi_i\}_{i=1}^N$ over $\Omega \subset \mathbb{R}^2$, *i.e.*, there exist coefficients $\{s_i\}_{i=1}^N$ and $\{f_i\}_{i=1}^N$ such that $S(\boldsymbol{x}) \approx \sum_{i=1}^N s_i\psi_i(\boldsymbol{x})$ and $f(\boldsymbol{x}) \approx \sum_{i=1}^N f_i\psi_i(\boldsymbol{x})$. By Parseval, this also means that $\int_\Omega S(\boldsymbol{x})f(\boldsymbol{x})\mathrm{d}\boldsymbol{x} \approx \boldsymbol{s}^\top\boldsymbol{f}$, where $\boldsymbol{s} \coloneqq (s_1,\dots,s_N)^\top$ and $\boldsymbol{f} \coloneqq (f_1,\dots,f_N)^\top$. In this work, we select a naive pixel representation with pixel pitch $\varpi > 0$, *i.e.*, $\psi_i(\boldsymbol{x}) \coloneqq \psi(\boldsymbol{x}-\boldsymbol{x}_i)$ with $\mathcal{X} \coloneqq \{\boldsymbol{x}_i\}_{i=1}^N$ evenly sampling Ω on N locations and $\psi(\boldsymbol{x})$ equals to $1/\varpi$ if $\boldsymbol{x} \in [-\varpi/2,\varpi/2]\times[-\varpi/2,\varpi/2]$, and 0 otherwise. Assuming a square FOV with side length L, we thus set $\varpi = L/n$.

Combined with the corruption of any possible signal-independent noise sources (*e.g.*, electronic and readout noise) appearing in the measurement process, the continuous model (3.4) becomes

$$Y \sim \mathcal{P}(\boldsymbol{s}^\top\boldsymbol{f}) + \mathcal{N}(0,\sigma^2), \tag{3.5}$$

where all extra noise sources are collectively modeled as a Gaussian noise with variance σ^2.

This work considers an estimation of the biological specimen, assimilated to its fluorophore density map $\boldsymbol{f}$, from the information captured from M distinct observations $\boldsymbol{y} \coloneqq (y_1,\cdots,y_M)^\top \in \mathbb{N}^M$, *i.e.*, M realizations of Y. Each observation y_j is associated with a discrete illumination patterns $\boldsymbol{s}_j$, the discrete representation of the field $S(\boldsymbol{x},\boldsymbol{\alpha}_j)$—focused or speckle—set by one complex amplitude configuration $\boldsymbol{\alpha}_j \in \mathbb{C}^J$. As justified by our experimental conditions (see section 5), this estimation is also realized in a high photon couting regime where the following approximation holds for Poisson distributed random vector,

$$Y \sim \boldsymbol{s}^\top\boldsymbol{f} + \mathcal{N}(0,\boldsymbol{s}^\top\boldsymbol{f}+\sigma^2). \tag{3.6}$$

Gathering the M illumination patterns in the matrix $\boldsymbol{S} \coloneqq (\boldsymbol{s}_1,\cdots,\boldsymbol{s}_M)$ and following the conclusions of section 2, we can thus compactly represent the sensing model as

$$\boldsymbol{y} = \boldsymbol{S}^\top\boldsymbol{f} + \boldsymbol{n}, \tag{3.7}$$

where $\boldsymbol{n} \coloneqq (n_1,\dots,n_M)^\top$ amounts to M realizations of the noise sources $n_i \sim \mathcal{N}(0,\sigma_i^2)$, with variance $\sigma_i^2 \coloneqq \boldsymbol{s}_i^\top\boldsymbol{f}+\sigma^2$. When we perform speckle illumination, the noise variance σ_i^2 is well approximated by a deterministic but unknown value given by $\sigma_f^2 \coloneqq \mathbb{E}[\boldsymbol{s}_i^\top\boldsymbol{f}+\sigma^2] = \bar{\boldsymbol{s}}^\top\boldsymbol{f}+\sigma^2$, which is independent of the measurement index i. In practice, we describe in subsection 3.4.4 a cross-validation technique that determines the influence of $\boldsymbol{n}$ (and its unknown variance σ_f^2) in the estimation of $\boldsymbol{f}$ from $\boldsymbol{y}$.

3.4. Imaging process. In this section, after a brief review of LE imaging by raster scanning of a focused beam, we present a more general imaging process allowing object observation with speckle illumination. The image reconstruction can then be seen as solving of an inverse problem—accounting for the indirect observation of the fluorophore density map—regularized by some prior information on the structure of this map.

3.4.1. Raster scanning imaging. As mentioned in subsection 1.1, RS is a conventional sensing strategy for lensless endoscopy [59]. It amounts to scanning the object of interest with

the focused light beam $S_{\text{foc}}(\boldsymbol{x})$ defined in (2.13). As explained in subsection 2.2, the action of the scanning mirrors allows us to move the center of S_{foc} along a trajectory Ω_{tr} in the FOV Ω, and to record one observation of the fluorophore density f for each trajectory position. Thus, the number of observations crucially depends on the total length of Ω_{tr} and its sampling rate, both tunable parameters of the experiment.

Let us assume that we get M observations $\boldsymbol{y} = (y_1, \cdots, y_M)$ for a light beam focused at M discrete locations $\Omega_{\text{d}} \coloneqq \{\boldsymbol{p}_i : i \in [M]\} \subset \Omega_{\text{tr}}$ that belong to the pixel grid $\mathcal{X}$ (see subsection 3.3). In this case, in a noiseless setting and in the discrete model developped in subsection 3.3, RS imaging directly estimates f from the observations, *i.e.*,

$$\hat{f}(\boldsymbol{p}_i) = y_i = \boldsymbol{s}_{\text{foc},i}^\top \boldsymbol{f} \approx f(\boldsymbol{p}_i), \tag{3.8}$$

where $\boldsymbol{s}_{\text{foc},i}$ is the discrete representation of S_{foc} when it is translated on $\boldsymbol{p}_i$.

In other words, we consider that the matrix $\boldsymbol{S}$ generated by the patterns $\{\boldsymbol{s}_{\text{foc},i}\}_{i=1}^M$ in (3.7) is well approximated by the identity operator over the FOV specified by Ω_{d}. However, even with a fiber arrangement optimized to produce point-like light beams, distant side lobes in the light pattern (see Figure 2a) limit the accuracy of this approximation. Moreover, setting unit amplitude α_j in the fiber endface requires to calibrate the SLM, and the FOV must be densely scanned by Ω_{tr}. This last point forces the RS trajectory to visit each point of a fine grid $\mathcal{X}$. We explain below how one can alleviate these limitations by adopting a compressive imaging strategy relying on random speckle illumination.

3.4.2. Inverse problem formulation. In the context of either focused RS imaging or for speckle illumination, we can *invert* the forward model given by (3.7). We do so by posing the reconstruction of the fluorophore density map $\boldsymbol{f}$ as the solving of an inverse problem, that is, the estimation of $\boldsymbol{f}$ from the (indirectly) observed data $\boldsymbol{y}$ for a general sensing matrix $\boldsymbol{S}$. A natural but naive way to find this estimate amounts to picking the density $\hat{\boldsymbol{f}}$ minimizing the squared ℓ_2-norm of the residual $\boldsymbol{S}^\top \hat{\boldsymbol{f}} - \boldsymbol{y}$. This minimization problem is unfortunately often *ill-posed*, mainly due to the noise corruption but also because $\boldsymbol{S}^\top$ is not necessarily well conditioned, *e.g.*, if the number of observations M is smaller than the number of pixels N. Consequently, we cannot guarantee the uniqueness or even the existence of a solution to this least squares minimization.

We overcome this issue by regularizing the inverse problem, *i.e.*, by adding extra objective functions g_1 and g_2 (or priors) in the minimization [12]. We solve this multi-criterion problem by minimizing the weighted sum of the objective functions:

$$\hat{\boldsymbol{f}} \in \arg\min_{\boldsymbol{u}} \|\boldsymbol{S}^\top \boldsymbol{u} - \boldsymbol{y}\|^2 + \rho_1 g_1(\boldsymbol{u}) + \rho_2 g_2(\boldsymbol{u}), \tag{3.9}$$

where $\rho_1, \rho_2 > 0$ are *regularization parameters* balancing between the data fidelity term (and thus the measurement noise level) and regularization terms g_1 and g_2. We detail in subsection 3.4.4 how to automatically estimate those parameters. In this work, we focus on two extra priors g_1 and g_2.

Image structure. We assume that $\boldsymbol{f}$ is made of smooth areas separated by sharp boundaries (possibly corresponding to tissue interfaces or cellular membranes). This prior, encoded in function g_1, corresponds either to the minimization of the total variation (TV) norm [55]

or the second-order total generalized variation (TGV^2_α) norm [13, 40] (see Appendix A for formal definitions). Promoting a small TV-norm leads to piecewise constant images and is then well suited for the synthetic data used in the simulations. As for TGV, it can be seen as a generalization of TV to higher-order image derivatives. Minimizing the TGV^2_α-norm leads to piecewise linear estimates where parameter $\alpha > 0$ mades a trade-off between edge-preserving and smoothness-promoting terms (as suggested by [40], we set $\alpha = 2$). TGV^2_α is better suited for fluorescent samples used for the experiments or when we estimate a vignetted map.

Non negativity. We know from the physics of fluorescence emission that $\boldsymbol{f} \succeq \mathbf{0}$. Function g_2 is then a convex indicator function on the set $\mathbb{R}^N_+$ that is equal to zero if the constraint is satisfied and to $+\infty$ otherwise, $g_2 = \imath_{\mathbb{R}^N_+}$.

3.4.3. Minimization algorithm. Since we do not require g_1 and g_2 to be differentiable, we cannot minimize (3.9) with methods relying on the computation of the gradient or the Hessian of the objective function. Instead, we resort to the family of proximal algorithms [46] that can deal with optimization of non differentiable functions. In a nutshell, these algorithms can minimize the sum of several convex non-smooth functions by splitting this optimization into an iterative algorithm relying on the computation of the *proximal operator* of each of these functions. This operator is defined as follows.

Definition 3.1. *(from [46]) Let ψ be a lower semicontinuous convex function from $\mathcal{S} \subset \mathbb{R}^d$ to $]-\infty, +\infty[$ such that the domain of ψ is non-empty. The proximal operator of $\psi : \mathcal{S} \to \mathcal{S}$ evaluated in $\boldsymbol{z} \in \mathcal{S}$ is unique and defined as*

$$\mathrm{prox}_\psi(\boldsymbol{z}) \coloneqq \underset{\boldsymbol{x}\in\mathcal{S}}{\arg\min}\ \tfrac{1}{2}\|\boldsymbol{z}-\boldsymbol{x}\|_2^2 + \psi(\boldsymbol{x}).$$

The evaluation of the proximal operator of a convex function ψ on $\boldsymbol{z}$ thus provides a minimizer of ψ that remains close to $\boldsymbol{z}$. For many smooth and non-smooth convex functions, this operator is closed form or fast to compute; this is the case of the ℓ_1-norm, the T(G)V-norm, the indicator function of a convex set, and the functions g_1 and g_2 defined above.

In this work, we solve (3.9) with a generalized version of the Chambolle-Pock (CP) primal-dual algorithm [17, 31]. This flexible algorithm allows for composing convex functions with linear operators in the minimization, *i.e.*, it can be used to solve the following type of problem [31],

$$\min_{\boldsymbol{u}}\ h(\boldsymbol{u}) + \textstyle\sum_{k=1}^K \varphi_k(\boldsymbol{A}_k\boldsymbol{u}). \tag{3.10}$$

Solving (3.9) with TV regularization thus corresponds to solving (3.10) with $K = 2$, $h(\boldsymbol{u}) = \imath_{\mathbb{R}^N_+}(\boldsymbol{u})$, $\varphi_1(\boldsymbol{u}) = \|\boldsymbol{u}-\boldsymbol{y}\|^2$, $\varphi_2(\boldsymbol{u}) = \|\boldsymbol{u}\|_1$, and, in the case where g_1 is the TV-norm, the matrix $\boldsymbol{A}_1 = \boldsymbol{S}^\top$ and $\boldsymbol{A}_2 = \nabla$ is the finite difference operator (defined in Appendix A). If g_1 is the TGV^2_α-norm, Appendix A provides the definition of the functions and the corresponding operators. The CP algorithm only requires computing the proximal operators of h, the Legendre-Fenchel conjugate of each φ_k, as well as the adjoints of the matrices $\boldsymbol{A}_k$'s. Notice that, as an alternative, the Generalized Forward Backward algorithm can also solve (3.9) by leveraging the fact that the function $\|\boldsymbol{S}^\top\boldsymbol{u}-\boldsymbol{y}\|^2$ is differentiable with respect to $\boldsymbol{u}$ [52].

3.4.4. Regularization parameter and stopping criterion. As brought up earlier, the optimal value of the parameter ρ balancing the data fidelity term and the regularization term g_1 in (3.9) is unknown *a priori*. From a maximum a posteriori standpoint, this parameter actually depends on the chosen regularization term g_1 and on noise corrupting the observations $\boldsymbol{y}$ (3.7). We can, however, set a value for ρ based on the optimization of a criterion depending on the estimate and the observations. Boufounos *et al.*, followed by Ward, proposed for instance to use cross-validation (CV) to estimate such a balancing parameter [11, 72]. To avoid overfitting noisy observations in the context of compressive sensing, the set of observations can be divided into two parts: a first set is used for the signal estimation, and a second, much smaller, is used to determine when to stop an iterative reconstruction algorithm, or to select appropriate regularization parameters.

We propose here to use the same CV approach. We first randomly split the vector of M observations in $\boldsymbol{y}$ in an estimation and validation vectors $\boldsymbol{y}_{\text{est}}$ and $\boldsymbol{y}_{\text{val}}$ of size M_{est} and $M_{\text{val}} = M - M_{\text{est}}$, respectively, with M_{val} representing a small fraction of M (see section 4). We use for this two restriction matrices, $\boldsymbol{R}_{\text{est}} \in \{0,1\}^{M_{\text{est}} \times M}$ and $\boldsymbol{R}_{\text{val}} \in \{0,1\}^{M_{\text{val}} \times M}$, such that $\boldsymbol{y}_{\text{val}} = \boldsymbol{R}_{\text{val}}\, \boldsymbol{y}$ and similarly for $\boldsymbol{y}_{\text{est}}$. In this context, the estimation problem (3.9) becomes

$$\hat{\boldsymbol{f}}_\rho \in \arg\min_{\boldsymbol{u}} \|\boldsymbol{R}_{\text{est}}(\boldsymbol{S}^\top \boldsymbol{u} - \boldsymbol{y})\|^2 + \rho g_1(\boldsymbol{u}) + \imath_{\mathbb{R}_+^N}(\boldsymbol{u}), \tag{3.11}$$

As in [11], we use the squared ℓ_2-norm of the residual to evaluate the quality of the produced estimate $\hat{\boldsymbol{f}}_\rho$. This is defined by

$$q(\hat{\boldsymbol{f}}_\rho) \coloneqq \|\boldsymbol{R}_{\text{val}}(\boldsymbol{S}^\top \hat{\boldsymbol{f}}_\rho - \boldsymbol{y})\|^2.$$

This quality function q is then used, first, to stop the CP algorithm when q reaches a minimum, and second, to select an appropriate value for ρ. This last operation is done by solving (3.11) for K values of ρ and ρ_{k^*} is considered as the best choice if $q(\hat{\boldsymbol{f}}_{\rho_{k^*}}) \leq q(\hat{\boldsymbol{f}}_{\rho_k})$, for all $k \in [K]$.

4. Compressive imaging with speckle illumination. We show in this section how the LE acquisition model (3.7) can follow a compressive sensing (CS) scheme [15], *i.e.*, when the density map $\boldsymbol{f} \in \mathbb{R}^N$ is observed with $M < N$ random speckle illuminations. This also aim at removing the beamforming calibration of the RS method by using speckle illumination. We thus consider a sensing model (3.7) where the patterns $\{\boldsymbol{s}_k\}_{k=1}^M$ composing $\boldsymbol{S} \in \mathbb{R}_+^{N \times M}$ are associated with M random configurations $\{\boldsymbol{\alpha}_k\}_{k=1}^M \subset \mathbb{C}^J$ of the MCF complex field amplitudes. As mentioned in subsection 1.2, speckle patterns realized with monochromatic light and a multiple scattering medium were shown, mostly experimentally, to be suitable candidates for generating the sensing matrix in a CS framework [41, 57, 42].

In CS theory, the success of the estimation of a signal of interest mainly depends on the ability of the sensing matrix to capture information about this signal during the acquisition process [53]. This theory offers tools to study the recovery of low-complexity signals (such as sparse images in a given basis) when the number of observations is much smaller than the number of pixels in the final image, *e.g.*, the restricted isometry property (RIP) and the null space property [15, 53]. But as mentioned in [24, 53, 64], those two properties are difficult to verify directly (*e.g.*, for deterministic matrices) without extra assumptions on the sensing

matrix. Most of the existing theoretical work in CS theory is about random matrices with zero-mean i.i.d. row entries, or random partial Fourier sampling. For instance, the behavior of a reconstruction based on observations acquired with sub-Gaussian or Bernoulli measurement matrices has been extensively studied from a theoretical point of view [15, 37]. However, those zero-mean random matrices with i.i.d. entries (or rows) are too ideal for real-life applications where the physics often limits the choice of sensing operator [24, 53].

As detailed in the next subsections, we study two practical compressive acquisitions for LE, and we show how to make them, at least partially, compatible with CS theory. The first strategy, called *speckle imaging* (**SI**), observes the object with M distinct speckle patterns generated by M i.i.d. complex amplitudes configurations $\{\boldsymbol{\alpha}_k\}_{k=1}^M$. The second, coined *partial speckle scanning* (**PSS**), adopts a structured compressive sensing strategy; while observing the object, only a fraction of the M speckles are generated randomly, and the others consist of shifted copies of the firsts.

4.1. Speckle imaging. In SI, the phases of the complex amplitudes of the J MCF cores are all independently picked uniformly at random over $[0, 2\pi]$ for each measurement, *i.e.*, given $\boldsymbol{\alpha}_k = (\alpha_{k1}, \cdots, \alpha_{kJ})^\top$, we have $\alpha_{kj} \sim_{\text{i.i.d.}} \exp(\mathrm{i}\mathcal{U}([0, 2\pi]))$, for $j \in [J]$ and $k \in [M]$. This unstructured compressive sensing schemes thus acquires M observations by successively illuminating the biological sample with M different unstructured light patterns $\boldsymbol{s}_j$, each obtained from a random configuration of the SLM. In this context, we first note that the sensing model (3.7) can be rewritten as

$$\boldsymbol{y} = \boldsymbol{S}^\top \boldsymbol{f} + \boldsymbol{n} = \sqrt{M}\, \boldsymbol{\Phi}\, \bar{\boldsymbol{S}} \boldsymbol{f} + \boldsymbol{n}, \tag{4.1}$$

where the sensing matrix $\boldsymbol{\Phi} \in \mathbb{R}^{M \times N}$ is a renormalization of $\boldsymbol{S}^\top$ defined as

$$\boldsymbol{\Phi} \coloneqq \tfrac{1}{\sqrt{M}} \boldsymbol{S}^\top \bar{\boldsymbol{S}}^{-1}, \quad \bar{\boldsymbol{S}} \coloneqq \operatorname{diag}(\bar{\boldsymbol{s}}) \in \mathbb{R}^{N \times N}, \tag{4.2}$$

and $\bar{\boldsymbol{s}} \coloneqq (\bar{s}_1, \ldots, \bar{s}_N)^\top \in \mathbb{R}_+^N$ is the discrete representation (over the pixel grid $\mathcal{X}$) of the mean speckle field $\bar{S}$ defined in (2.6). The rationale of the renormalized model (4.1) follows the conclusion of section 2. Since by construction each row $\bar{\boldsymbol{S}}^{-1}\boldsymbol{s}_k$ of the random matrix $\sqrt{M}\boldsymbol{\Phi}$ in (4.2) corresponds to sampling $S(\boldsymbol{x}; \boldsymbol{\alpha}_k)/\bar{S}(\boldsymbol{x})$ on the pixel grid $\mathcal{X}$ (provided that $\bar{S}(\boldsymbol{x})$ evolves slowly compared to S, which basically means that $1/\sigma_c \gg 1/D$), Proposition 2.3 shows that each Φ_{jk} are sub-exponential for a random speckle illumination generated with random complex amplitudes.

However, the entries of $\boldsymbol{\Phi}$ are biased: their mean is $1/\sqrt{M}$ since $\bar{S} = \mathbb{E}S$. Such a bias is detrimental to the application of CS theory; for instance, it prevents $\boldsymbol{\Phi}$ to satisfy the RIP property [53]. Following [51, 63], we can consider the following *debiased observation model* that amounts to modify both the observations and the sensing model:

$$\boldsymbol{z} \coloneqq \boldsymbol{y} - (\bar{\boldsymbol{s}}^\top \boldsymbol{f})\mathbf{1}_M = \sqrt{M}\tilde{\boldsymbol{\Phi}}\bar{\boldsymbol{S}}\boldsymbol{f} + \boldsymbol{n}, \tag{4.3}$$

where $\tilde{\boldsymbol{\Phi}}$ is the debiased sensing matrix

$$\tilde{\boldsymbol{\Phi}} \coloneqq \boldsymbol{\Phi} - \tfrac{\mathbf{1}_M \mathbf{1}_N^\top}{\sqrt{M}} = \tfrac{1}{\sqrt{M}} \boldsymbol{S}^\top \bar{\boldsymbol{S}}^{-1} - \tfrac{\mathbf{1}_M \mathbf{1}_N^\top}{\sqrt{M}} \bar{\boldsymbol{S}} \bar{\boldsymbol{S}}^{-1} = \tfrac{1}{\sqrt{M}} (\boldsymbol{S}^\top - \mathbf{1}_M \bar{\boldsymbol{s}}^\top) \bar{\boldsymbol{S}}^{-1}. \tag{4.4}$$

In this new model, we thus collect M observations $\boldsymbol{z}$ of the object $\boldsymbol{f}$ thanks to the sensing matrix $\tilde{\boldsymbol{\Phi}} = (\tilde{\boldsymbol{r}}_1, \dots, \tilde{\boldsymbol{r}}_M)^\top / \sqrt{M} \in \mathbb{R}^{M \times N}$. By construction, its random entries are zero-mean. Moreover, comparing (4.4) to (2.8), we see that each row $\tilde{\boldsymbol{r}}_k = \bar{\boldsymbol{S}}^{-1}(\boldsymbol{s}_k - \bar{\boldsymbol{s}})$ of $\sqrt{M}\tilde{\boldsymbol{\Phi}} = (\tilde{\boldsymbol{r}}_1, \dots, \tilde{\boldsymbol{r}}_M)^\top$ corresponds—for a slowly varying mean field $\bar{S}$—to the discrete representation of $\tilde{R}(\boldsymbol{x}; \boldsymbol{\alpha}_k)$ as defined in (2.8). Therefore, the sub-exponential norm of the entries of $\tilde{\boldsymbol{\Phi}}$ are bounded by $1/\sqrt{M}$.

However, the debiased model (4.3) is impractical; the quantity $(\bar{\boldsymbol{s}}^\top \boldsymbol{f})\mathbf{1}_M$ is unknown. Up to a slight increase of the noise level, we can solve this problem by assuming that the noise $\boldsymbol{n}$ corrupting the observations is zero-mean. In this case, since $\mathbb{E}\frac{1}{M}\mathbf{1}_M^\top \boldsymbol{y} = \mathbb{E}\frac{1}{M}\sum_{i=1}^M y_i = \frac{1}{M}\sum_{i=1}^M \mathbb{E}(\boldsymbol{s}_i^\top \boldsymbol{f} + n_i) \approx \bar{\boldsymbol{s}}^\top \boldsymbol{f}$, where the expectation refers to the randomness of both $\boldsymbol{n}$ and all the $\{\boldsymbol{s}_i\}_{i=1}^M$, we can reach this realistic sensing model

$$\tilde{\boldsymbol{y}} \coloneqq \boldsymbol{y} - \tfrac{1}{M}(\mathbf{1}_M^\top \boldsymbol{y})\mathbf{1}_M = \sqrt{M}\tilde{\boldsymbol{\Phi}}\bar{\boldsymbol{S}}\boldsymbol{f} + \tilde{\boldsymbol{n}}, \tag{4.5}$$

with $\tilde{\boldsymbol{n}} \coloneqq \boldsymbol{n} + \boldsymbol{n}'$ and $\boldsymbol{n}' \coloneqq (\bar{\boldsymbol{s}}^\top \boldsymbol{f})\mathbf{1}_M - \frac{1}{M}(\mathbf{1}_M^\top \boldsymbol{y})\mathbf{1}_M$. Since $\mathbb{E}\frac{1}{M}(\mathbf{1}_M^\top \boldsymbol{y}) = \bar{\boldsymbol{s}}^\top \boldsymbol{f}$ and each y_i are independent r.v.s, we can compute that the variance of the entries of $\boldsymbol{n}'$ decay like $\mathcal{O}(1/M)$ when M increases. Therefore, the variance $\tilde{\sigma}^2$ of each components of $\tilde{\boldsymbol{n}}$ behaves like $\tilde{\sigma}^2 = \sigma^2 + \mathcal{O}(1/M)$, which is close to σ^2, the variance of each n_k, if M is large.

In model (4.5), the entries of $\tilde{\boldsymbol{\Phi}}$ are zero-mean and sub-exponential. For matrices $\boldsymbol{A}$ with zero-mean, random i.i.d. entries with bounded sub-exponential norm, Adamczak *et al.* showed that, provided $M = \mathcal{O}(s \ln^2(N/s))$, such matrices satisfy the RIP property with high probability when normalized by $\sqrt{M}$ [1]. This means that there exists a constant $0 < \delta < 1$ such that, for all s-sparse vectors $\boldsymbol{u}$ in $\mathbb{R}^N$—with at most s non-zero components—we have

$$(1-\delta)\|\boldsymbol{u}\|^2 \leq \|\tfrac{1}{\sqrt{M}}\boldsymbol{A}\boldsymbol{u}\|^2 \leq (1+\delta)\|\boldsymbol{u}\|^2.$$

Respecting the RIP implies that robust reconstruction of any s-sparse vector $\boldsymbol{x}$ from its (possibly noisy) observations $\boldsymbol{A}\boldsymbol{x}$ can be achieved via, *e.g.*, ℓ_1-minimization or greedy methods [53]. Foucart *et al.* showed that such matrices also satisfy a modified version of the RIP based on ℓ_1-norm [28] and propose a recovery algorithm able to recover s-sparse vectors with M in $\mathcal{O}(s \ln(N/s))$.

Unfortunately, the analysis of the autocorrelation of $\tilde{R}$ made in subsection 2.4 shows that all entries of $\tilde{\boldsymbol{\Phi}}$ are (locally) correlated. This is no surprise; while each row $\tilde{\boldsymbol{r}}_k$ of $\sqrt{M}\tilde{\boldsymbol{\Phi}}$ belongs to $\mathbb{R}^N$, only J parameters—the J random phases of the complex amplitudes $\boldsymbol{\alpha}_k$—were used to generate the pattern $\tilde{\boldsymbol{r}}_k$. Thus, provided that $N \geq J$ (*e.g.*, if $J = \mathcal{O}(100)$ and $N = \mathcal{O}(128^2)$ as in subsection 4.1.1), spatial correlations are inherent in each $\tilde{\boldsymbol{r}}_k$. Therefore, we cannot readily use the results of [1, 28] to characterize our sensing scheme. Note that the existence of the above-mentioned correlations is not an impossibility per se; after all, random partial Fourier sensing matrices, made of M randomly sampled rows of a Fourier matrix, do present such correlations but respect the RIP under certain conditions [15]. Studying if our sensing matrix $\tilde{\boldsymbol{\Phi}}$ satisfies the RIP is thus an open question. However, the *coherence* of $\tilde{\boldsymbol{\Phi}}$ also provides (weaker but computable) guarantees for the recovery of s-sparse signals [24, 29]. This quantity (defined hereafter) measures our ability to distinguish two distinct atoms of the sensing basis in the signal measurements; the lower the coherence the better is the signal recovery (as a limit case, an orthonormal basis has zero coherence and all signals can be recovered).

Definition 4.1. *(from [24, Def. 2]) The coherence $\mu(\boldsymbol{\Theta})$ of a matrix $\boldsymbol{\Theta}$ is the largest absolute inner product between any two columns $\boldsymbol{\theta}_i, \boldsymbol{\theta}_j$ of $\boldsymbol{\Theta}$:*

$$\mu(\boldsymbol{\Theta}) = \max_{1 \le i \ne j \le N} \frac{|\langle \boldsymbol{\theta}_i, \boldsymbol{\theta}_j \rangle|}{\|\boldsymbol{\theta}_i\| \|\boldsymbol{\theta}_j\|}.$$

The analysis provided by the coherence is quite limited compared to an analysis based on the RIP. However, we can still use this information to compare different sensing matrices. A low coherence is a good indicator of the efficiency of the sensing matrix in a CS framework.

There also exists a limit to the resolution achievable by our sensing model. Under the far-field approximation, (2.9) shows that

$$\tilde{R}(\boldsymbol{x}; \boldsymbol{\alpha}) \approx \tilde{R}_{\text{ff}}(\boldsymbol{x}; \boldsymbol{\alpha}) \coloneqq \tfrac{1}{J} \textstyle\sum_{j,k=1}^{J} \alpha_j \alpha_k^* e^{\frac{2\pi \mathrm{i}}{\lambda z}(\boldsymbol{q}_j - \boldsymbol{q}_k)^\top \boldsymbol{x}} - 1 = \mathcal{F}\big[\Theta(\cdot; \boldsymbol{\alpha})\big](-\tfrac{2\pi}{\lambda z}\boldsymbol{x}),$$

with $\Theta(\boldsymbol{u}; \boldsymbol{\alpha}) \coloneqq \frac{1}{J} \sum_{j,k=1}^{J} \alpha_j \alpha_k^* \delta(\boldsymbol{u} - (\boldsymbol{q}_j - \boldsymbol{q}_k)) - \delta(\boldsymbol{u})$. With $D = \max_{j,k} \|\boldsymbol{q}_j - \boldsymbol{q}_k\|$ being the smallest length such that $\Theta(\boldsymbol{u}; \boldsymbol{\alpha}) = \Theta(\boldsymbol{u}; \boldsymbol{\alpha}) \cdot \mathrm{disk}(\frac{\boldsymbol{u}}{D})$, we find

$$\tilde{R}_{\text{ff}}(\boldsymbol{x}; \boldsymbol{\alpha}) = (\tilde{R}_{\text{ff}}(\cdot; \boldsymbol{\alpha}) * H)(\boldsymbol{x}), \quad \text{with } H(\boldsymbol{x}) = H(-\boldsymbol{x}) = (\tfrac{2\pi D}{\lambda z})^2 \mathcal{F}[\mathrm{disk}](\tfrac{2\pi D}{\lambda z}\boldsymbol{x}).$$

Therefore, for any row $\tilde{\boldsymbol{r}}_k$ of $\sqrt{M}\tilde{\boldsymbol{\Phi}}$, which discretizes $\tilde{R}(\boldsymbol{x}; \boldsymbol{\alpha}_k) \approx \tilde{R}_{\text{ff}}(\boldsymbol{x}; \boldsymbol{\alpha}_k)$, there exists a symetric filter $\boldsymbol{h}$ (the discretization of H) whose size scales like $\mathcal{O}(\lambda z / D)$ and such that $\tilde{\boldsymbol{r}}_k = \tilde{\boldsymbol{r}}_k * \boldsymbol{h}$. In other words, as also stressed in subsection 2.4, the patterns $\tilde{\boldsymbol{r}}_k$ cannot have faster variations than those allowed by the spectrum of $\boldsymbol{h}$; the components of this pattern are correlated and display "speckle grains" whose size are related to the size of $\boldsymbol{h}$.

The existence of $\boldsymbol{h}$ alters the model (4.5); since $\langle \boldsymbol{a} * \boldsymbol{h}, \boldsymbol{b} \rangle = \langle \boldsymbol{a}, \boldsymbol{b} * \boldsymbol{h} \rangle$ for any vector $\boldsymbol{a}$ and $\boldsymbol{b}$, we find

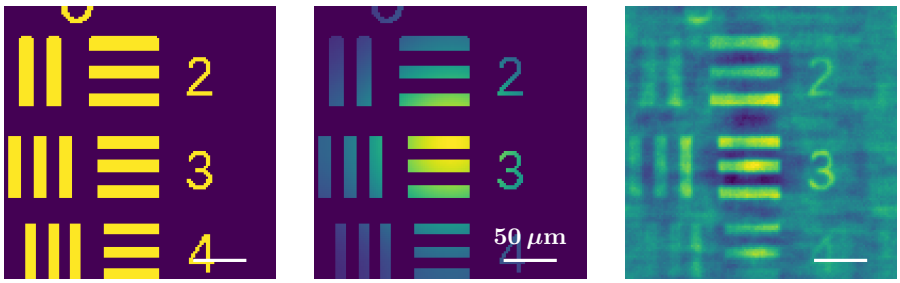
$$\tilde{\boldsymbol{y}} = \sqrt{M}\tilde{\boldsymbol{\Phi}}(\boldsymbol{h} * \bar{\boldsymbol{S}}\boldsymbol{f}) + \tilde{\boldsymbol{n}}. \tag{4.6}$$

In conclusion, even if $\tilde{\boldsymbol{\Phi}}$ in this last model was a random matrix with i.i.d. sub-exponential entries, (4.6) shows that, at best, we can estimate $\boldsymbol{h} * \bar{\boldsymbol{S}}\boldsymbol{f}$—a version of $\bar{\boldsymbol{S}}\boldsymbol{f}$ that is blurred by $\boldsymbol{h}$. Since the size of this filter is controlled by $\lambda z / D$, the impact of $\boldsymbol{h}$ can be mitigated by either decreasing distance z between the distal end of the fiber and the object (while still keeping the far-field regime valid) or increasing D when designing the MCF. Model (4.6) also shows that the pixel pitch ϖ used to discretize f can be adjusted to the speckle grain $\lambda z / D$ since a higher resolution cannot be achieved. We will follow this procedure in subsection 5.2 for actual LE fluorescence imaging.

In the following sections, we use the physical model (3.7) for data simulations. For the reconstruction process, we consider debiased observations $\tilde{\boldsymbol{y}}$ in (4.5), the associated sensing matrix $\tilde{\boldsymbol{\Phi}}$, and we adapt the CV optimization method (3.11) of subsection 3.4.4 into

$$\hat{\boldsymbol{f}}_\rho \in \arg\min_{\boldsymbol{u}} \ \|\boldsymbol{R}_{\text{est}}(\sqrt{M}\tilde{\boldsymbol{\Phi}}\bar{\boldsymbol{S}}\boldsymbol{u} - \tilde{\boldsymbol{y}})\|^2 + \rho g_1(\boldsymbol{u}) + \imath_{\mathbb{R}_+^N}(\boldsymbol{u}), \tag{4.7}$$

with quality criterion $\tilde{q}(\hat{\boldsymbol{f}}_\rho) \coloneqq \|\boldsymbol{R}_{\text{val}}(\sqrt{M}\boldsymbol{\Phi}\bar{\boldsymbol{S}}\hat{\boldsymbol{f}}_\rho - \tilde{\boldsymbol{y}})\|^2$. As will be clear below, the inclusion of $\bar{\boldsymbol{S}}$ in this minimization—induced by inserting (4.5) in (3.11)—allows us to expand the area where $\boldsymbol{f}$ is estimated outside of the support of $\bar{\boldsymbol{S}}$.



(a) (b) (c)

Figure 5: (a) 128×128 ground truth $\boldsymbol{f}$ of standard USAF transmission target. (b) Vignetted ground truth $\bar{\boldsymbol{S}}\boldsymbol{f}$. (c) Acquisition with RS strategy as explained in subsection 3.4. Observations were generated according to (3.7) with focused illumination pattern and AWGN (SNR of $\boldsymbol{\Phi}\boldsymbol{f}$ equal to 40 dB). The maximum intensity of the RS image is around three times lower than the one of the ground truth.

As explained in subsection 4.1.1, this optimization scheme leads to a quality of the final estimate similar to the one obtained with the biased sensing model but achieved in a much shorter reconstruction time.

4.1.1. Simulations. We simulate observations $\boldsymbol{y}$ of 128×128 USAF transmission target $\boldsymbol{f}$ (see Figure 5a) according to (3.7) for $M/N \in \{0.1, 0.2, \dots, 1\}$. We reconstruct estimates of $\boldsymbol{f}$ by first solving the original inverse problem (3.11) with $(\boldsymbol{\Phi}, \boldsymbol{y})$ and then the debiased model (4.7) with $(\tilde{\boldsymbol{\Phi}}, \tilde{\boldsymbol{y}})$. For comparison, we also simulate RS observations and measurements acquired with an ideal zero-mean Gaussian matrix $\boldsymbol{\Phi}_{\mathrm{G}}$. Since the synthetic USAF transmission target is piecewise constant, we select the TV-norm for g_1 in (3.11) and (4.7).

We measure the quality of an estimate $\hat{\boldsymbol{u}}$ with both the signal-to-noise ratio (SNR) and the weighted SNR metrics:

$$\mathrm{SNR}(\hat{\boldsymbol{u}}, \boldsymbol{u}) \coloneqq 20 \log_{10} \left(\frac{\|\hat{\boldsymbol{u}}\|}{\|\hat{\boldsymbol{u}} - \boldsymbol{u}\|} \right), \quad \mathrm{WSNR}(\hat{\boldsymbol{u}}, \boldsymbol{u}) \coloneqq \mathrm{SNR}(\bar{\boldsymbol{S}}\hat{\boldsymbol{u}}, \bar{\boldsymbol{S}}\boldsymbol{u}).$$

The WSNR attenuates the reconstruction artefacts at the limit of the FOV where the vignetting of $\bar{\boldsymbol{S}}$ is the strongest.

For each value of M above, M observations of $\boldsymbol{f}$ are generated according to model (3.7), with a noise variance σ^2 set such that the SNR of $\boldsymbol{\Phi}\boldsymbol{f}$ is equal to 40 dB. For all the simulations, the illumination patterns were generated for a Fermat's spiral core arrangement ($J = 120$ cores and diameter $D = 113\,\mu$m) and the following parameters: $\lambda = 1\,\mu$m, $z = 500\,\mu$m, $3\sigma_c = 3.2\,\mu$m and pixel pitch $\varpi = 2\,\mu$m.

We perform the reconstruction with the CP algorithm (see subsection 3.4.3) and a maximum number of internal iterations equal to 5 000, never reached in practice. The stopping criterion and regularization parameter are set as described in subsection 3.4.4 with $M_{\mathrm{val}} = 256$, and initial values $\rho^{(0)} = 10$ and $\hat{\boldsymbol{f}}_{\rho^{(0)}} = \boldsymbol{0}_N$. The initial guess of the algorithm for each $k > 0$ is given by previous estimate $\hat{\boldsymbol{f}}_{\rho^{(k-1)}}$ and $\rho^{(k)} = 0.5\rho^{(k-1)}$. The maximum number of iterations on the estimation of ρ is 20.

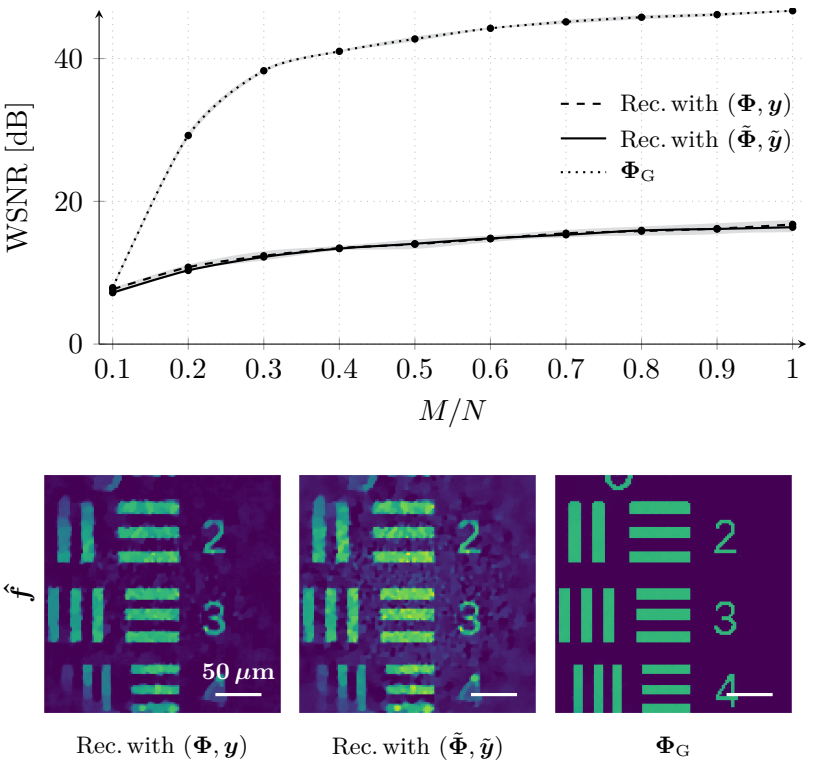


Figure 6: SI strategies. Above: mean WSNR (over 10 trials) of the restored USAF target *versus* M/N for three SI strategies: (*i*) acquisition and reconstruction with $(\boldsymbol{\Phi}, \boldsymbol{y})$ (dashed line), (*ii*) acquisition with $\boldsymbol{\Phi}$ and reconstruction with $(\tilde{\boldsymbol{\Phi}}, \tilde{\boldsymbol{y}})$ (solid line), and (*iii*) acquisition and reconstruction with ideal (unrealistic) sensing matrix $\boldsymbol{\Phi}_{\mathrm{G}}$ (dotted line). The gray areas represent the standard deviations. Synthetic observations were generated according to (3.7) (SNR of $\boldsymbol{\Phi}\boldsymbol{f}$ equal to 40 dB). Below: estimates of $\boldsymbol{f}$ obtained for $M/N = 0.3$. All estimates share the same intensity scale.

4.1.2. Comparison results. Simulation results are visible on Figure 6. For comparison purpose, we also show the estimate obtained with RS strategy (see Figure 5c). Even if they require to store M speckles patterns and to perform a costly reconstruction, estimates of the fluorophore density map obtained with SI methods have a better quality compared to the estimate obtained with RS (WSNR equal to 1.68 dB with optimal normalization, see Figure 5c). The poor quality of the RS estimate is indeed mainly due to the imperfections[3] of the focused PSF (see Figure 2a) that are not corrected during the trivial reconstruction described in (3.8).

Acquisition and reconstruction with a Gaussian sensing matrix outperforms the other two frameworks involving speckle illumination. While unrealistic, such a Gaussian framework is useful as it sets an upper bound on the achievable reconstruction quality of LE imaging. For instance, it shows that the CS regime, where enough measurements are collected to ensure a reconstruction quality close to the measurement SNR of 40 dB, starts around $M/N = 0.3$ after a sharp transition in the curve. After that value, an increase in the number of measurements

[3]For instance, the central peak represents only 2% of the total energy of the focused beam [3, 4].

does not lead to a significant gain in WSNR. The same behavior is observed for reconstructions based on speckle acquisition even if the change in the curve rate is less outstanding. WSNR of images reconstructed with $(\tilde{\boldsymbol{\Phi}}, \tilde{\boldsymbol{y}})$ is similar to the one obtained for a reconstruction performed with $(\boldsymbol{\Phi}, \boldsymbol{y})$, despite the extra noise introduced in the debiased model. However, we observed significantly shorter reconstruction times for the debiased model when $M/N \leq 0.7$. We also noticed that, while the reconstruction with $(\boldsymbol{\Phi}, \boldsymbol{y})$ fails when we do not constraint the estimate to be non negative, the debiased model performs almost as well without enforcing the positivity of the map. In the rest fo the paper, all reconstructions are performed with the debiased model.

4.2. Partial Speckle Scanning. One disadvantage of SI is that, despite the good quality of its reconstruction, it requires long acquisition and reconstruction times compared to RS imaging that simply uses scanning mirrors and direct reconstruction (see (3.8)). This is due to the strong contrast that exists between the time it takes to change the SLM configuration in SI—about 1/60 s in our setup [36, see DVI frame rate]—and the elapsed time between two consecutive tilts of the galvanometric mirrors, which is about 1,000 times faster [14].

We here propose a hybrid acquisition framework, called *partial speckle scanning* (PSS), combining the advantages of both techniques while keeping a high reconstruction quality. This method strongly relies on two properties of the LE considered in this paper: (*i*) the ability to easily generate speckles by randomly programming the SLM and (*ii*) the MCF memory effect.

The PSS strategy acquires M observations but unlike the SI strategy, a single SLM configuration allows us to collect $M_P \leq M$ observations by translating the speckle. Following subsection 2.2, this is achieved by applying different tips to the input wave front to the MCF with the scan mirrors. Mathematically, given a set of $P = M/M_P$ complex amplitudes $\{\boldsymbol{\alpha}_j\}_{j=1}^{P}$ randomly generated as in the SI model, and M_P mirror tilts $\{\boldsymbol{\theta}_k\}_{k=1}^{M_P}$, the PSS sensing matrix $\tilde{\boldsymbol{\Phi}}$ used in the model (4.5) corresponds to

$$\tilde{\boldsymbol{\Phi}} = [\tilde{\boldsymbol{\Phi}}_1^\top, \cdots, \tilde{\boldsymbol{\Phi}}_P^\top]^\top, \quad \text{with } \tilde{\boldsymbol{\Phi}}_j^\top = [\tilde{\boldsymbol{r}}_j^{(0)}, \tilde{\boldsymbol{r}}_j^{(1)}, \cdots, \tilde{\boldsymbol{r}}_j^{(M_P-1)}]^\top. \tag{4.8}$$

and where, according to (2.11), $\tilde{\boldsymbol{r}}_j^{(k)}$ is the discrete representation of $\tilde{R}(\boldsymbol{x}, \mathrm{diag}(\boldsymbol{\gamma}(\boldsymbol{\theta}_k))\boldsymbol{\alpha}_i) \approx \tilde{R}(\boldsymbol{x} + \boldsymbol{\theta}_k, \boldsymbol{\alpha}_i)$—the shifted residual field—for each $j \in [P]$, $k \in [M_P]$, and $\boldsymbol{\gamma}$ defined in (2.10).

In this work, we consider a single line scanning mode, *i.e.*, shifts are applied to the speckle patterns in only one (abitrary) unit direction $\boldsymbol{u} \in \mathbb{R}^2$ and $\boldsymbol{\theta}_k = k\delta\,\boldsymbol{u}$ for some translation step $\delta > 0$ of the speckle patterns in the plane $\mathcal{Z}$. Line scanning is fast and accurate because it only needs the rotation of one galvanometric mirror.

When designing the PSS framework, the adjustment of the shift δ between two illumination patterns (which we perform in the next section) faces two competing effects. First, if δ is smaller than a speckle grain, *i.e.*, if too small compared to $\lambda z/D$, the shifting model (4.8) introduces too much correlations between neighbouring rows of $\tilde{\boldsymbol{\Phi}}$ and there is not much variety in the M_P observations acquired with a fixed SLM configuration. Moreover, in addition to the column dependency mentioned in subsection 4.1, our sensing matrix further deviates from an ideal decorrelated sub-exponential random matrix. Second, to approximate $\tilde{\boldsymbol{\Phi}}$ by a block-circulant matrix (*e.g.*, to boost the computation of matrix-vector multiplication), δ must be small enough such that approximation (2.12) still holds. This imposes us to respect the

paraxial approximation for all translation steps, *i.e.*, we must have $M_P\delta \ll z$. Note that, while there exist sensing constructions based on subsampled random circulant matrices defined from sub-Gaussian random filters [24, 53], there are no known constructions for sub-exponential random filters. Despite this absence, our simulations below confirm that the PSS sensing compares favorably to the SI and the Gaussian sensing schemes.

An advantage of the PSS scheme is that its scanning time t_{acq} is reduced compared to the SI acquisition time; a desirable advantage when observing fluorescent biological samples subject to photo-bleaching (see subsection 3.1). We can compute this time from

$$t_{\text{acq}} = t_{\text{galva}}(\xi P + M - P), \quad \text{with } M = PM_P, \tag{4.9}$$

where $1/t_{\text{galva}}$ is the scan mirror rate and $\xi := t_{\text{SLM}}/t_{\text{galva}}$ with $1/t_{\text{SLM}}$ the SLM frame rate. If t_{SLM} is large compared to t_{galva}, *i.e.*, if ξ is large, for a fixed M, t_{acq} can be kept small if we keep the number of distinct speckle patterns P small. In our setup, $\xi \approx 10^3$.

4.2.1. Simulations. As a first simulation, we aim to choose a convenient shift δ between two consecutive speckles. We simulate M observations $\boldsymbol{y}$ of 128×128 USAF transmission target $\boldsymbol{f}$ for $M/N = 0.3$. $M_P = 2$ observations are acquired before changing the configuration of the SLM, *i.e.*, $P/N = 0.15$. We sequentially set the shift δ to the values $\{0, 1, 2, 3, 4, 6, 8, 10\}$ [μm]. For each, we reconstruct the estimate $\hat{\boldsymbol{f}}$ of $\boldsymbol{f}$ by solving the debiased model (4.7) with $(\tilde{\boldsymbol{\Phi}}, \tilde{\boldsymbol{y}})$. Exactly like for the SI strategy, g_1 in (4.7) is the TV-norm.

We also conduct a second simulation where we compare the PSS strategies for different values of M_P. We simulate and reconstruct observations as described in the previous paragraph but with fixed value of δ (chosen according to the result of the first experiment), for $M/N \in \{0.1, 0.2, \ldots, 1\}$ and $M_P \in \{1, 2, 4, 8, 16, 32, 64\}$. The value $M_P = 1$ corresponds to the SI strategy. We generate the observations according to (3.7) with σ^2 set to reach a measurement SNR of 40 dB. The speckles are generated for a Fermat's spiral core arrangement ($J = 120$ cores and diameter $D = 113\,\mu$m) with the following parameters: $\lambda = 1\,\mu$m, $z = 500\,\mu$m, $3\sigma_c = 3.2\,\mu$m and a pixel pitch $\varpi = 2\,\mu$m. The time t_{SLM} is set to 100 ms, slightly higher than the actual SLM time [36], and t_{galva} is set to $100\,\mu$s [14]. The reconstruction is performed as already described in the SI simulations (see subsection 4.1.1).

4.2.2. Results. Results in Figure 7 suggest to consider a shift $\delta \geq 4\,\mu$m (corresponding to $\delta/\lambda \geq 4$). In this case, the theoretical autocorrelation of field $\tilde{R}$ is close to zero and the coherence of the sensing matrix is lower (see Definition 4.1) than the ones for $\delta < 4\,\mu$m. The quality of the estimate is similar to the quality obtained with the SI strategy.

Figure 8 shows the results of the second experiment performed with $\delta = 4\,\mu$m. We observe a reconstruction quality similar for all M_P values at fixed M/N ratio. It indicates that the choice of δ based on Figure 7 is good: for $\delta = 4\,\mu$m, the shifted patterns get sufficiently decorrelated such that the quality of the estimates for $M_P > 1$ is similar to the one obtained with SI ($M_P = 1$). However, we are more interested in considering the quality obtained for a fixed acquisition time t_{acq} rather than a fixed number of measurements M. Images shown in Figure 8 correspond to (M, M_P) pairs located on the same t_{acq} level curve (see the upper plot). For a fixed acquisition time (or photobudget), choosing M_P as large as possible is the best option. To minimize the photo-bleaching of the sample and increase the time resolution of the acquisition, we would like to choose t_{acq} as small as possible while keeping a high WSNR. In

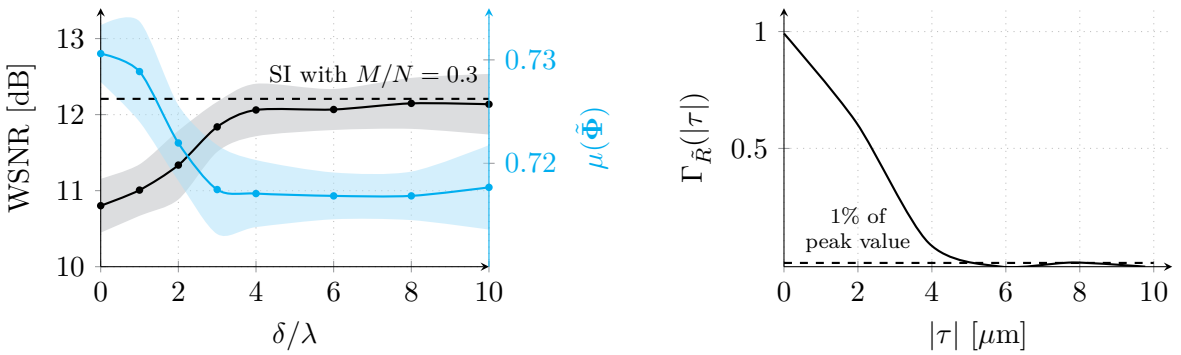


Figure 7: PSS strategy: choice of δ. (Left) Mean WSNR (over 25 trials) of the restored USAF target (black) and mean coherence (over 25 trials) of the sensing matrix $\tilde{\mathbf{\Phi}}$ (cyan) *versus* the relative shift δ/λ between two consecutive replicas of the same speckle. The light area represents the standard deviation. Synthetic observations were generated according to (3.7) (SNR of $\mathbf{\Phi} \boldsymbol{f}$ equal to 40 dB). Each illumination pattern was used twice and $M/N = 0.3$. (Right) Theoretical autocorrelation of the residual field $\tilde{R}$ defined in (2.16) as a function of $|\tau|$ (extracted from Figure 3).

this case, the best choice is the PSS strategy with $M_P = 64$: we achieve high WSNR (around 17 dB) with an acquisition time around 30 s.

Remark 4.2 (Approximation of the sensing matrix). Up to now, the proposed PSS strategy only decreases the acquisition time. Compared to SI for the same number of measurements, the number of speckles to be recorded to form $\tilde{\mathbf{\Phi}}$ is identical. Moreover, the reconstruction time is similar between SI and PSS since the complexity (in $\mathcal{O}(MN)$) of the matrix-vector product involving $\tilde{\mathbf{\Phi}}$ is not optimized. By considering the translation rule (2.12), this can be potentially improved by approximating $\tilde{\mathbf{\Phi}}$ with a block-Toeplitz matrix $\tilde{\mathbf{\Phi}}_{\text{app}}$. Then, each matrix-vector product involving this new matrix can benefit from the FFT and the complexity is then reduced to $\mathcal{O}(N \log N)$. The previous storage of M speckle patterns is reduced to P patterns. We test this matrix approximation in Figure 9. This figure is the same as Figure 8 but for a reconstruction performed with $\tilde{\mathbf{\Phi}}_{\text{app}}$. As expected, when M_P increases, $\|\boldsymbol{\theta}_{M_P}\| = M_P \delta$ takes bigger values and the paraxial approximation is less respected. The error made by approximation (2.12) increases. This leads to a decrease in the reconstruction quality. However, for $M_P \leq 16$, the reconstruction quality is still similar to the one of the SI strategy (around 15 dB). Therefore, in this approximation of $\tilde{\mathbf{\Phi}}$ by $\tilde{\mathbf{\Phi}}_{\text{app}}$, one must find a trade-off between a fast acquisition time and the quality of the produced estimate. Unlike SI, PSS has thus the potential to (*i*) reduce the pattern storage and (*ii*) speed up the reconstruction by exploiting the structure of the sensing matrix. Fast reconstruction algorithms are crucial for *in vivo* imaging or to reach higher image resolution. For now, those two advantages are more a perspective than a proof of concept since deviation to pure speckle translation (given by (2.12)) currently prevents us to boost the reconstruction. We postpone this analysis for a future study, and we only consider the original matrix $\tilde{\mathbf{\Phi}}$ in our following experiments.

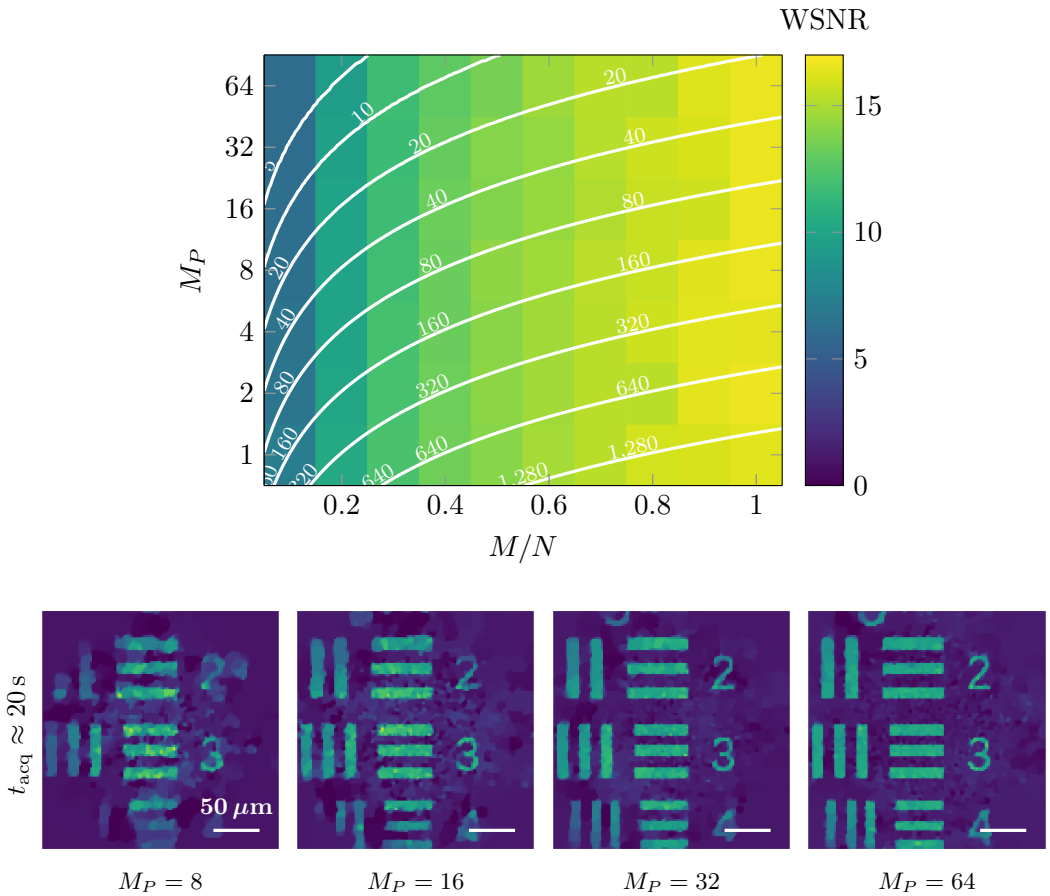


Figure 8: PSS strategy. Above: mean WSNR (over 10 trials) of the restored USAF target *versus* M_P and M/N ratio. Synthetic observations were generated according to (3.7) with AWGN (SNR of $\boldsymbol{\Phi f}$ equal to 40 dB). Number M_P of replicas of the same (shifted) speckle pattern belongs to $\{1, 2, 4, 8, 16, 32, 64\}$ and shift δ between replicas is 4 μm. Level curves of the acquisition time t_{acq} (in seconds) are superimposed to the SNR (white solid lines). Below: estimates of $\boldsymbol{f}$ obtained for $M_P \in \{8, 16, 32, 64\}$ and $t_{\text{acq}} \approx 20\,\text{s}$. The PSS strategy with $M_P = 1$ corresponds to the SI strategy with $(\tilde{\boldsymbol{\Phi}}, \tilde{\boldsymbol{y}})$ (see Figure 6). Estimates share the same intensity scale.

5. Fluorescence imaging experiments. In this last section, we apply the SI and PSS methods to an actual lensless endoscope in the context of fluorescence imaging. We first describe the experimental setup, material and methods, before explaining how these sensing strategies can improve the quality of the reconstructed images compared to the RS technique.

5.1. Experimental setup. A simplified view of the experimental setup is presented in Figure 10 and described below in functional blocks.

Spatial light modulator. A continous wave laser operating at 491 nm (Cobolt lasers, Sweden) is expanded and impinges upon a liquid crystal SLM (X10468-03, Hamamatsu, Japan). A set of relay optics, depicted by a single lens and objectives *Obj* in Figure 10, images the SLM

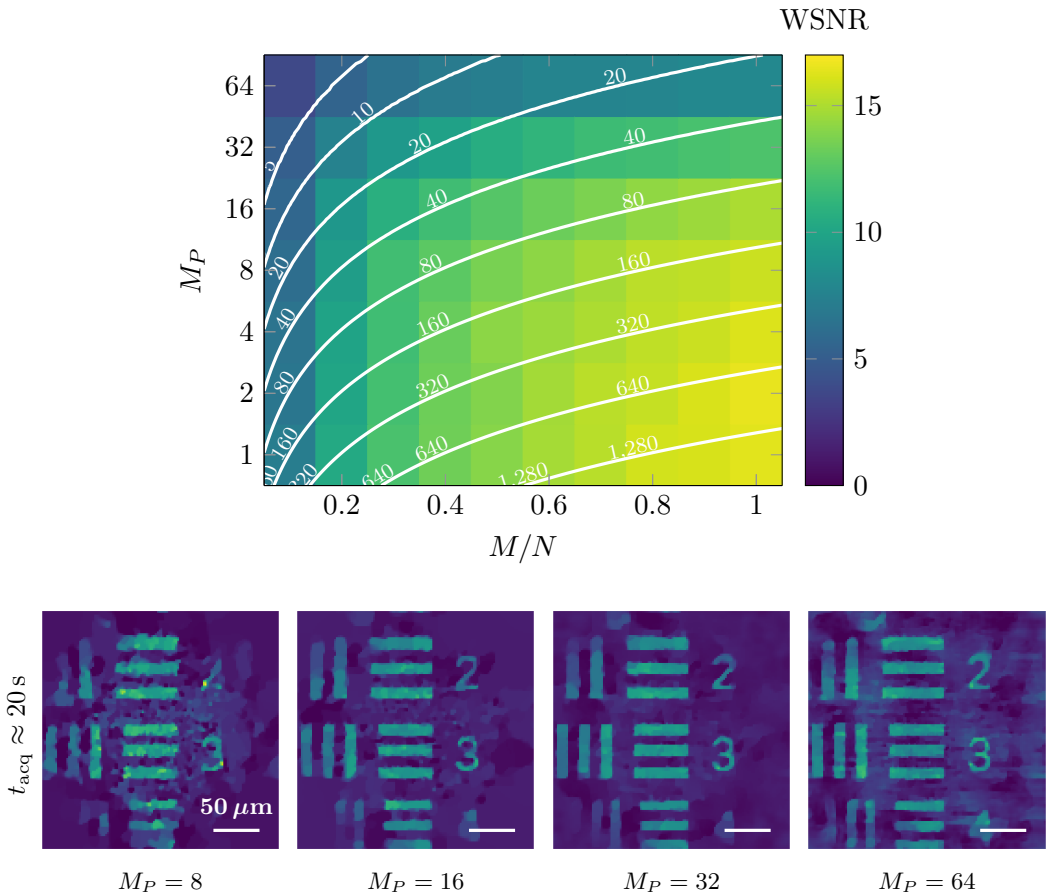


Figure 9: PSS approximated strategy. Above: mean WSNR (over 10 trials) of the restored USAF target *versus* M_P and M/N ratio. Synthetic observations were generated according to (3.7) with AWGN (SNR of $\boldsymbol{\Phi f}$ equal to 40 dB). Number M_P of replicas of the same (shifted) speckle pattern belongs to $\{1, 2, 4, 8, 16, 32, 64\}$ and shift δ between replicas is $4\,\mu$m. Sensing matrix $\tilde{\boldsymbol{\Phi}}$ was approximated by a block-circulant operator (instead of a block-Toeplitz operator since the FOV is limited) for the reconstruction. Level curves of the acquisition time t_{acq} (in seconds) are superimposed to the SNR (white solid lines). Below: estimates of $\boldsymbol{f}$ obtained for $M_P \in \{8, 16, 32, 64\}$ and $t_{\text{acq}} \approx 20\,\text{s}$. The PSS strategy with $M_P = 1$ corresponds to the SI strategy with $(\tilde{\boldsymbol{\Phi}}, \tilde{\boldsymbol{y}})$ (see Figure 6).

to the proximal endface of the MCF. In order to maximize the injection of the light beams into the individual cores, a convex lenslet array whose centers are matched to the individual cores is displayed on the SLM. This results in a beamlet array which is efficiently coupled into the MCF and whose relative phases can be tuned independently with the SLM.

Multicore fiber. Imaging through the MCF is analogous to phased arrays for beamforming where the relative phases between the cores (antennae) can be calibrated and tuned to generate and shift a focused beam. The cores of these fibers are single mode at the operating wavelength and exhibit an inter-core coupling term less than 20 dB [60]. These factors have two important advantages: (*i*) operations such as RS or defocusing can be performed with conventional

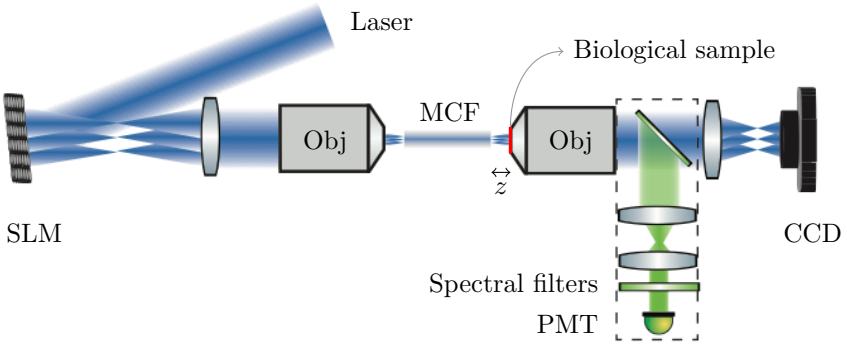


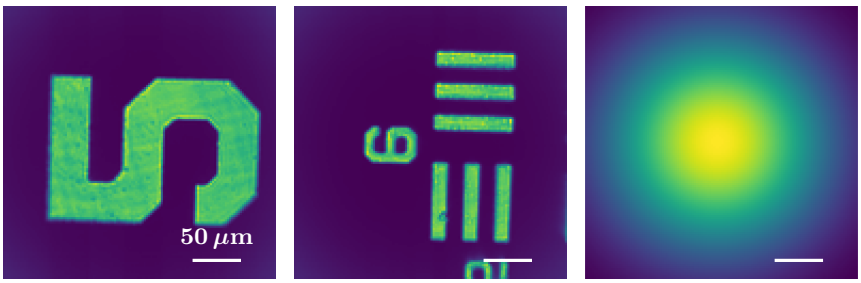
Figure 10: Simplified view of the experimental setup used to characterize the performance of the CS-based acquisition strategies. Relay optics (Obj) image the spatial light modulator (SLM) to the proximal endface of the multicore fiber (MCF). An imaging system with two channels is placed at a distance z of the MCF distal end. The main purpose of the first channel is to image the speckle patterns with a charge-coupled device (CCD) sensor. The second channel is dedicated to observations collection through a single pixel detector or photomultiplier tube (PMT).

optical elements at the distal end, and (*ii*) the resulting speckle patterns are highly resilient to external perturbations both thermal and mechanical (except for a global shift). These are significant advantages over single multimode fiber (MMF) where the significant off-diagonal coupling terms preclude stability and fast imaging with conventional optical elements. For a more comprehensive discussion of the imaging properties of the golden spiral MCFs, we refer the reader to [60].

Generation of the speckle patterns. Multiple illumination patterns are generated a few hundred microns away from the distal end as a combination of (*i*) randomizing the relative phase of the injected beamlets into the MCF resulting in a speckle pattern at the distal end of the fiber, and (*ii*) translations of the speckle pattern with a global tilt of the beamlets.

Calibration and imaging. The distal end of the MCF is placed at the focal plane of a second imaging system with two channels: one imaging the distal end onto the camera (CCD) and a second one detecting the signal of interest (PMT). The first channel serves for the recording of the sensing matrix, the visual inspection of the samples and the generation of an image close to the ground truth. The sensing matrix is populated by acquiring a multi-exposure image of each speckle pattern and fusing them to generate a synthetic high-dynamic range. The second channel employs a single pixel detector (PMT) (H7240-50, Hamamatsu, Japan) upon which the signal is detected. In the following experiments, the sample is either a standard USAF transmission target or fluorescent beads. In either case, we operate in a high photon count regime as assumed in subsection 3.2.

In the interest of maximum flexibility, we employ the SLM to function as a series of optical components such as a microlens array to maximize the coupling into the fiber, and as a galvanometric scanner. This also allows us to test and compare between the proposed imaging scheme and conventional RS techniques. However, for the speckle illumination based compressive imaging proposed in the paper, the SLM can be replaced with conventional optical



USAF #1 USAF #2 Mean speckle field

Figure 11: Estimates of USAF ground truth. Images of standard USAF transmission target. They are estimated using the first channel of the imaging system (see Figure 10). The CCD sensor acquires M images of the product between a speckle and density map f. Their average divided by the mean speckle field (Gaussian fit estimated from the M light patterns) leads to the estimated ground truths.

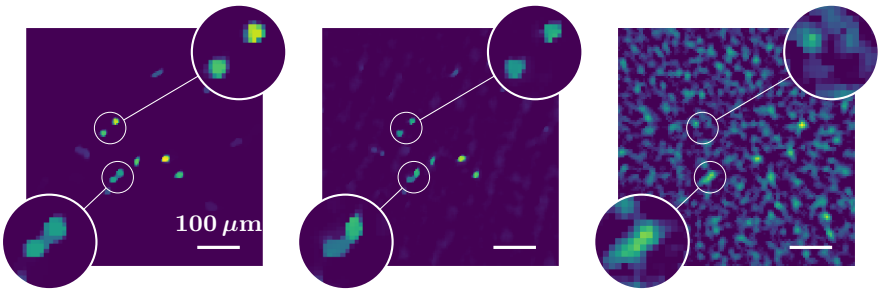
elements such as microlens arrays, mirror scanners and thin diffusers as in our earlier works [3].

5.2. Material and methods. Samples used for the experiments are either standard USAF transmission target (see Figure 11) or 5 μm fluorescent Beads. USAF sample itself is not fluorescent: it is a layer of metal deposited on glass where the features (the bars) are transparent. To make it fluorescent, we apply a layer of highlighter on top of it. We use the setup depicted in Figure 10 and described in subsection 5.1 with a Fermat's golden spiral MCF containing $J = 118$ cores and with standard deviation $\sigma_c = 0.8\,\mu$m (or, equivalently, $d = 1.9\,\mu$m). The distance between the endface of the fiber and the sample plane is $z = 500\,\mu$m.

5.2.1. Acquisition and reconstruction. In all experiments, the data acquisition follows those two steps: (*i*) recording the speckle patterns with the CCD sensor to build and store the sensing matrix $\mathbf{\Phi}$ (pixel size equal to 2.2 μm), and (*ii*) illuminating the sample with M speckles and measuring the signal on the single pixel detector. Light patterns are either all different from each other (SI strategy) or shifted versions of each other (PSS strategies) due to the application of global tilt terms on the SLM ($\delta = 1.1\,\mu$m).

The first experiment was designed to compare the SI and PSS strategies: $M = 4096$ measurements of Beads #3 sample were acquired for $M_P = 1$ (SI) and then $M_P = 64$ (PSS). The second experiment acquired $M = 4096$ observations of two different parts of the USAF target (see Figure 11) with $M_P = 64$. Finally, $M = 4096$ observations of three other Beads samples were acquired with $M_P = 64$.

Unlike the number of observations M that is an acquisition parameter, the number of pixels N of the estimate (or equivalently, the pixel pitch ϖ) can be chosen after the acquisition process. Ideally, ϖ should match the diffraction limited point spread function of the device, *i.e.*, the speckle grain size. In this case, we can take full advantage of the CS-friendly statistical properties of the sensing matrix and we avoid spurious correlations in the estimate. For the considered experimental parameters, the average grain size is $r \approx 3.5\,\mu$m (see subsection 2.4). We thus set the pixel pitch ϖ to r, and select $N = 80 \times 80$ for USAF samples and $N = 160 \times 160$



SI
$M = 4096$
$t_{\text{acq}} = 409.6\,\text{s}$

PSS
$M = 4096$
$t_{\text{acq}} = 6.8\,\text{s}$

SI
$M = 68$
$t_{\text{acq}} = 6.8\,\text{s}$

Figure 12: Vignetted estimates of `Beads #3` sample. Left and middle images are obtained with two different acquisition strategies (SI and PSS) but the same number of observations M ($\rho = 860$ and $\rho = 690$). Middle and right images are again obtained with two different strategies (PSS and SI) but the same acquisition time $t_{\text{acq}} = 6.8\,\text{s}$. Left and middle estimates share the same intensity scale. The intensity of the right estimate is 50 times lower compared to the other two images.

for `Beads` samples (original sizes are 128×128 and 256×256, respectively).

We obtained better results by solving a slightly different problem from (4.7): given the zero-mean sensing matrix $\tilde{\boldsymbol{\Phi}}$ and the debiased observations $\tilde{\boldsymbol{y}}$, we solve

$$\widehat{\bar{\boldsymbol{S}}\boldsymbol{f}}_\rho \in \arg\min_{\boldsymbol{u}} \|\boldsymbol{R}_{\text{est}}(\sqrt{M}\tilde{\boldsymbol{\Phi}}\boldsymbol{u} - \tilde{\boldsymbol{y}})\|^2 + \rho g_1(\boldsymbol{u}) + \imath_{\mathbb{R}_+^N}(\boldsymbol{u}), \tag{5.1}$$

i.e., we estimate the *vignetted* fluorophore density map $\bar{\boldsymbol{S}}\boldsymbol{f}$ instead of $\boldsymbol{f}$. This adaptation is possible by setting g_1 in (5.1) to the TGV^2_α-norm. This norm, which promotes piecewise linear images, is well adapted to our piecewise constant images multiplied by the smooth vignetting $\bar{\boldsymbol{s}}$. As in the synthetic case, we use the CP algorithm with a maximum number of iterations equal to 5000 (that was also never reached). The stopping criterion of CP internal iterations is set as described in subsection 3.4.4 with $M_{\text{val}} = 256$, while the parameter ρ is chosen by visual inspection. Initial values for ρ are $\rho^{(0)} = 2500$ (`USAF`) or $\rho^{(0)} = 10^4$ (`Beads`) and $\rho^{(k)} = 0.8\rho^{(k-1)}$ for $k > 0$ ($k_{\max} = 20$). Initial guess of the algorithm is $\hat{\boldsymbol{f}}_{\rho^{(0)}} = \boldsymbol{0}_N$ and then $\hat{\boldsymbol{f}}_{\rho^{(k-1)}}$ for $k > 0$.

5.3. Results. Reconstruction results are visible in Figures 12 to 14. Regarding Figure 12, as expected, the SI strategy provides a better estimate for `Beads #3` sample: beads are better resolved and there are less artifacts in the background. However, this quality is obtained at the cost of an acquisition time 60 times longer compared to the PSS strategy. Middle and right images are again obtained with two different strategies (PSS and SI) but the same acquisition time $t_{\text{acq}} = 6.8\,\text{s}$. In this case, the SI strategy fails to reconstruct the beads.

For `USAF` samples and `Beads #1`, `#2` and `#4`, we performed the image reconstruction from the observations acquired with PSS strategy for $M/N \in \{0.3, 0.5, 0.64\}$ (see Figures 13 and 14). We note that in order to minimize photo-bleaching, we keep the incident laser power extremely

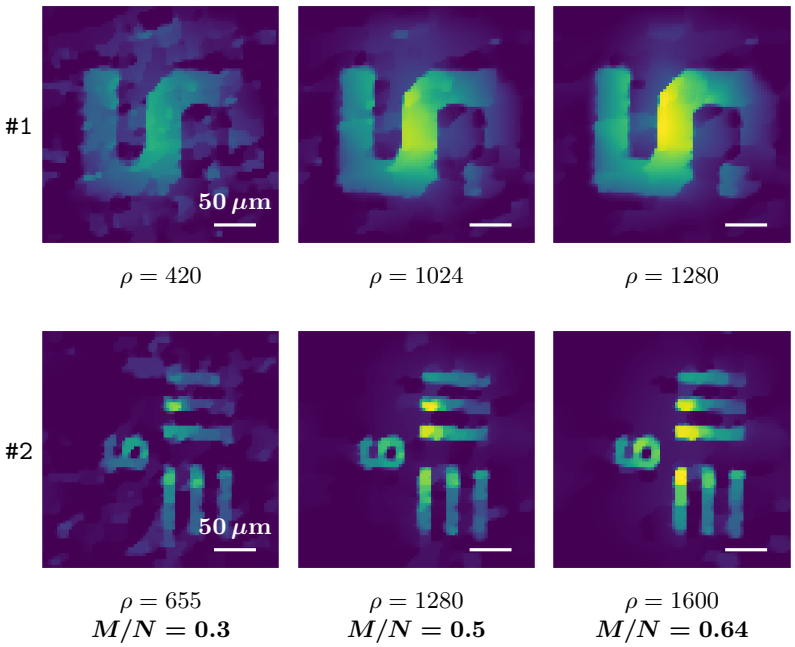


Figure 13: Vignetted estimates of USAF samples with PSS strategy. All 80×80 estimates are reconstructed from $M = 4096$ observations (corresponding to $M/N = 0.64$) acquired with PSS strategy ($M_P = 64$). Estimates for each sample share the same intensity scale.

low (few 100 s of μW over the entire FOV). This precludes us from acquiring ground truth fluorescence images of the beads since the sensitivity of a standard camera is much lower than the single pixel detector.

6. Conclusion. We have proposed new sensing methods to image fluorescent object in the context of lensless endoscopy. Our procedure departs from the classical raster scanning imaging by using two constructions leveraging speckle patterns to illuminate the sample, namely the unstructured speckle imaging (SI) and the partial speckle scanning (PSS). Our work also relies on proving that these speckles correspond to random fields following a sub-exponential distribution at each location of the sample plane (see subsection 2.5). After normalization and discretization, they are thus good candidates to build efficient sensing matrices, such as sub-exponential random matrices [1].

To use speckle illumination to collect observations, several challenges must be solved with regard to the design of the fiber, the acquisition strategy, as well as the reconstruction scheme.

First, the arrangement of the single mode cores must be optimized to achieve narrow autocorrelation of the speckle field (see subsection 2.4), *i.e.*, a grain size as small as possible, and to minimize the magnitude of the side lobes. Fermat's golden spiral shape shows a low side lobes level and a very good contrast between the intensities of the central peak and the side lobes [60]. A small grain size combined with an appropriate choice of the pixel pitch leads to a sensing matrix with nearly independent columns, a desirable property in CS.

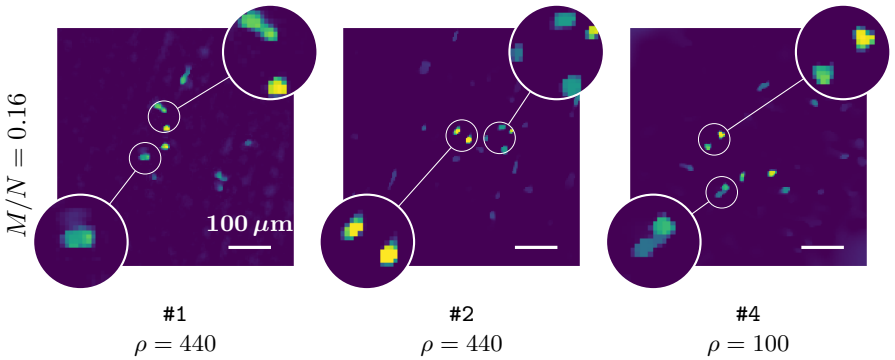


Figure 14: Vignetted estimates of `Beads` samples (`#1`, `#2` and `#4`). All 160×160 estimates are reconstructed from $M = 4096$ observations acquired with PSS strategy ($M_P = 64$). It corresponds to $M/N = 0.16$.

Second, the acquisition strategy must be thought to minimize the acquisition time to avoid as far as possible the loss of fluorescence of the sample and to reach a frame rate suitable for *in vivo* imaging of cellular processes (*e.g.*, the propagation of nerve impulses). The SI strategy offers good reconstruction quality for far fewer measurements compared to RS (see Figure 6). However, the corresponding acquisition time is unrealistic for real biological applications, as requiring to change speckle patterns using slow SLM (with a frame rate of the order of 10 ms). We overcome this limitation using the PSS strategy exploiting the memory effect of the fiber via the use of scan mirrors. Their rate is around 1,000 times higher compared to a change of the SLM configuration. With an appropriate value for the speckle shift δ (see Figure 7), we reach a reconstruction quality similar to that of SI (see Figure 8).

Finally, the reconstruction scheme, *i.e.*, the formulation and solving of the inverse problem, must encode prior information on the structure of the fluorophore density map. We proposed a debiased formulation including the minimization of TV or TGV^2_α-norm. We showed that the sensing matrix can be approximated by a block-circulant matrix to reduce the number of stored patterns. When $M_P \leq 16$, the reconstruction quality is still close to SI (see Figure 9).

As a perspective, let us mention that we only considered line scanning for the PSS strategy. However, if we use two scan mirrors, other scanning trajectories could be considered to further minimize the number of measurements compared to traditional RS. If the mirrors are driven parallely, there would be no increase in the acquisition time. In this case, we could perform speckle scanning in a way similar to RS: a single speckle would scan $M \leq N$ positions in the field of view (FOV). In addition, recent advances in ultrafast scanners employing acousto-optic deflectors [54] and resonant galvanometers [26, 3] will serve to speed up the acquisitions both in the case of focused or speckle illumination. Faster spatial light modulators such as digital micromirror devices [30] and deformable mirrors [8] can also provide a massive speedup to SI, albeit at the cost of experimental integration.

The 2-D setup presented in this work is quite unrealistic for real *in vivo* imaging: (*i*) in practice, the LE will be required to image 3-D volumes and (*ii*) we will not have access

to the fiber endface to image the speckle pattern. Regarding those challenges, we would like to highlight two interesting research directions. The first one is the design of a 3-D acquisition and reconstruction framework. One of the challenges will be to deal with the depth dependency of the speckle autocorrelation [47]. The second one is the study of blind imaging taking advantage of the memory effect like in [7, 62]. In this case, the estimation is based on the autocorrelations of the measurements vector and of the discretized speckle field (known *a priori*). Another solution to deal with the inaccessibility to the fiber endface would be to perform a single initial calibration of the field and then, compute it. Given the high resilience of the fiber to bending, this would alleviate the tedious calibration of mulitple intensity patterns. This would provide a route to the deployment of LEs in a more realistic medical or diagnostic environment.

Appendix A. Total variation and second-order total generalized variation.

The total variation (TV) and the total generalized variation (TGV) of an image are both quantities related to the first- or higher-order derivatives of this image. We start by defining the discrete *gradient* operator ∇ and the *symmetrized derivative* operator ε.

Definition A.1 (from [35]). *The general discrete gradient operator ∇ applicable to $N \times k$ tensor fields is defined as*

$$\nabla : \mathbb{R}^{N\times k} \to \mathbb{R}^{N\times 2k},\ \boldsymbol{u} \mapsto (\nabla \boldsymbol{u}) \coloneqq (\nabla_1 \boldsymbol{u}, \nabla_2 \boldsymbol{u}),$$

with $k \in \mathbb{N}_0$ and $\nabla_i \in \mathbb{R}^{N\times N}$ the first spatial derivative of the tensor field in direction $\boldsymbol{e}_i$, aligned with axis i of the 2-D image.

Definition A.2 (from [35]). *The symmetrized derivative operator ε is*

$$\varepsilon : \mathbb{R}^{N\times 2} \to \mathbb{R}^{N\times 4},\ \boldsymbol{u} \mapsto \varepsilon(\boldsymbol{u}) \coloneqq \frac{1}{2}\left((\nabla \boldsymbol{u}) + (\nabla \boldsymbol{u}) S_{23}\right),$$

where $S_{23} \in \{0,1\}^{4\times 4}$ is a matrix permuting the second and the third columns of $(\nabla \boldsymbol{u})$.

Applying ε on the gradient of an image provides information about its second derivative. The concept of symmetrized tensors is detailed in [13].

Definition A.3. *Let $\boldsymbol{u}_i \in \mathbb{R}^k$ be the i^{th} row of $\boldsymbol{u}$. The $L_{p,q}$-norm of $\boldsymbol{u} \in \mathbb{R}^{N\times k}$ is defined as*

$$\|\boldsymbol{u}\|_{p,q} = \left(\sum_{i=1}^{N} \|\boldsymbol{u}_i\|_p^q\right)^{\frac{1}{q}}.$$

Regarding the previous definitions, the TV-norm of $\boldsymbol{u} \in \mathbb{R}^N$ is defined in a discrete setting as the ℓ_1-norm of the gradient magnitude of $\boldsymbol{u}$,

$$\mathrm{TV} : \mathbb{R}^N \to \mathbb{R},\ \boldsymbol{u} \mapsto \mathrm{TV}(\boldsymbol{u}) \coloneqq \|\nabla \boldsymbol{u}\|_{2,1}. \tag{A.1}$$

Minimizing the TV-norm of an image in an estimation problem like (4.7) leads to piecewise constant estimate. If the original image is not piecewise constant, staircasing artifacts will appear [13]. To alleviate those limitations when working with real, piecewise smooth, images,

we can resort to the TGV-norm. This norm was introduced by Bredies et al. [13] in 2010 and can be considered as the generalization of TV to higher-order image derivatives. The second-order TGV^2_α-norm is defined as

$$\mathrm{TGV}^2_\alpha : \mathbb{R}^N \to \mathbb{R},\ \boldsymbol{u} \mapsto \mathrm{TGV}^2_\alpha(\boldsymbol{u}) \coloneqq \min_{\boldsymbol{w}\in\mathbb{R}^{N\times 2}} \|\nabla \boldsymbol{u} - \boldsymbol{w}\|_{2,1} + \alpha\|\varepsilon(\boldsymbol{w})\|_{2,1}, \tag{A.2}$$

where $\alpha > 0$ is a parameter making a trade-off between the edge-preserving term and the smoothness-promoting term. Deriving $\mathrm{TGV}^2_\alpha(\boldsymbol{u})$ is not as easy as $\mathrm{TV}(\boldsymbol{u})$ because an additional minimization problem has to be solved. We re-write (4.7) in the following way,

$$\hat{\boldsymbol{z}}_\rho \in \arg\min_{\boldsymbol{z}} \|\boldsymbol{R}_{\mathrm{est}}(\sqrt{M}\tilde{\boldsymbol{\Phi}}\bar{\boldsymbol{S}}\boldsymbol{R}_{\boldsymbol{u}}\boldsymbol{z} - \tilde{\boldsymbol{y}})\|^2 + \rho\|(\nabla\boldsymbol{R}_{\boldsymbol{u}} - \boldsymbol{R}_{\boldsymbol{w}})\boldsymbol{z}\|_{2,1} + \rho\alpha\|\varepsilon(\boldsymbol{R}_{\boldsymbol{w}}\boldsymbol{z})\|_{2,1} + \imath_{\mathbb{R}^N_+}(\boldsymbol{R}_{\boldsymbol{u}}\boldsymbol{z}),$$

where $\boldsymbol{z} = (\boldsymbol{u}, \boldsymbol{w})$, $\boldsymbol{R}_{\boldsymbol{u}}$ and $\boldsymbol{R}_{\boldsymbol{w}}$ are restriction operators keeping only the first column of $\boldsymbol{z}$ and the last two columns of $\boldsymbol{z}$, respectively. $\hat{\boldsymbol{f}}_\rho$ is given by $\boldsymbol{R}_{\boldsymbol{u}}\hat{\boldsymbol{z}}_\rho$.

Funding. Part of this research is funded by the FNRS under Grant n° T.0136.20 (project Learn2Sense).

Projects related to lensless endoscopy in Fresnel Institute are supported by Agence Nationale de la Recherche FR NAIMA and National Institutes of Health US R21EB025389 & R21MH117786.

REFERENCES

[1] R. Adamczak, A. E. Litvak, A. Pajor, and N. Tomczak-Jaegermann, *Restricted Isometry Property of Matrices with Independent Columns and Neighborly Polytopes by Random Sampling*, Constructive Approximation, 34 (2011), pp. 61–88.

[2] L. V. Amitonova and J. F. de Boer, *Compressive imaging through a multimode fiber*, Opt. Lett., 43 (2018), pp. 5427–5430.

[3] E. R. Andresen, G. Bouwmans, S. Monneret, and H. Rigneault, *Toward endoscopes with no distal optics: video-rate scanning microscopy through a fiber bundle*, Optics letters, 38 (2013), pp. 609–611.

[4] E. R. Andresen, G. Bouwmans, S. Monneret, and H. Rigneault, *Two-photon lensless endoscope*, Optics Express, 21 (2013), pp. 20713–20721.

[5] E. R. Andresen, S. Sivankutty, V. Tsvirkun, G. Bouwmans, and H. Rigneault, *Ultrathin endoscopes based on multicore fibers and adaptive optics: status and perspectives*, Journal of Biomedical Optics, 21 (2016), p. 121506.

[6] N. Bender, H. Yilmaz, Y. Bromberg, and H. Cao, *Customizing speckle intensity statistics*, Optica, 5 (2018), pp. 595–600.

[7] J. Bertolotti, E. G. Van Putten, C. Blum, A. Lagendijk, W. L. Vos, and A. P. Mosk, *Non-invasive imaging through opaque scattering layers*, Nature, 491 (2012), pp. 232–234.

[8] T. G. Bifano, J. Perreault, R. K. Mali, and M. N. Horenstein, *Microelectromechanical deformable mirrors*, IEEE Journal of Selected Topics in Quantum Electronics, 5 (1999), pp. 83–89.

[9] M. A. Bolshtyansky and B. Y. Zel'dovich, *Transmission of the image signal with the use of a multimode fiber*, Optics Communications, 123 (1996), pp. 629–636.

[10] V. Boominathan, J. K. Adams, M. S. Asif, B. W. Avants, J. T. Robinson, R. G. Baraniuk, A. C. Sankaranarayanan, and A. Veeraraghavan, *Lensless Imaging: A computational renaissance*, IEEE Signal Processing Magazine, 33 (2016), pp. 23–35.

[11] P. T. Boufounos, M. F. Duarte, and R. G. Baraniuk, *Sparse Signal Reconstruction from Noisy Compressive Measurements using Cross Validation*, in 2007 IEEE/SP 14th Workshop on Statistical Signal Processing, Madison, WI, USA, 2007, pp. 299–303.

[12] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge, 2004.
[13] K. Bredies, K. Kunisch, and T. Pock, *Total Generalized Variation*, SIAM Journal on Imaging Sciences, 3 (2010), pp. 492–526.
[14] Cambridge Technology Inc., *62XXH Series Galvanometer Scanners*, (2019).
[15] E. J. Candès and M. Wakin, *An Introduction To Compressive Sampling*, IEEE Signal Processing Magazine, 25 (2008), pp. 21–30.
[16] A. M. Caravaca-Aguirre, S. Singh, S. Labouesse, M. V. Baratta, R. Piestun, and E. Bossy, *Hybrid photoacoustic/fluorescence microendoscopy through a multimode fiber using speckle illumination*, (2018), pp. 1–10.
[17] A. Chambolle and T. Pock, *A First-Order Primal-Dual Algorithm for Convex Problems with Applications to Imaging*, Journal of Mathematical Imaging and Vision, 40 (2010), pp. 120–145.
[18] P. C. Cheng, *The contrast formation in optical microscopy*, in Handbook of Biological Confocal Microscopy: Third Edition, 2006, ch. 8, pp. 162–206.
[19] D. Choudhury, D. K. McNicholl, A. Repetti, I. Gris-Sánchez, T. A. Birks, Y. Wiaux, and R. R. Thomson, *Compressive optical imaging with a photonic lantern*, (2019).
[20] T. Cizmar and K. Dholakia, *Exploiting multimode waveguides for pure fibre-based imaging*, Nature Communications, 3 (2012).
[21] C. A. DiMarzio, *Fourier Optics*, tech. report, 2020.
[22] D. L. Donoho, *Compressed Sensing*, IEEE Transactions on Information Theory, 52 (2006), pp. 1289–1306.
[23] M. F. Duarte, M. A. Davenport, D. Takhar, J. N. Laska, T. Sun, K. F. Kelly, and R. G. Baraniuk, *Single-Pixel Imaging via Compressive Sampling*, IEEE Signal Processing Magazine, 25 (2008), pp. 83–91.
[24] M. F. Duarte and Y. C. Eldar, *Structured compressed sensing: From theory to applications*, IEEE Transactions on Signal Processing, 59 (2011), pp. 4053–4085.
[25] K. Exner and W. Villiger, *The "Optical Power" of the Atmosphere and its Measurement*, The Astrophysical Journal, 21 (1905), pp. 368–370.
[26] G. Fan, H. Fujisaki, A. Miyawaki, R.-K. Tsay, R. Y. Tsien, and M. H. Ellisman, *Video-rate scanning two-photon excitation fluorescence microscopy and ratio imaging with cameleons*, Biophysical journal, 76 (1999), pp. 2412–2420.
[27] B. A. Flusberg, E. D. Cocker, W. Piyawattanametha, J. C. Jung, E. L. Cheung, and M. J. Schnitzer, *Fiber-optic fluorescence imaging*, Nature methods, 2 (2005), pp. 941–950.
[28] S. Foucart and G. Lecue, *An IHT Algorithm for Sparse Recovery from Subexponential Measurements*, IEEE Signal Processing Letters, 24 (2017), pp. 1280–1283.
[29] S. Foucart and H. Rauhut, *A Mathematical Introduction to Compressive Sensing*, Birkhäuser, New York, NY, 2013.
[30] Q. Geng, C. Gu, J. Cheng, and S.-c. Chen, *Digital micromirror device-based two-photon microscopy for three-dimensional and random-access imaging*, Optica, 4 (2017), pp. 674–677.
[31] A. González, L. Jacques, C. De Vleeschouwer, and P. Antoine, *Compressive optical deflectometric tomography: A constrained total-variation minimization approach*, Inverse Problems and Imaging, 8 (2014), pp. 421–457.
[32] J. W. Goodman, *Introduction to Fourier optics*, Roberts and Company Publishers, 2005.
[33] J. W. Goodman, *Speckle phenomena in optics: theory and applications*, Roberts and Company Publishers, 2007.
[34] J. W. Goodman, *Statistical Optics*, John Wiley & Sons, 2015.
[35] S. Guérit, L. Jacques, B. Macq, and J. A. Lee, *Post-Reconstruction Deconvolution of PET Images by Total Generalized Variation Regularization*, in 23rd European Signal Processing Conference (EUSIPCO), 2015.
[36] Hamamatsu, *LCOS-SLM (Liquid Crystal on Silicon-Spatial Light Modulator): X10468/X13267/X13138 series*, tech. report.
[37] L. Jacques and P. Vandergheynst, *Compressed Sensing:"When sparsity meets sampling"*, in Optical and Digital Image Processing - Fundamentals and Applications, Wiley-Blackwell, 2010, pp. 1–30.
[38] O. Katz, Y. Bromberg, and Y. Silberberg, *Compressive ghost imaging*, Applied Physics Letters, 95 (2009), pp. 93–96.

[39] M. K. Kim, C. H. Park, C. Rodriguez, Y. K. Park, and Y. H. Cho, *Superresolution imaging with optical fluctuation using speckle patterns illumination*, Scientific Reports, 5 (2015), pp. 1–10.
[40] F. Knoll, K. Bredies, T. Pock, and R. Stollberger, *Second order total generalized variation (TGV) for MRI.*, Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine, 65 (2011), pp. 480–91.
[41] A. Liutkus, D. Martina, S. Popoff, G. Chardon, O. Katz, G. Lerosey, S. Gigan, L. Daudet, and I. Carron, *Imaging with nature: Compressive imaging using a multiply scattering medium*, Scientific reports, 4 (2014), pp. 1–7.
[42] B. Lochocki, K. Abrashitova, J. F. de Boer, and L. V. Amitonova, *Ultimate resolution limits of speckle-based compressive imaging*, Optics Express, 29 (2021), p. 3943.
[43] J. Min, J. Jang, D. Keum, S. W. Ryu, C. Choi, K. H. Jeong, and J. C. Ye, *Fluorescent microscopy beyond diffraction limits using speckle illumination and joint support recovery*, Scientific Reports, 3 (2013), pp. 1–6.
[44] E. Mudry, K. Belkebir, J. Girard, J. Savatier, E. Le Moal, C. Nicoletti, M. Allain, and A. Sentenac, *Structured illumination microscopy using unknown speckle patterns*, Nature Photonics, 6 (2012), pp. 312–315.
[45] V. Ntziachristos, *Going deeper than microscopy: The optical imaging frontier in biology*, Nature Methods, 7 (2010), pp. 603–614.
[46] N. Parikh and S. Boyd, *Proximal algorithms*, Foundations and Trends in optimization, 1 (2013), pp. 123–231.
[47] M. Pascucci, S. Ganesan, A. Tripathy, O. Katz, V. Emiliani, and M. Guillon, *Compressive three-dimensional super-resolution microscopy with speckle-saturated fluorescence excitation*, (2017), pp. 1–36.
[48] A. Perperidis, K. Dhaliwal, S. McLaughlin, and T. Vercauteren, *Image computing for fibre-bundle endomicroscopy: A review*, Medical Image Analysis, 62 (2020).
[49] M. Plöschner, T. Tyc, and T. Čižmár, *Seeing through chaos in multimode fibres*, Nature Photonics, 9 (2015).
[50] D. Psaltis and C. Moser, *Imaging with Multimode Fibers*, Optics and Photonics News, 27 (2016), pp. 24—-31.
[51] M. Raginsky, R. M. Willett, Z. T. Harmany, and R. F. Marcia, *Compressed sensing performance bounds under poisson noise*, IEEE Transactions on Signal Processing, 58 (2010), pp. 3990–4002.
[52] H. Raguet, J. Fadili, and G. Peyré, *Generalized Forward-Backward Splitting*, SIAM Journal on Imaging Sciences, 6 (2013), pp. 1199–1226.
[53] H. Rauhut, J. Romberg, and J. A. Tropp, *Restricted isometries for partial random circulant matrices*, Applied and Computational Harmonic Analysis, 32 (2012), pp. 242–254.
[54] G. D. Reddy, K. Kelleher, R. Fink, and P. Saggau, *Three-dimensional random access multiphoton microscopy for functional imaging of neuronal activity*, Nature neuroscience, 11 (2008), pp. 713–720.
[55] L. Rudin, S. Osher, and E. Fatemi, *Nonlinear total variation based noise removal algorithms*, Physica D: Nonlinear Phenomena, 60 (1992), pp. 259–268.
[56] J. H. Shapiro and R. W. Boyd, *The physics of ghost imaging*, vol. 11, 2012.
[57] J. Shin, B. T. Bosworth, and M. A. Foster, *Compressive fluorescence imaging using a multi-corefiber and spatially-dependent scattering*, Optics letters, 42 (2017), pp. 109—-112.
[58] S. Sivankutty, D. Kogan, V. Tsvirkun, G. Bouwmans, E. Andresen, M. Guillon, M. Alonso, D. Oron, and H. Rigneault, *Non-interferometric calibration of the phase transmission matrix in lensless endoscopy (conference presentation)*, in Adaptive Optics and Wavefront Control for Biological Systems V, vol. 10886, International Society for Optics and Photonics, 2019, p. 108860H.
[59] S. Sivankutty, V. Tsvirkun, G. Bouwmans, D. Kogan, D. Oron, E. R. Andresen, and H. Rigneault, *Extended field-of-view in a lensless endoscope using an aperiodic multicore fiber*, Optics Letters, 41 (2016), p. 3531.
[60] S. Sivankutty, V. Tsvirkun, O. Vanvincq, G. Bouwmans, E. R. Andresen, and H. Rigneault, *Nonlinear imaging through a Fermat's golden spiral multicore fiber*, Optics Letters, 43 (2018), p. 3638.
[61] L. Song, E. J. Hennink, I. T. Young, and H. J. Tanke, *Photobleaching kinetics of fluorescein in quantitative fluorescence microscopy*, Biophysical Journal, 68 (1995), pp. 2588–2600.
[62] N. Stasio, C. Moser, and D. Psaltis, *Calibration-free imaging through a multicore fiber using speckle*

scanning microscopy, Optics Letters, 41 (2016), p. 3078.
[63] P. Sudhakar, L. Jacques, X. Dubois, P. Antoine, and L. Joannes, *Compressive Imaging and Compressive Imaging and Characterization of Sparse Light Deflection Maps*, (2018), pp. 1–35.
[64] T. Tanaka, *Performance analysis of L1-norm minimization for compressed sensing with non-zero-mean matrix elements*, in IEEE International Symposium on Information Theory, IEEE, 2018, pp. 401–405.
[65] A. J. Thompson, C. Paterson, M. A. A. Neil, C. Dunsby, and P. M. W. French, *Adaptive phase compensation for ultracompact laser scanning endomicroscopy*, Optics Letters, 36 (2011), p. 1707.
[66] R. Y. Tsien, L. Ernst, and A. Waggoner, *Fluorophores for confocal microscopy: Photophysics and Photochemistry*, in Handbook of Biological Confocal Microscopy: Third Edition, 2006, ch. 16, pp. 267–279.
[67] V. Tsvirkun, S. Sivankutty, K. Baudelle, R. Habert, G. Bouwmans, O. Vanvincq, E. R. Andresen, and H. Rigneault, *Flexible lensless endoscope with a conformationally invariant multi-core fiber*, Optica, 6 (2019).
[68] S. Turtaev, I. T. Leite, K. J. Mitchell, M. J. Padgett, D. B. Phillips, and T. Čižmár, *Comparison of nematic liquid-crystal and dmd based spatial light modulation in complex photonics*, Optics express, 25 (2017), pp. 29874–29884.
[69] G. C. Valley, G. A. Sefler, and T. Justin Shaw, *Multimode waveguide speckle patterns for compressive sensing*, Optics Letters, 41 (2016), p. 2529.
[70] R. Vershynin, *Introduction to the non-asymptotic analysis of random matrices*, (2010), pp. 210–268.
[71] R. Vershynin, *High-Dimensional Probability: An Introduction with Applications in Data Science*, (2016).
[72] R. Ward, *Compressed sensing with cross validation*, IEEE Transactions on Information Theory, 55 (2009), pp. 5773–5782.