Compressed sensing imaging: towards low-phototoxicity and high-speed microscopy

Maxime Woringer[1,2] *, under the supervision of Mustafa Mir[2], Christophe Zimmer[1] and Xavier Darzacq[2]
[1]: IMOD lab, Pasteur Institute, Paris & [2]: Darzacq lab, UC Berkeley

March–October 2016

Contents

Abstract

To image live cells, the light intensity at the sample is bounded by two limiting factors: first, photodamage and photobleaching impose an upper bound on the total and instantaneous incident light intensity. Second, a reduced light intensity results in higher SNR and images of lower quality. Compressed sensing proposes a paradigm to efficiently image sparse and compressible signals, providing a better SNR vs. photodamage trade-off than the traditional Nyquist-derived sampling schemes.

Here, we partially developed a widely applicable compressed sensing scheme that can be fitted with little or no hardware modification to a wide range of microscopes, from a high-performance lattice light sheet microscope to a standard epifluorescence microscope.

We first demonstrated on simulations the feasability of compressed sensing techniques under noisy conditions and designed a scheme where the signal of a 3D stack is compressed in z during at the imaging step. We demonstrated the applicability of this scheme both on extensive simulations and on preliminary implementation on a lattice light sheet microscope.

An implementation an generic epifluorescence microscope together with better quality reconstructions are still in progress.

*Maxime Woringer, maxime.woringer@ens.fr

Introduction

Microscopy plays a central role in life sciences and constitutes in itself a rapidly-evolving field that requires sustained collaboration between wet-bench biologists, physicists, chemists, computer scientists, etc. As technology progresses, it enables researchers to focus on highly dynamic processes, that have been proven to be of high importance in biological systems.

To observe highly dynamic structures requires live-sample imaging which has always been a challenging task, primarily for the following reasons:

1. **Photobleaching**: The limited number of photons emitted by one fluorophore: once a fluorescent molecule has been excited, it can only be imaged for a limited amount of time due to photobleaching. The highly reactive environment of a live cell combined with the intense laser excitation applied to the sample greatly limits the number of photons that can be collected from one fluorophore before it bleaches

2. **Contrast**: High background fluorescence: the complex mixture of cell constituents includes aromatic compounds, conjugated residues and many slightly-fluorescent components that add background noise. Furthermore, traditional fluorescence microscopy techniques illuminate the whole sample, beyond the depth of field, which leads to diffuse background noise.

3. **Phototoxicity**: This leads to the formation of reactive compounds such as highly oxidizing reactive oxygen species (ROS). These compounds have various visible consequences, ranging from the formation of stress vacuoles in cells to apoptosis, event triggering developmental defects in small organisms.

To partially tackle these issues , several major technological breakthroughs were achieved in the past ten years and provided biologists with better labeling dyes (with increased quantum efficiency and lifetime Grimm et al., 2015), better labeling techniques (allowing to conjugate non-protein dyes to proteins of interest – Los et al., 2008), better hardware (electron-multiplying CCD cameras and high-power lasers), and better microscopes (B.-C. Chen et al., 2014).

Among these advances, light sheet microscopes (Huisken et al., 2004) can be given a specific emphasis: whereas in a traditional microscope, the same objective is used to illuminate the sample and collect the emitted light, light sheet microscopes use two objectives at a 90° angle, the first one that illuminates only one plane of the sample, the other being a standard observation objective (Figure 1b). With this microscope, further refined into a lattice light sheet microscope (B.-C. Chen et al., 2014, detailed in Section 6.1), off-focus illumination and phototoxicity are greatly reduced because the setup allows to only illuminate a very thin "sheet" that matches the depth of field of the detection objective.

Although the signal we aim to measure is continuous, camera detectors only provide a discretized, noisy image of the object. Thus, an appropriate sampling scheme (altogether with carefully chosen hypotheses) is needed to accurately reproduce the intensity at the sample. Traditional imaging makes use of the asumption that the collected signal is bandlimited and samples at the *Shannon-Nyquist* rate (Kim, 2011 and Wescott, 2010).

However, microscopy images are not only bandlimited, they also exhibit a much stronger structure, with a lot of redundancies. Such signals with redundancies, denominated as *compressible signals*, can be sampled with a much higher efficiency, as described by the theory of compressed sensing, whose foundations were set by E.J. Candes, J. Romberg, and T. Tao, 2006, provided that an *ad-hoc* sampling scheme can be implemented. In this theory, a "compressed" version of the signal is directly acquired with a sensor that collects linear measurements of the sample (sometimes achieving a more than 10-fold decrease in the acquisition time). Then, a reconstruction algorithm "decompresses" the collected sequences.

The aim of this work is to evaluate the feasability of the implementation of such a compressed-sensing imaging scheme within the lattice light sheet microscope. After presenting the theory behind lattice light sheet microscopy and the canonical setup of compressed sensing, we address in order several theoretical and implementation-related questions: (1) how to design a sampling scheme that can be implemented in such a microscope? (2) which sensing pattern provide the best reconstructions? (3) what is the influence of both Gaussian and Poisson noise in the acquisitions? (4) can reconstructions be further improved by incorporating more information about the data, that is, by performing dictionary learning?

After investigating these questions based on theoretical grounds and simulations, we then turn to the practical implementation of the scheme on the microscope and image biological samples.

Movies and code are provided as a webpage[1].

1 Material and Methods

1.1 Lattice light sheet microscope

In a traditional epifluorescence microscope, the laser excitation and observation are performed through the same objective. This leads to an inefficient scheme where the sample is illuminated beyond the depth of field. To tackle this issues, light sheet microscope (Huisken et al., 2004) employ a paradigm where two orthogonal objectives are used to perform excitation and observation (Figure 1). In this setup, the excitation laser is concentrated on a thin "sheet". Light sheet microscopes usually achieve a z sectioning of about 1 µm.

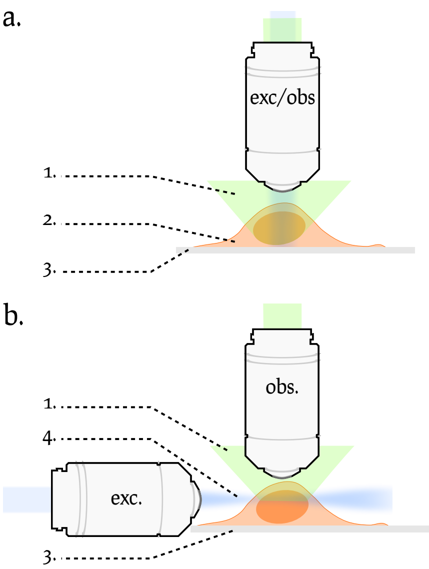

Figure 1: Comparison between **(a)** standard epifluorescence microscopy and **(b)** light sheet microscopy. In epifluorescence microscopy, the whole sample is illuminated whereas only part of it is in the focal plane, resulting in background fluorescence and photodamage. In light sheet microscopy, illumination is confined to the focal plane of the objective. **1.** emitted light from the illuminated sample **2.** sample labeled with a fluorescent dye **3.** glass coverslip **4.** excitation light. **exc.** excitation objective, **obs.** observation objective.

[1]Several movies of biological samples together with analysis code and supplementary materials can be found on this page: `http://www.eleves.ens.fr/home/woringer/CS/`

Several refinements of the light sheet allowed to obtain even better z sectioning, and thus less phototoxicity and photobleaching. One of them is the lattice light sheet microscope (B.-C. Chen et al., 2014). In the lattice light sheet microscope the sheet derives from the interference of an array of non-diffracting Bessel beams. Indeed, purposely induced interference is used to reduce the amplitude of the higher order rings of the Bessel beams (see Section 6.1 for a more detailed presentation of the theory behind lattice light sheet microscopes).

More concretely, an array of interfering Bessel beams can be engineered as follows. First, note that in a microscope, several planes are conjugated. The most important ones are the pupil/sample plane (or object/direct domain) and the back pupil plane (or Fourier domain). The two are conjugated through Fourier transforms (see Kim, 2011 for an introduction to Fourier optics). Second, notice that the Fourier transform of a first order, first kind Bessel function is an annulus. Then, a Bessel beam can be created by simply putting an annulus mask in the Fourier domain. Third, an array of Bessel beam can be engineered by restraining the illumination of the annulus to several spots. The spacing of the spots then determines the spacing of the lattice, and thus the magnitude of the interference patterns.

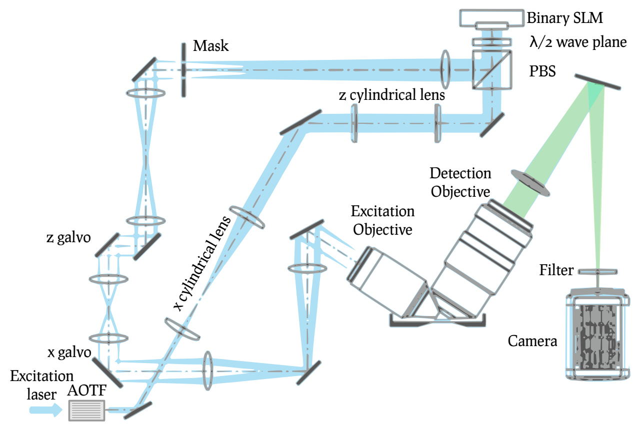

Figure 2: Optical path of the lattice light sheet microscope. Laser excitation (bottom left) is first modulated in intensity by an AOTF and then compressed in the z direction and stretched in the x direction. The deformed beam is then sent to the SLM where a mask pattern is applied in the image plane. In the next Fourier plane, an annular mask creates the Bessel beam. Then, the z galvo keeps the sheet in focus and synchronized with the observation objective while the x galvo scans the sample to create the light sheet. The light sheet is then focused by the excitation objective that is positioned orthogonally to the observation objective. Sample lies in between the two objectives. **blue**: excitation part, **green**: observation part. Adapted from B.-C. Chen et al., 2014.

In practice, a spatial light modulator, a liquid crystals device, or SLM in the direct domain is used to create the array of Bessel beam. This array is then filtered by an annulus mask in the Fourier domain (then filtering the unwanted frequencies). A schematic of the optical light path of a lattice light sheet microscope is presented Figure 2

1.2 Compressed sensing, a (very) informal introduction

1.2.1 Signals are compressible...

Many signals, including microscopy images, exhibit a structure of *compressible signals*. Indeed, although the uncompressed size of a 3D volume image can be very high (e.g: 512 px × 1024 px × 300 px @ 16 bits/px = 315 MB), a simple *lossless* file compression can dramatically reduce its size on disk (the same image, 16 bits/px, lossless `PNG` compressed weighs 48 Mo): in this lossless process, the redundant information has been factorized, thus reducing the file size.

Another way to represent compressible signals is to find a representation (which would be a basis or a dictionary in our case) where the signal has sparse coefficients, that is, most of its coefficients are zero in this representation.

For instance, the image of a mitotic cell X (Figure 3.a) can be decomposed into a sparse series of coefficients C (Figure 3.b) associated with a *sparsifying basis* B (Figure 3.c, inset), such that $X = CB$ (or in a more realistic manner: $||X - CB||_{\ell_2} < \epsilon$ and C sparse, ie: $||C||_{\ell_0} \leq k_0$). In the pictured case, the *sparsifying basis* has been learned from a dataset that contains similar features as the sample images (the dataset was a series of mitotic cells). One can see that in a *sparsifying basis* that incorporates prior information on the structure of mitotic cells, the particular sample (Figure 3.a) exhibits a sparse representation, as shown by (Figure 3.c), i.e. its ordered coefficients decay quickly to zero (red curve), compared to the initial image (blue curve).

1.2.2 ...because they have structure

This property of compressibility is tightly related to the fact that such images are far from being an aggregation of pixels of random intensity: they exhibit some *structure*, and such structure is captured by the `PNG` file compression algorithm.

Stated differently, this means that a significant part of the data that goes through the camera is in fact highly redundant, and could possibly be dropped without any effect on the data quality. Compressed sensing is all about trying to acquire only the relevant information, thus speeding (sometimes dramatically) the process. For instance, if one could directly acquire an image in a PNG format, a $\sim$ 6x speedup could be envisioned.

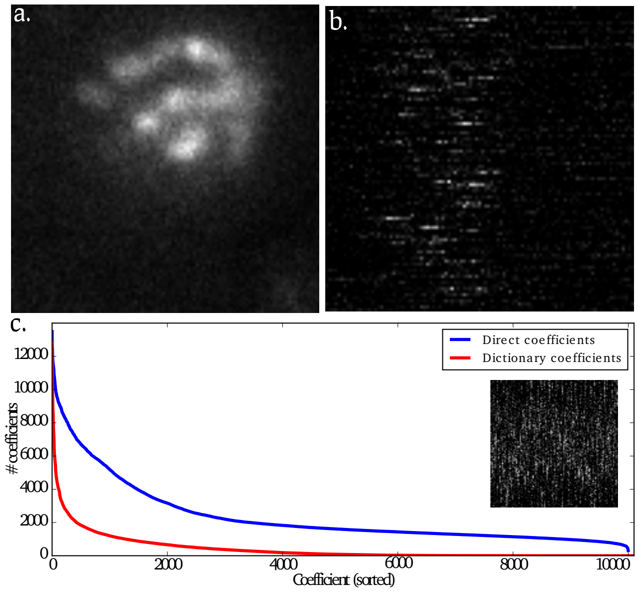

Figure 3: Microscopy images are compressible. **(a).** zoom on an image acquired using the lattice light sheet microscope. Most pixels have nonzero values, making the signal *non sparse*. However **(b).**, this signal can be made sparse by using a decomposition such as non-negative matrix factorization. **(c)** such a decomposition uses a matrix prior learned from the data: a dictionary **(inset)**. Compressibility can be appreciated by looking at the decay of the coefficients in both representations: with the dictionary decomposition, many more coefficients have zero or near-zero values.

1.2.3 Some structure can be captured in a non-adaptive way

File compression algorithms (such as lossless `PNG`), however, perform their tasks in an adaptive way: one or several passes are first performed over the whole dataset and then the most frequent features are "factorized". We call such transforms *adaptive* because the nature of the filter will differ depending on the input.

A brilliant result from the compressed sensing theory states (E.J. Candes, J. Romberg, and T. Tao, 2006) that under some conditions, a series of *non-adaptive* measurements can be performed and that the input signal can be recovered *exactly* using a simple algorithm performing ℓ_1-norm minimization, provided that it is sparse/compressible. We present below the intuition of this reconstruction:

1.2.4 Toy problem

Assume a 1D image X (such as an image made of only one line) of length $n = 100$. Let, for the sake of the example that this image contains two Dirac-like features (here: two non-zero componenents), leading to a sparsity of 98 (Figure 4). *Sparsity* is defined as the number of zero components, also denoted as the ℓ_0 pseudonorm: $||X||_{\ell_0} = \#\{x_j, \text{ st. } x_j = 0\}$.

A classical view in signal theory is to consider that X goes through a linear detector (a virtual sensor that incorporates the *imaging scheme*, its imperfections, etc): one never measures the real signal X, but rather a **measurement** of X through a sensor A, also denoted as a **sensing matrix**, $Y = AX$. A perfect sensor is modeled as the identity matrix, where Y is a copy of X.

The sensing matrix A has dimensions $m \times n$. In the ideal sensor case, the matrix has dimension $n \times n$, but one can also design sensing matrices with a number of lines $m < n$. Note that any square matrix of rank n provides a *lossless* transformation of X: $Y = AX \Leftrightarrow X = A^{-1}Y$. Indeed, in the case where A is a Fourier matrix, X can be recovered from Y by applying the inverse (discrete) Fourier transform.

Compressed sensing focuses on sensing matrices of low rank ($m < n$, and most of the time $m \ll n$). Such cases are referred here as a compressed sensing setting and a measurement matrix can be (non exhaustive list):

- a truncated Fourier transform (where only the m first low-frequency components have been kept, the others being discarded), see Figure 4.c, left,
- a random matrix with uniformly distributed coefficients over $[0, 1]$ (ie: $\forall (i,j) \in 1 \ldots m \times 1 \ldots n, A_{ij} \sim \mathcal{U}([0,1])$), see Figure 4.c, right.

In the compressed sensing setting, one wants to recover X from the knowledge of $Y = AX$ and A. However, it easy to see that the problem is not well posed because $Y = AX$ is underparametrized and thus has infinitely many solutions. Figure 4.d shows several solutions of this problem.

Fortunately, one can add constraints to the system, such as constraining the sparsity of X. This allow to formulate the compressed sensing problem as an optimization problem:

$$\min_x ||x||_{\ell_0} \text{ s.t. } y - Ax = 0 \tag{1}$$

In this case, a beautiful theorem by E.J. Candes, J. Romberg, and T. Tao, 2006 states that the vector X of length n can be exactly recovered from $m \ll n$ measurements (that is, Y). This can be seen in Figure 4.d : although the truncated Fourier transform of two Dirac forms a highly underparametrized system, the constrained problem where one looks for a sparse solution of the system exhibits the exact solution (upon numerical approximation and assuming numerical convergence).

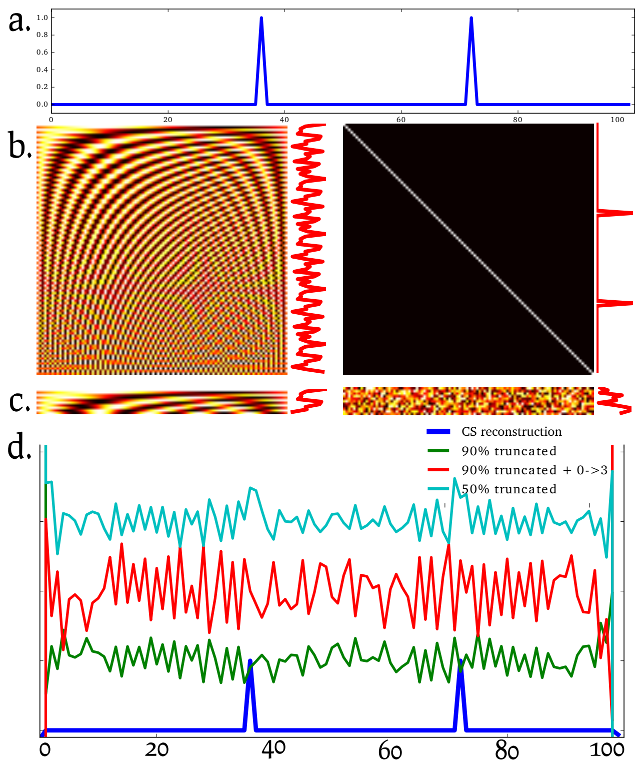

Figure 4: Toy compressed sensing problem. **(a).** Assume a sparse input signal with only two non-zero coefficients, **(b).** this input signal can be *measured* by various measurement matrices, here (left) a Fourier basis and (right) an ideal Dirac basis, the red curves on the right of each basis represent the measured signal Y, that is AX. **(c).** compressive measurement can also be performed, in this case, the measurement matrix is not a square anymore. Such matrices can for instance be a truncated Fourier basis (left) or a random matrix (right). Again, the red curves on the right show the measured signal, that now have a smaller length than with square measurements matrices. **(d).** recovering X from such an undersampled Y is a ill-posed problem: all the curves represent $X_{candidate}$ solution of the $Y = AX$ problem. However, it can be seen on this particular example that the solution with higher sparsity matches the original signal. **(blue)** solution to the problem obtained by ℓ_1 minimization (closely matches the original X, **(green)** signal recovered by inverse Fourier transform from 90% of the Fourier coefficients of X, **(red)** same as green, but the 10% remaining coefficients were assigned the value 3.0, **(light blue)** same as green, but recovered from only 50% of the coefficients. An offset has been added to the curves in order to differentiate them but vertical scale is the same.

Here, we considered a very simplified model (the signals are Dirac-like), but this result generalizes to sparse signals in general, as detailed in Section 6.2.1.

1.3 Noisy setting

1.3.1 A theory to account for noise in compressed sensing

The above presentation (together with Section 6.2) highlighted the foundations of compressed sensing in a noiseless setting, in which there is a strict equality between the measured signal y and the input signal x measured through the measurement matrix A: $y = Ax$.

Unfortunately, this situation never arises in real cases, because images are contaminated with various levels of noise. The most traditional model to incorporate noise into the compressed sensing setting is to relax the equality hypothesis by adding an error tolerance term ϵ. The relaxed compressed sensing setting then reads (Eq. 2):

$$\min_x ||x||_{\ell_0} \text{ s.t. } ||Ax - y||_{\ell_2} \leq \epsilon \tag{2}$$

Such a relaxation indeed allows some noise (of bounded ℓ_2 norm) to interfere with the measurement. Also, since $\epsilon > 0$, the space of solutions to the problem is bigger, and thus usually contains sparser solutions than when $\epsilon = 0$. Solving such problem thus performs some *denoising* on the input. Such denoising can either be beneficial or detrimental to the problem, depending on the value of ϵ.

In this setting, the question of the uniqueness of the solution makes little sense and is replaced by the notion of stability: how big is the change in the reconstructed x when small perturbations (noise) are applied to the measurement Ax.

Remarkably, the concept of spark of a matrix $\mu(A)$ (introduced in Supplementary Section 6.2) can be extended to the noisy case, and stability can be proven:

Theorem 1 (Stability of the relaxed problem)**.** *Denote x^ϵ a solution to the relaxed problem (2) and assume that the solution x to the non relaxed problem obeys $||x||_{\ell_0} \leq \frac{1}{2}\left(1 + \frac{1}{\mu(A)}\right)$, then x^ϵ obeys:*

$$||x^\epsilon - x||_{\ell_2} \leq \frac{4\epsilon^2}{1 - \mu(A)(2||x||_{\ell_0} - 1)}$$

Although interesting from a theoretical standpoint, this result makes some asumptions on the structure of the noise that happen to be violated in the case of microscopy imaging. We thus move to a more physics-oriented approach to understand the effect of noisy acquisitions on compressed sensing settings.

1.3.2 Sources of noise in microscopy

Several types of noises contaminate microscopy images:

1. **Poisson noise:** in an experimental setup, light is collected on a matrix-based detector (CCD for instance). Furthermore, the emitted fluorescence is a stochastic process such that the average number of photons per time unit corresponds to the average power of the flow. We can then consider that each pixel of the detector integrates (counts) the number of photons it receives in one time step. Then, at each time step, the number of photons detected can be seen as a realization of a Poisson random variable whose intensity is the "real" light brightness we want to estimate. Thus, for a given intensity λ at one spatial position (in one pixel), the probability to collect k photons in a t-long time interval is as follow: $\mathbb{P}(N = k) = \frac{e^{-\lambda t}(\lambda t)^k}{k!}$.

2. **Thermal noise:** Going from potentially single photons per pixel to a detectable voltage on the acquisition card of the computer requires first a very high gain, and second a very high resilience to noise and interferences. Indeed, in EMCCD cameras, to achieve one-photon sensitivity, a quasi-single electron has to be amplified, and any single noise in the preamplification part will result in very high distortions after amplification. Furthermore, for such small currents, thermal variations are sufficient to induce noise. To tackle this issue, camera sensors and amplifiers operate at very low temperature (-100 °C to -50 °C) and the amplification (called electron-multiplication – EM) is realized at the readout stage. All in all, thermal noise is usually low compared to Poisson noise under traditional operating conditions.

3. **Structural error:** In many cases, the measurement matrix A is not known with infinite precision. For instance, the synchronization between the optical and mechanical parts of the microscope is never perfect, and a non-negligible systematic error between the "theoretical" measurement matrix and the actual measurement matrix can occur (in addi-

tion to random fluctuations in the process). In contrast with the two types of noise introduced above, this type of error is deterministic, and has been little studied.

4. **Other types:** it is always possible to add other types of noise, such as quantization noise, etc. but those are usually negligible compared to Poisson noise and thermal noise.

A topical issue with imaging and compressed sensing is the fact that in contrast with traditional, widely studied Gaussian noise, the Poisson noise that dominates the acquisitions is unbounded and signal-dependent. Such settings have been little studied (Raginsky, Willett, et al., 2010, Willett and Raginsky, 2009).

1.3.3 Influence of structural errors

Consider the basis pursuit formulation of compressed sensing:

$$x^\star = \arg\min_x ||x||_{\ell_1} \text{ s.t. } ||y - Ax||_{\ell_2} \leq \epsilon$$

where x is s-sparse. One notices that this model incorporates *additive noise* by the means of the ϵ term, accounting for some discrepancies between the ideal measurement of the sparse vector Ax and its actual measurement y. However, one can wonder what happens if some structural perturbations are introduced to the sensing matrix, that is $\hat{A} = A + E$. In that case, it is easy to see that the measurement is also contaminated by *multiplicative noise*: Ex. A mathematical treatment of such situation has been developed in Herman and Strohmer, 2009.

Assume relative bounds on the following quantities:

- $\frac{||E||_2}{||A||_2} \leq \epsilon_A$ the relative magnitude of the structural error compared to the measurement matrix,
- $\frac{||E||_2^{(s)}}{||A||_2^{(s)}} \leq \epsilon_A^{(s)}$ where the norm $||\cdot||_2^{(s)}$ denotes the maximum of the norm over all $m \times s$ subvector with s elements,
- $\frac{||e||_2}{||y||_2} \leq \epsilon_y$ the relative magnitude of the additive error term compared with the measured vector.
- $\frac{\sigma_{\max}^{(s)}}{\sigma_{\min}^{(s)}} \leq \kappa_A^{(s)}$ where $\sigma_{\max}^{(s)}$ ($\sigma_{\min}^{(s)}$, respectively) denotes the maximum (resp. minimum) over all s-colums submatrices of A of the maximum (resp. minimum) of the singular value of A restricted to s columns.

It is easy to see that when A satisfies a s-RIP of constant $\delta_s \ll 1$ then $\kappa_s^{(s)} \approx 1$. It can then be shown that (denoting $\tilde{x}$ the original vector before compression):

$$||x^\star - \tilde{x}||_{\ell_2} \leq C_1 \left(\kappa_A^{(s)} \epsilon_A^{(s)} + ||\epsilon_y||_{\ell_2}\right) ||y||_{\ell_2}$$

This relationship gives a bound of the reconstruction error that depends on the relative structural error.

It should be noted that such structural stability analysis has been very little studied.

1.3.4 Effect of noise on the measurements

Consider a 1-dimensional sample of brightness $(\lambda_j)_{1\leq j<n}$ sensed with a series of m measurements performed with a measurement matrix $(a_{ij})_{1\leq i<m, 1\leq j<n}$, where the measurement matrix represents a light pattern applied at the sample (a laser power modulation, thus being bounded by nominal laser power). Then, using the m lines of the measurement matrix, m measurements can be performed, denoted $(y_i)_{1\leq i<m}$, that is:

$$y_i = \sum_{j=0}^{n} \lambda_j a_{ij}$$

where the y_i precisely stands for the number of photons (provided the matrices and inputs were properly scaled). Then, the actual number of photons detected at the camera follows a Poisson process $\mathcal{P}$ of intensity $\sum_{j=0}^{n} \lambda_j a_i j$:

$$y_i \sim \mathcal{P}\left(\sum_{j=0}^{n} \lambda_j a_i j\right)$$

Now, assume that one acquires one compressed measurement (one projection of the image onto one of the vectors of the measurement basis) using similar parameters as the standard acquisition scheme. This way, since the camera settings are assumed to be the same, thermal noise should affect compressed and standard measurements the same way. Indeed, if one acquisition in the standard scheme is contaminated by Gaussian, thermal noise $\sigma_{SN} \sim \mathcal{G}(\mu, \sigma^2)$, the sensed light y_i follows a law similar to:

$$y_i \sim \mathcal{P}\left(\sum_{j=0}^{n} \lambda_j a_{ij}\right) + \sigma_{SN}$$

In a case where only Gaussian additive noise is considered, one can establish a direct relationship between noise levels in the standard acquisition scheme and in the compressed sensing scheme (see Arias-Castro and Eldar, 2011). However, when signal-dependent noise such as Poisson noise apply, there is no easy way to compare it with noise levels in the direct domain.

Indeed, the noise applies in the transformed domain rather than in the direct domain. This noisy measurement will then be processed by a nonlinear compressed sensing reconstruction algorithm. Such algorithms rarely have accurate estimates on noise sensitivity (although bounds are often derived).

Furthermore, it should be noticed that if the signal is very sparse in the direct domain, then the noiseless measurement will have small magnitude (even though the saqrse regions might have a high magnitude), and the measurement term might become negligible compared to the intrinsic thermal noise.

Finally, several algorithms have been designed to specifically address the effects of such noises.

1.4 Algorithms

The compressed sensing problem (Eq. 1) can be reformulated into several equivalent or very similar problems (such as the relaxed version). Additionally, it has been shown that a very useful and equivalent reformulation can be obtained by replacing the ℓ_0 non-convex pseudonorm by the convex ℓ_1 norm. This leads to a high variety of algorithms, ranging from linear programs to Lasso-derived algorithms. In total, dozens, if not hundreds of algorithms have been proposed, each one exhibiting specific features (speed, parallelization, accuracy, invariance to noise, specific constraints such as positivity, etc). Section 6.3 details the principle of a toy algorithm and characteristics of several algorithms benchmarked for this work.

For most of the reconstructions, we used the general purpose ℓ_1- `magic` routines (Emmanuel Candes and Justin Romberg, 2005) for their speed of reconstruction and ease of use. We also used the `SPIRALTAP` algorithm (Harmany, Marcia, and Willett, 2012). This algorithm specifically incorporates a Poisson likelihood in the optimization criterion and enforces positivity of the reconstructed signal, making it particularly suitable for imaging purposes. We developed a port in Python of the original Matlab® code (see Section 6.7).

1.5 Sparsifying transforms and dictionary learning

1.5.1 How to engineer sparsity: sparsifying basis

The theory and algorithms presented above provide both theoretical guarantees and empirical evidence that signals can be recovered (decompressed) from much less information than usually required by the Shannon-Nyquist sampling theorem. The power of this theory is that it makes only one asumption on the measured signal: its sparsity.

Now looking at traditional microscopy imaging (see for instance Figure 3a), it is clear a major part of the image is dark, and one can say that the image is sparse in a sense. However, it turns out that such a "sparse" image is not amenable to compressed sensing. Indeed, whereas Figure 3a is at most approximately 50 % sparse, compressed sensing usually deals with signals that are $>$ 90 % sparse.

This lack of sparsity might be a major hurdle to apply compressed sensing to microscopy signals. Fortunately, it turns out that:

1. If the signal has a sparse representation in some known basis, all the compressed sensing theory applies with very little modification.

2. One can show that an uncertainty principle holds that states that a signal cannot be spread out in all of its representations.

These properties are briefly examined from a theoretical standpoint in Section 6.4. They can be interpreted as follows: even though the signal one want to acquire is not sparse, it is very likely that there exists a basis that transforms it into a sparse signal. Once this sparsifying transform D has been found, the compressed sensing problem can bery easily be adapted to suit a sparsifying transform:

Assume that for an input vector x (of sparsity $||x||_{\ell_0} = s$, but not necessarily a sparse signal) there exists a *known* basis D such that $x' = Dx$ of sparsity $||x'||_{\ell_0} = s'$ is sparse enough for compressed sensing ($s' \ll s$). A compressed signal y is acquired using a measurement basis A. One can then define a "surrogate" sensing matrix A' such as $A' = AD$ and apply a compressed sensing reconstruction algorithm to recover the sparse vector x' from y, knowing A', that is, solve the problem:

$$\min_{x'} ||x'||_{\ell_0} \text{ s.t. } ADx' = y$$

Finally, a simple change of basis allows to bijectively switch to x. Such a generalization of compressed sensing is better vizualized by Figure 5.

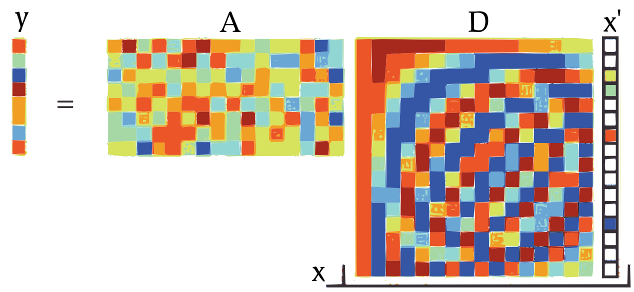

Figure 5: Compressed sensing with a sparsifying basis. Assume that the original vector x can be decomposed into the matrix product of a known basis D and an (unknown) sparse vector x'. In this setup, a compressed sensing reconstruction algorithm can be applied: one can recover x' from the measured y. Finally, the non-sparse signal x can be recovered by a matrix product: $x = Dx'$. (adapted from Baraniuk, 2007)

Here follow two comments:

1. *Do we need to measure differently?* In contrast with the measurement basis that has to be physically implemented in the measurement setup, the use of a sparsifying basis can be seen as a *software trick* since any sparsifying basis can be applied once the measurement has been acquired. This is a huge asset of this method, because the sparsifying basis can be optimized without the constraints of the physical setup.

2. *Examples of sparsifying basis.* Many sparsifying bases have been introduced, following the results derived from wavelets theory: Discrete Fourier Transforms, curvelets, etc. In the domain of microscopy, little information has been published, although Studer et al., 2012 used a traditional wavelet transform as a sparsifying basis.

1.5.2 Dictionary learning: beyond sparsifying basis

Notice that the sparsifying transform doesn't need to be restricted to a basis, and frames such as dictionaries can also be used. Learned dictionaries is a machine learning approach that allows to learn an *ad-hoc* sparsifying dictionary from sample data. It has proven particularly successful when no sparsifying transform can be "guessed" by simply looking at the data. Section 6.4 details the dictionary learning problem.

Several algorithms have been published to perform dictionary learning. The most famous is probably `K-SVD`, (described for instance in Elad, 2010).

However, when employed with images, algorithms that enforce a positive dictionary (constituted only by positive elements) have been shown to give better performances. Such algorithms fall onto the category of non-negative matrix factorization (`NMF`). An efficient algorithm is described in Lee and Seung, 2001, and the traditional implementation is criticized and discussed in Le Roux, Weninger, and Hershey, 2015. The optimization problem solved by `NMF` is simply, for $X = [x_1, x_2, \dots, x_k]$ positive vectors:

$$\min_{D,\Gamma} ||X - D\Gamma||_F^2 \text{ s.t. } ||\Gamma||_0 \leq \eta, \forall i \text{ and } D \geq 0$$

1.6 Input data

In addition to the images directly acquired in the compressed sensing setting (which will be presented in details in Section 3.3.3), the following datasets were used.

1.6.1 Simulations

Noiseless images We first performed simulations on ideally sparse images. These images consist of a few non-zero, uniformly distributed pixels on top of a dark, zero-valued background, as depicted in Figure 6a. The non-zero pixels have a mean value of ~ 100. Of course, these images are very far from a real life images (such as Figure 8b). However, with such images, we aim at simulating the output after application of a sparsifying transform, thus operating in the intermediate domain described above and in Section 6.4.

Noisy images To first assess the sensitivity of the reconstruction algorithms to noise, we produced noisy counterpart of each noiseless image. To do so, Gaussian, additive noise and Poisson noise were added to the noiseless signal. The magnitude of the Gaussian noise was progressively varied in order to produce a SNR (signal-to-noise ratio) varying between 0.2 and 10.

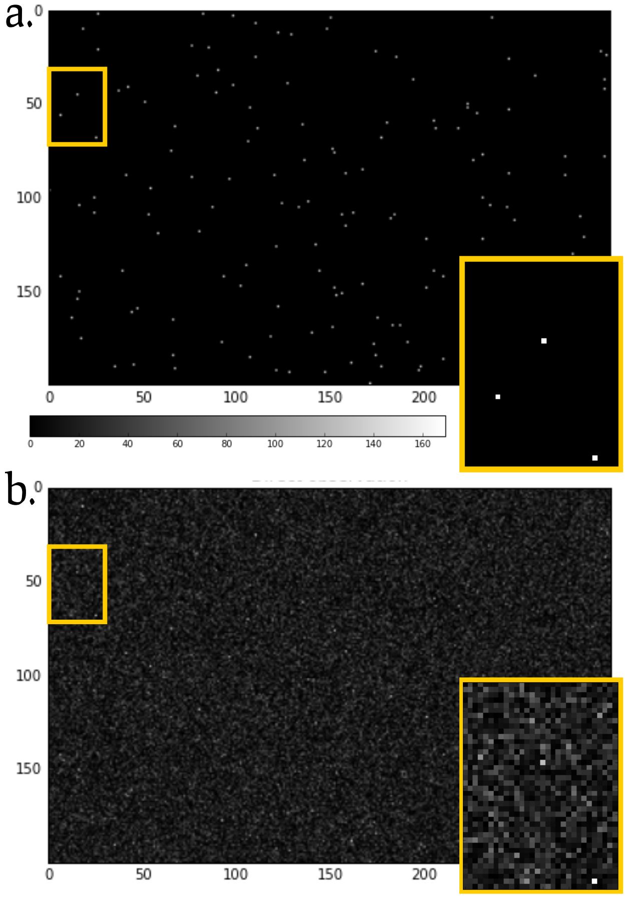

Figure 6: Simulations of ideally sparse images **(a).** noiseless image, **(b).** same image with noise added, with a SNR of ~ 1. SNR is reduced due to the presence of Gaussian additive noise, and the signal is also contaminated by Poisson noise, **(inset)** zoom on a 40×30 pixels region containing 3 non-zero elements.

1.6.2 Beads

As a first implementation step, the first acquisitions were performed on fluorescent beads. When scattered on a coverslip, the beads provide very bright, high signal-to-noise ratio signal and most importantly show little photobleaching (Figure 7). Such signal is ideal to calibrate the method for two reasons:

- The bead signal is very stable, thus the camera will measure exactly the same object over the acquisition.
- Beads are point/PSF-like features, which should be near-optimally acquired using a Fourier basis.

1.6.3 Live samples

As a biologically-relevant sample, developing fly embryos were imaged by Mustafa Mir in 3D over time. The strain used for imaging expresses a fluorescently labelled version of the DNA-associated protein histone-2B (H2B fused to a green fluorescent protein or H2B-GFP). This allows to follow the chromosomes and the cellular divisions during the early development of the fly embryo.

Early development of the fly embryo is a highly dynamic process, where synchronized cellular divisions occur every ~ 15 minutes. Furthermore, during these divisions critical developmental events occur during which short-lived, embryo-scale protein gradients form and progressively refine the morphology of the embryo (Garcia et al., 2013).

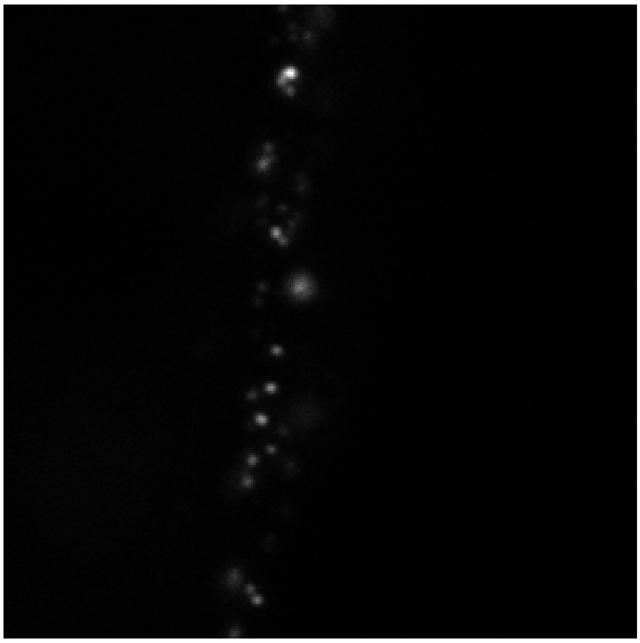

Figure 7: Sample images of beads acquired in the direct, uncompressed mode. This `256x256` pixels image acquired using a 10 ms exposure time and is extracted from a 3D stack. Beads seem to be lined vertically: this corresponds to a profile view of the coverslip. Off-focus beads appear with a halo whereas in-focus beads appear as sharp and intense spots. Pixel size: 100 nm (whole picture: 25.6 x 25.6 µm).

So far, most of the obeservations on the fly embryo are performed using a confocal microscope, a technique that achieves sectioning by scanning a conjugated pinhole over the whole sample, resulting in a high dose of light at sample and photodamage. Such high intensity at the sample is usually a limiting factor for imaging and overexposed flies tend to develop developmental defects in later division stages. Thus, techniques such as the lattice light sheet microscope combined with a reduced exposure time thanks to a compressed sensig acquisition mode could have valuable and direct applications.

Finally, the choice of a *H2B*-labelled fly is a particularly interesting sample because it features two visually and functionally highly distinct chromatin states: interphase and mitosis:

- during interphase, the chromosomes are decondensed and no individual chromosome is visible,

- during mitosis, the fly chromosomes are condensed and form clearly distinguishible "blobs" with a high contrast with respect to the background. These chromosomes can be followed over time, allowing to probe for the spatial consistency of the reconstructions.

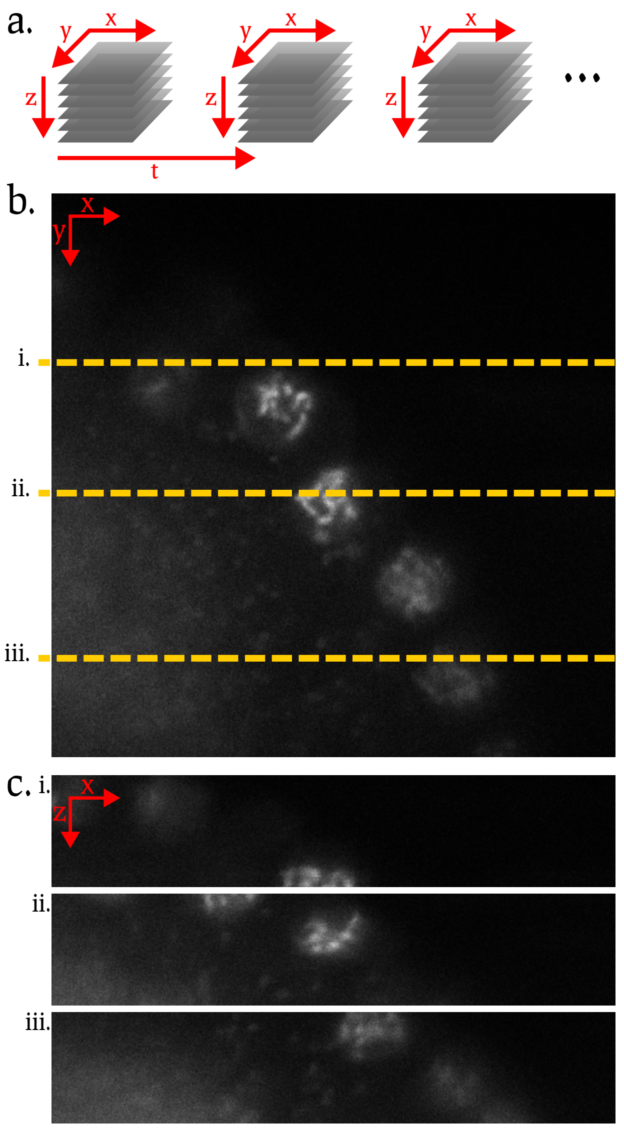

Figure 8: Subregion of a developing fly embryo during mitosis imaged with the lattice light sheet microscope. At that developmental stage, nuclei stand at the surface of the embryo and cellular divisions are synchronized. **(a).** scheme used to perform the acquisition. At each time point, a series of images is taken in the (x, y) plane and at various z positions, forming a 3D-stack, **(b).** sample image in the (x, y) plane, four dividing nuclei are clearly visible (circular features on the first diagonal of the image). inside those nuclei, condensed chromosomes are clearly visible. Yellow lines mark the positions in y of the sections presented in **(c).** sections in the (x, z) plane of the image presented in panel (b). Image width: 51.2 µm. Images by Mustafa Mir.

Fly embryos were prepared for imaging and imaged using the lattice lightsheet microscope. `512x512x101` pixels 3D stacks were acquired at a few tens of seconds time step, 101 2D images (parallel to the (x, y) plane) were acquired per 3D stacks with an exposure time of 100ms/2D plane. A sample image is shown in Figure 8b and a view of an (x, z) plane is shown in Figure 8c.

(x, y) resolution is approximately 300 nm. axial, z resolution is approximately 400 nm. 100 to 300 planes from the 3D stack are located approximately 200 nm apart, in order to fulfill the Shannon-Nyquist criterion, where the PSF bandlimits the signal. Images were acquired as 16 bits/pixel, resulting in sevral gigabytes of data per movie (approximately 100 MB per 3D stack).

After acquisition, these images were used as a reference to eveluate the performance of compressed sensing algorithms. To do so, the images were compressed in z (see details in Section 2.1) using a predefined measurement matrix and additive Gausssian and Poisson noise were added to the compressed images. After decompression using a compressed sensing reconstruction algorithm, the images could then be compared with the original reference.

2 Results – Simulations

Prior to the implementation of a physical setup to acquire compressed images, we perform a series of simulations and computations in order to demonstrate the theoretical feasibility of such system under realistic speeds and noise levels.

2.1 Several schemes can be adapted to the lattice lightsheet microscope

To find a scheme that can actually be implemented into a physical system is a traditional challenge in the compressed sensing field. Indeed, contrary to acquiring a big dataset and then compressing it, to acquire data in a compressed manner requires to design a sensor/sensing mode that can actually sense in a basis that is incoherent with the signal naturally sparse basis. Such a challenge is a major limitation in imaging, where the signal-to-noise ratio is usually limited (Willett, Marcia, and Nichols, 2011).

First, achieving compression in the (x, y) domain seems hardly achievable, and there seem to be no easy solution for the lattice lightsheet since the whole camera sensor is conjugated with the sample. Indeed, in such a setting, each individual (x, y) position on the sample is uniquely mapped to a (x, y) position of the detector, preventing us from acquiring an incoherent linear measurement in this plane. Note that if we were using a laser scanning confocal microscope, compression in the (x, y) plane could have been possible by the use of a digital micromirror device – DMD, see Ye et al., 2009.

However, for 3D acquisition, one can imagine a compression in z, the third dimension, allowing to perform compression in one dimension:

2.1.1 Implementing a 1D compression...

Indeed, one can design a scheme in which the third dimension (z) of a 3D stack is acquired in a compressed fashion. Such an acquisition is possible thanks to two features of the microscope:

- The z-piezo, that allows the user to manually command the z position of the stage/sample.
- The AOTF that allows to precisely tune the light intensity.

It is then possible to design a scheme, depicted in Figure 9, where during each measurement (that is, camera frame) the stage sweeps through the whole sample while the AOTF shows a specific, user-defined pattern: a vector of the measurement matrix. For instance, assuming a Hadamard transform as the measurement matrix, Figure 10 jointly shows the position of the stage in z and the illumination intensity.

Note that even though the data acquired is still 3D, the compression only operates in the z dimension and at each (x, y) position corresponds a 1D compressed signal. This signal can be seen as independent from the other adjacent 1D signals. They thus can be reconstructed independently, allowing for a massively parallel reconstruction scheme.

Furthermore, note that in this dimension, one has an almost absolute freedom of choice with respect to the measurement matrix, the only constraints that apply are the following:

- the measurement matrix has to be positive,
- its variation in intensity should be compatible with the AOTF speed
- it should not have a higher resolution than the PSF.

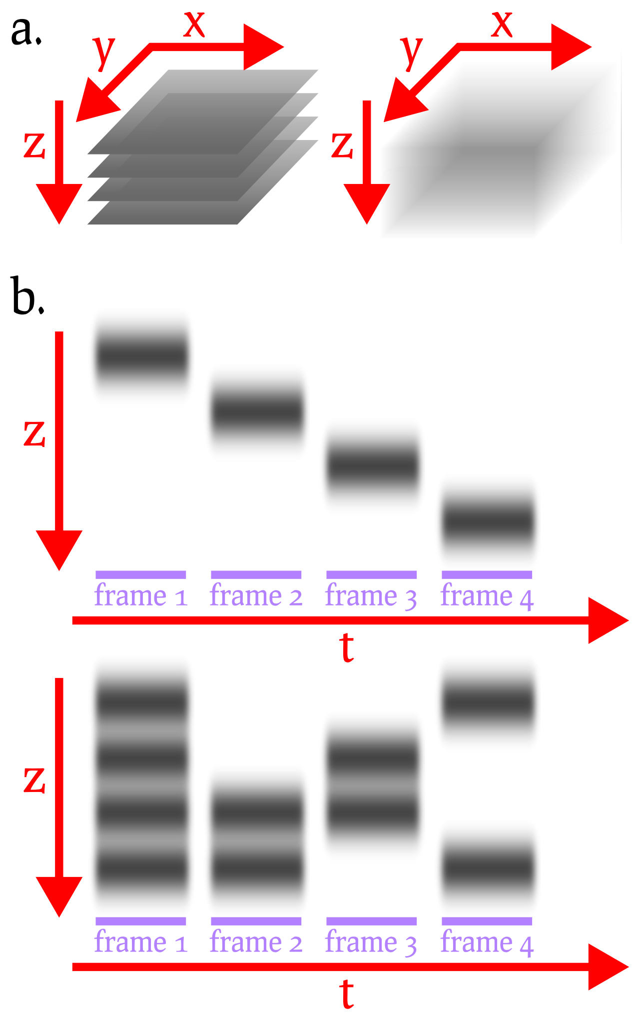

Figure 9: Comparison between the traditional mode to acquire 3D volumes and the compressed sensing mode. **(a).** in the traditional imaging scheme (left), each camera frame contains the information from a given z position whereas in the compressed sensing mode (right) each camera frame conveys part of the information over the whole z stack. Such spread of the information reflects the incoherence property of the measurement matrix, **(b).** comparison of the two imaging schemes: (top) standard scheme, at each camera frame only one position is illuminated (bottom) at each camera frame, a selected number of frames is illuminated with a given light intensity in a way that the final intensity pattern forms an incoherent basis. Here a Hadamard transform has been pictured as the measurement matrix.

Regarding this last point, we know that the signal (such as the image of a sample) coming through an objective is bandlimited (see Kim, 2011). Thus, the conditions to apply the Shannon-Nyquist theorem are fullfilled and one can fully sample the signal at a frequency twice the highest frequency of the signal, that is at a step of half the full-width at half maximum of the PSF. Thus, there is no need to apply measurement matrices that have a resolution higher than the order of the PSF width.

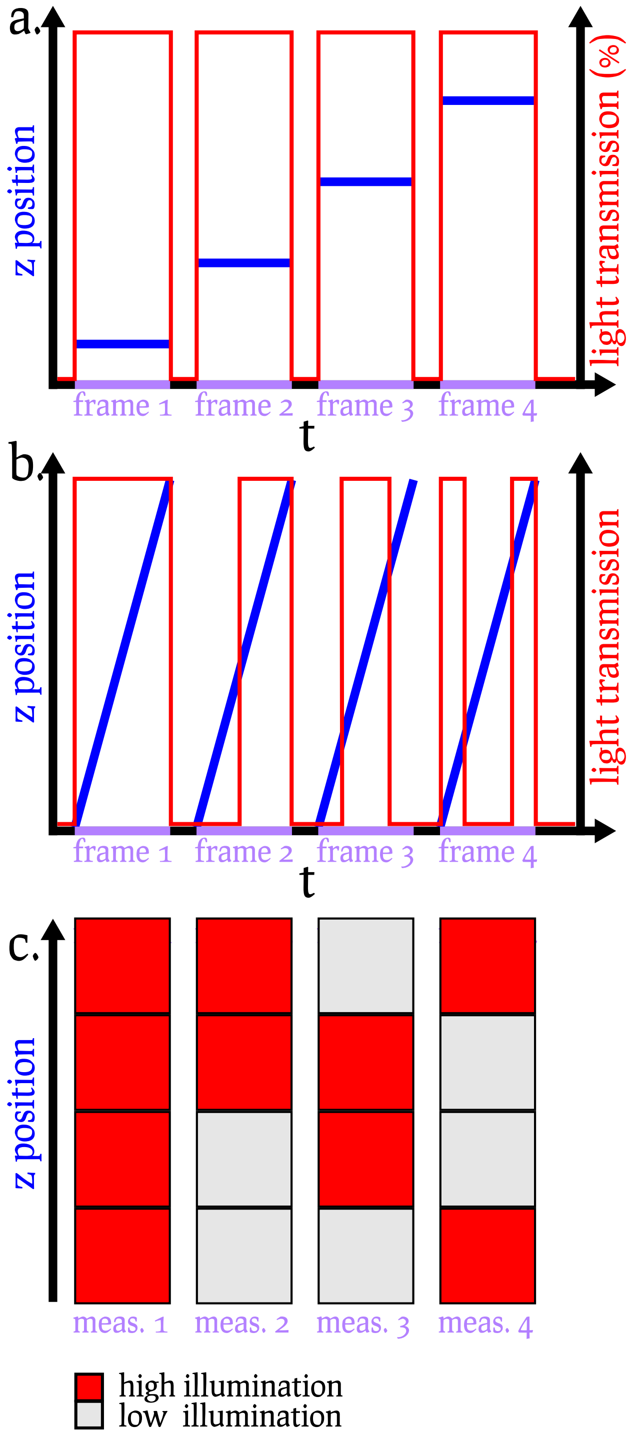

Figure 10: Comparison of the implementation of the traditional imaging scheme with a compressed sensing mode (using a Hadamrd transform as the measurement matrix). **(a).** traditional imaging scheme. To image a 4-plane 3D stack, the camera successivelly exposes four times at four different positions, **(b).** implementation of a Hadamrd transform in z: during each camera frame (taking the same camera exposure time as the traditional scheme), the z-piezo scans the whole sample. At the same time, the AOTF modulates the light intensity. The light pattern is such that a Hadamard transform matrix is applied as a measurement matrix (depicted in **(c).**).

2.1.2 . . . seems compatible with the hardware setup. . .

Let us review now the physical constraints that might apply to the setup and possibly hinder its practical implementation. In order to apply a measurement matrix with comparable exposure times as the standard scheme (for instance, 20 ms/frame over a 50 µm thick sample, that is 150 frames):

1. The AOTF should be able to closely follow a pattern (it should be fast and accurate enough). Ideally, its update frequency should be $>$ 1000 times the exposure time, that is 50 kHz,
2. The z-piezo should be able to scan fast enough (50 µm in 20 ms, and thus oscillate at 25 Hz)
3. Synchronization between the z-piezo and the AOTF should either be extremely good (tens of µs) or accurately measured in order to be software corrected.

A quick review of the datasheets shows that such a compressive scheme seems to be compatible with the hardware isntalled on the lattice light sheet. A more extensive analysis of the hardware limitations of the setup is detailed in section 3.3.3.

2.1.3 . . . and allows the implementation of a wide range of sensing matrices

A very interesting feature of this imaging scheme is that almost any (positive) measurement matrix can be applied to the sample.

However, the positivity constraint is an important restriction. Indeed, many of the traditional sensing matrices cannot be applied to the imaging. For instance, Gaussian measurement matrices have negative values and thus are not suitable for imaging. Furthermore, although it is always possible to offset a matrix such that it has no negative coefficient, this usually strongly degrades the theoretical properties of the matrix. Indeed, adding an offset w to a matrix A is equivalent to translating all the measurement vectors by the vector $w \times (1, 1, \dots, 1)$. As w grows, the initial values of A become negligible and thus all the columns of the matrix become more and more colinear. This effect can also be seen by recomputing the RIP in the presence of an offset and to show that the measured sparse vectors are no longer orthogonal.

Bernoulli and uniform basis Fortunately, the so-called Bernoulli basis (where each entry of the matrix follows a Bernoulli law of parameter $p \in [0, 1]$) and uniform basis (each entry follows a uniform law: $a_{ij} \sim \mathcal{U}([0,1))$) are matrices that provide good guarantees.

Fourier basis Another approach consists in taking a truncated and offset discrete Fourier transform (in an approach similar to E.J. Candes, J. Romberg, and T. Tao, 2006). An interesting property of this approach is that a simple transform allows to switch from the offset Fourier basis to the traditional Fourier signal. Then, the compressed signal can be seen as a bandlimited Fourier series decomposition, thus providing a guarantee on the reconstructed signal: the reconstruction from the algorithm cannot be worse than the signal recovered from the bandlimited Fourier transform. From an empirical standpoint, the worse reconstructions we could achieve with the compressed sensing solver were always of better quality than taking the inverse Fourier transform from the bandlimited signal. Furthermore, Fourier matrices show minimal coherence with the Dirac basis, making it particularly suitable to observe Dirac-like features in the direct domain. Due to these theoretical guarantees, unless specified, acquisitions and simulations are realized with a Fourier basis.

Expander graphs Yet another orthogonal approaches can also be undertaken. It relies on the construction of sensing matrices derived from *expander graphs* (Raginsky, Willett, et al., 2010, Raginsky, Jafarpour, et al., 2011), that show interesting properties for imaging. However, after a closer look at their design, they seem to be answering another question. Indeed, they ensure a useful energy-conservation property (the energy received by the sample is constant for each measurement). Unfortunately, when modulating the laser power with the AOTF, the energy at the sample is no longer constant, making schemes derived from such graphs are less suitable for implementation in the lattice light sheet.

Matrix optimization Finally, instead of postulating a measurement matrix, one can also derive it from the dictionary used for the reconstruction. Such an approach ensures that the couple (sensing matrix, dictionary) has maximal incoherence, giving the best theoretical guarantees. Such a method is called *matrix optimisation* and several methods have been proposed (Elad, 2006) to achieve it. A recent method performs matrix optimization under positivity constraint (Mordechay and Schechner, 2014).

However, although promising, we were unable to obtain optimized sensing matrices using the code provided by these last authors.

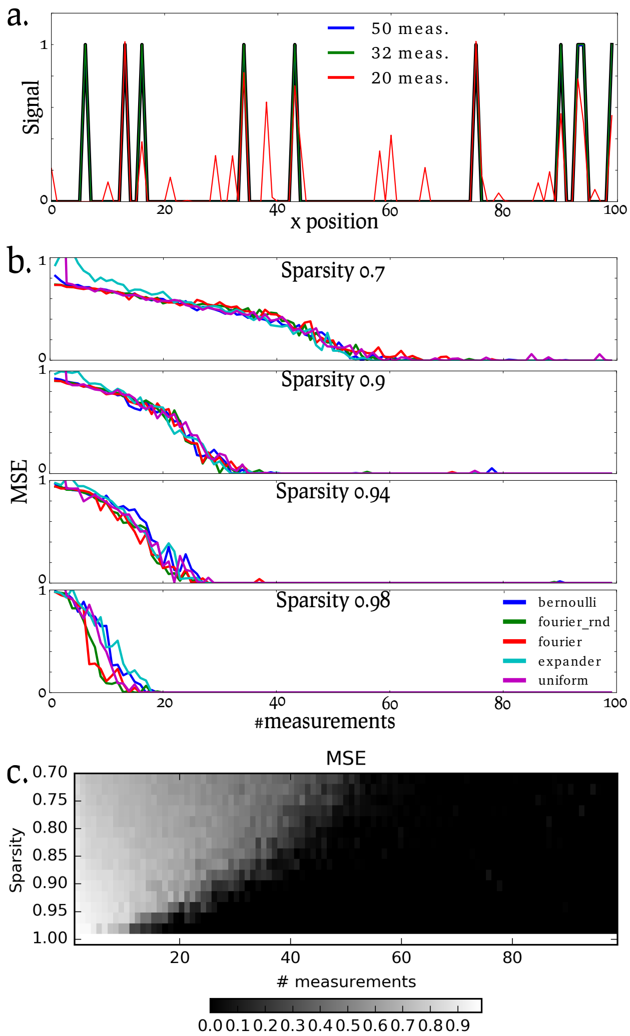

Figure 11: Sparse signals can be recovered in a noiseless setting. Reconstructions are performed on a 1D vector of length 100 whose sparsity is varies from 70% to 98%, **(a).** Sample signal of sparsity 90% (black, 10 nonzero components) and its reconstruction from a decreasing number of measurements using a Bernoulli measurement matrix. Reconstruction is exact for 32 and 50 measurements. With 20 measurements, significant errors are visible. **(b).** mean square error (MSE, computed as $\frac{\sum_{i=0}^{100}(x-\hat{x})^2}{\sum_{i=0}^{100}x^2}$) as a function of the number of measurements for various measurement basis (colors) and four different sparsities of the input signal (top: 70%, middle: 90%, middle again: 94%, bottom: 98%), **(c).** a phase transition is visible when the ℓ_2 reconstruction error is plotted as a function of sparsity and the number of measurements, gray level represent the MSE. Reconstruction made with SPIRALTAP (although the results are similar using a different algorithm). Panels (b) and (c) represent the average over 10 replicates. **bernoulli**: Bernoulli random matrix used as measurement matrix, **fourier**: Fourier matrix (10 first basis vectors), **fourier$_{\mathbf{rnd}}$**: 10 vectors from a Fourier basis, sampled at random, **expander**: measurement matrix based on expander graphs, **uniform**: uniform random matrix.

In Figure 11, we present on a toy reconstruction problem without noise the performance of several measurement basis at various sparsities of the input signal and the number of measurements. It is clear that the sparser the signal the better the reconstruction, and obviously, the higher the number of measurements (and thus the lower the compression factor) the lower the mean square error (computed as $\frac{\sum_{i=0}^{100}(x-\hat{x})^2}{\sum_{i=0}^{100}x^2}$).

Several conclusions can be drawn from this first simulation:

- Note that a *phase transition* occurs in the bottom panel (Eldar and Kutyniok, 2012, chapter 7), that is there is a clear threshold at which the reconstruction is exact ($MSE < 10^{-10}$),
- High compression ratio can be achieved, provided that the signal is sparse enough. For instance, it seems that a 10 times compression can be achieved provided that the signal is $\geq 95\%$ sparse. In the following, we will stick to a compression ratio of 10 as a target.
- In such an artificial setting, it is difficult to assess which measurement matrix will be the most suitable for imaging.

2.2 Gaussian additive noise has a limited impact on reconstruction

We now turn to a case where we apply noise to the measured signal and evaluate the sensitivity of the reconstruction algorithm. We start from the same setting as in the previous section and add additive Gaussian noise in order to match a predefined SNR (defined as $\rho = \frac{\mathbb{E}(X)}{\sqrt{\mathbb{V}ar(X)}}$).

Note that since the compressed sensing noise applies in a domain that is incoherent with the original sample, it might be difficult to calibrate the SNR. However, due to the fact that this specific noise is additive, an analytical relation can be computed to match the SNR in the direct and the projected domain.

Indeed, assume a given SNR ρ in the direct domain for a signal $\tilde{X}$ of length n contaminated by additive gaussian noise of standard deviation σ, and compute it as $\rho = \frac{\mathbb{E}(X)}{\sqrt{\mathbb{V}ar(X)}} = \frac{\mathbb{E}(X)}{\sigma}$, that is:

$$X = \tilde{X} + \sigma N, \text{ with } N \sim \mathcal{N}(0,1)$$

Trivially, one has $\mathbb{V}ar(X) = \sigma^2$. Now, consider the variance of the measured signal $Y = AX$ (with an neat abuse of notations regarding that N is a Gaussian vector):

$$Y = AX = A(\tilde{X} + \sigma N) = A\tilde{X} + \sigma NA$$
$$\implies \langle \mathbb{V}ar(Y) \rangle = \frac{\sigma^2}{m^2}\sum_{i=1}^{m}\sum_{j=1}^{n} a_{i,j}^2 = \frac{\sigma^2}{m^2} tr(A^T A)$$

This implies that for a signal of mean value $\mathbb{E}(X) = \mu$ in the direct domain and to achieve a given SNR of ρ, one can, as a first approximation add noise in the projected domain with a variance $\frac{\sigma^2}{m^2} tr(A^T A)$.

Figure 12 presents the empirical performance of one reconstruction algorithm in noisy conditions. In contrast with Figure 11 the chosen vector has length 200. Several conclusions can be drawn:

- it is possible to reconstruct accurate signals in noisy conditions (SNR $<$ 10),
- as expected the MSE decreases with the number of measurements,
- furthermore, when the MSE is plotted as a function of the noise term σ instead of the SNR, the error exhibits a linear scaling with σ, compatible with the guarantees of reconstruction under additive noise.

This analysis, although useful to demonstrate the applicability of the method, do not give any clue on how it compares with noisy images acquired in the direct domain. To perform such comparison, we have found that relying only on means square error sometimes provided unfair results (often in favour of the compressed sensing approach). Indeed, many criticisms have been expressed against MSE in the imaging field (see for instance Zhou Wang and Bovik, 2009). We thus try to present more *ad-hoc* metrics when needed, altogether with the raw images for visual inspection.

To assess the relative impact of noise in the direct domain vs. noise in the transformed domain, we proceed as follow (principle of this simulation is depicted in Figure 13):

1. A noiseless 2D image is generated as described in section 1.6.1.
2. For a given SNR, either additive Gaussian noise is added in the direct domain, or the corresponding

noise level is set in the projected domain.

3. The projected image is decompressed and the output of the algorithm is compared with its counterpart with a matched noise in the direct domain.

4. The process is repeated using various SNR

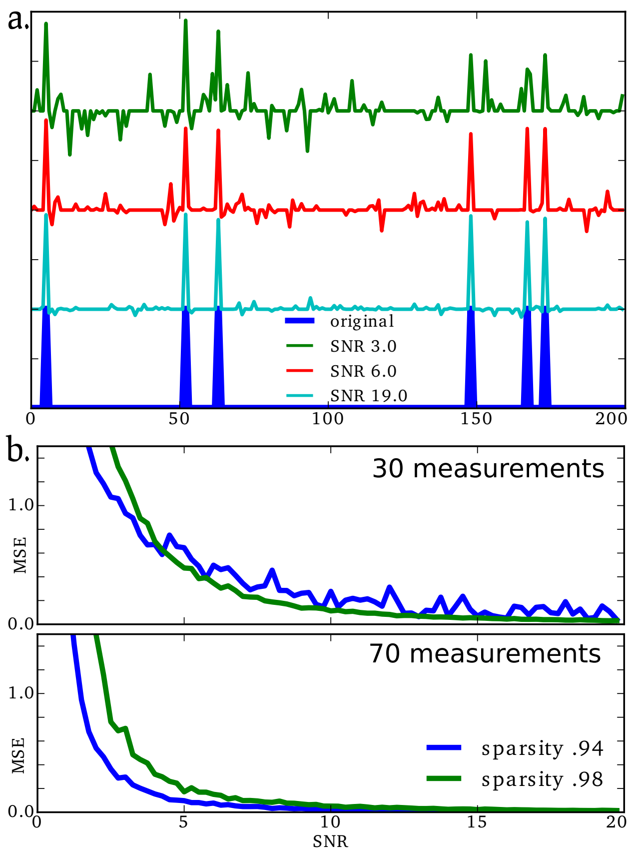

Figure 12: Performance of the ℓ_1-magic reconstruction algorithm under noisy measurement. SNR correspond to the equivalent noise level in the direct domain. **(a).** Sample reconstructions of a signal of sparsity .97 (bottom, blue) when measured with increasing SNR (green, red and azure curves), **(b).** Evolution of the mean square error as a function of the SNR in the direct domain for signals of different sparsities and various number of measurements. SNR is computed as the peak-SNR (PSNR).

The results of the simulation are presented in Figure 14. We used two metrics. The first is simply the PSNR (peak signal-to-noise ratio, defined as $PSNR(x, \tilde{x}) = 10 \log_{10} \frac{\max(x)^2}{MSE(x,\tilde{x})}$). PSNR is expressed in decibels (dB). Using this metric, one can see that the compressed sensing approach seems to clearly outperform the blurry result by significant margin (several decibels), and the higher difference aappears in low-SNR regions, which precisely are the region of interest in the case of fast imaging.

However, and maybe obviously, this approach is very unfair to the blurry image. Indeed, as seen above in Section 1.3, compressed sensing acts as a *denoising* routine, thus achieving artificially high PSNR on sparse samples, whereas an image contaminated by Gaussian noise will exhibit very low PSNR.

To tackle this issue, we propose a different metric for the specific purpose of this experiment. Since the input image is binary, we seek how many non-zero pixels are accurately recovered. A simple way to define a recovered pixel is as follows: assume there are ν non-zero pixels in the original image. In both the noisy images, take the ν pixels of higher intensity and see how many of them matches the non-zero pixels in the original image. The results of this approach are depicted in Figure 14b.

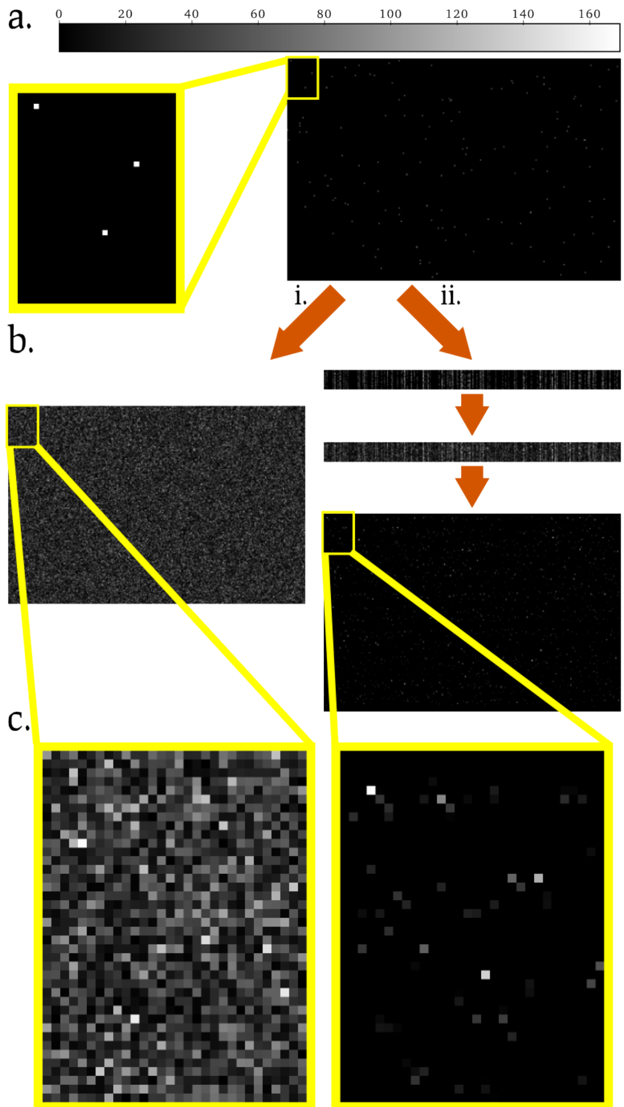

Figure 13: Procedure to compare the effect of noise applied in the direct vs. the projected domain: **(a).** a binary 2D image is generated. Then **(b).** two mutually exclusive transformations are applied, either **(i).** Gaussian noise is applied in the direct domain or **(ii).** the image is compressed ten times in one of the dimensions and a matched additive Gaussian noise is applied before reconstruction. Finally **(c).** the images resulting from the two processes can be compared, either on their general appearance or on selected subregions. Here a 10 times compression has been applied. matched SNR is 1.

Note that this approach is now more fair to the addition of Gaussian noise. Indeed, we consider that if the pixel has a higher intensity than the background, then it can be recovered. Note that since the features we want to recover are isolated pixels, the selection of the pixels of maximum intensity is the best guess one can have (this is a maximum likelihood configuration).

Now, this method slightly disadvantages the compressed sensing approach. Indeed, it might happen that the reconstructed image shows exactly the right pattern, but one non-zero pixel appears slightly blurred over two adjacent positions, leading to one of the two adjacent pixels of high intensity to be counted as a false detection. This is particularly visible in the high-SNR portion of the curve, where the reconstructed curve never reaches 100% for this specific reason.

Finally, in the low-SNR region, a very interesting fact appears: the compressed sensing method captures more spots than the acquisition in the direct domain, suggesting that compressed sensing could show competitive results with highly noisy signals where traditional imaging performs poorly.

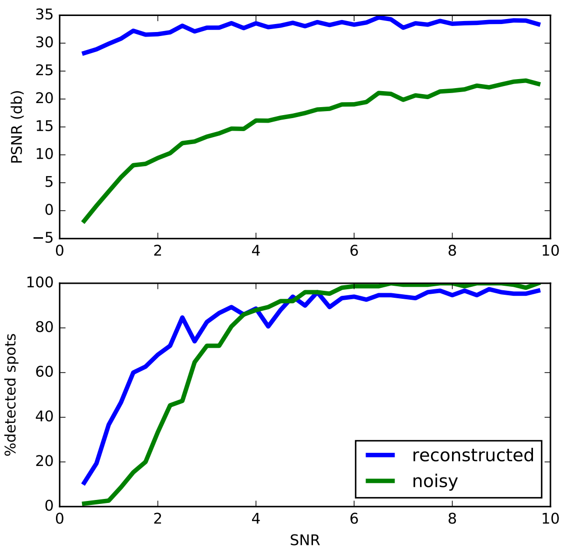

Figure 14: Performance of compressed sensing reconstructions compared with traditional imaging under noisy conditions. Two sets of noisy images with various SNR are generated, as described in Figure 13. **(top)** PSNR as a function of the SNR, the blue curve is the compressed sensing reconstruction and the green curve imaging in the direct space, **(bottom)**. Same experiment, but the metric is now the proportion of bright pixels actually recovered. Reconstructions of a 10 times compressed signal.

Conclusion Using these simple simulations on binary images, we were able to get a working example and competitive reconstructions in noisy settings, possibly including where direct imaging is made difficult by the noise. These binary images are very far from biological samples, but we chose such simplified images as proxies for of an intermediate representation when an *ad-hoc* dictionary has been learned. In the next section (2.3), we turn to optimizing compressed sensing together with dictionary learning.

2.3 Dictionary learning improves reconstructions

Recall Figure 3. Biological images do not match the sparsity requirements we exhibited in our previous simulations: although biological images are at most ~ 60 % sparse, our simulations showed interesting results for >95% signals. This remark justifies the use of dictionary learning to find an *ad-hoc* sparsifying transform.

2.3.1 Dictionary learning provides significantly improved performances over inverse Fourier transform

In the case where the measurement matrix is a Fourier basis, there are two, non-equivalent ways to reconstruct an image:

1. Use a compressed sensing reconstruction algorithm, as described above,
2. Set the unmeasured coefficients to zero and perform a traditional inverse discrete Fourier transform.

The latter method can formally be seen as a lower bound on the quality of reconstruction. We thus compare the performance of a compressed sensing, with and without dictionary learning with respect to a reconstruction derived from inverse discrete Fourier transform, as depicted on a sample image of a dividing fly embryo in Figure 15.

Traditional MSE are presented together with sample reconstructions and absolute errors for visual inspection in Figure 16. Several conclusions can be drawn:

First, the bandlimited signal reconstructed by inverse Fourier transform clearly is a degraded version of the original image, as shown by the MSE between the iFFT reconstruction and the original (Figure 16a, left).

Second, even without a dictionary, the compressed sensing reconstruction algorithm outputs an image relatively similar to the original, having its high frequency features adequately resolved. Interestingly, this reconstruction ends up with a very high MSE. Indeed, the global background is not properly reconstructed, which leads to an accumulation of small error ters, dramatically increasing the MSE. Furthermore, this method introduces several artifacts that appear as shadow chromosomes (that is chromosome-like features where there was nothing in the initial image). These observations lead us to always proceed with a visual inspection of the results rather than relying only on the MSE.

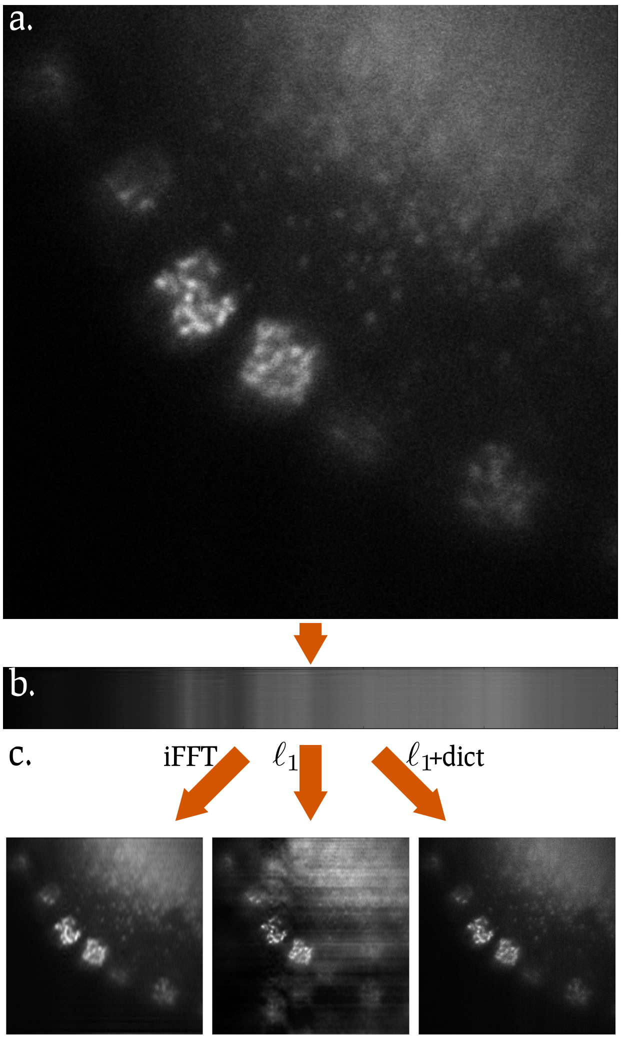

Figure 15: Pipeline to assess the performance of dictionary learning. **(a).** original image used for reconstructions. Note the feature-rich region of the nuclei, where individual chromosomes are visible, and the top-right region with less defined characteristics, **(b).** compressed version of the image. A 10x compression is applied in the vertical dimension using a Fourier matrix. 90% of the data is thus discarded, and **(c).** three reconstruction algorithms are applied: **(iFFT)** the discarded coefficients are assumed to be zero and an inverse Fourier transform is performed, **(ℓ_1).** traditional compressed sensing reconstruction is applied or **(ℓ_1 + dict).** a dictionary is used for reconstruction. The dictionary has been learned on a set of ~ 10 planes coming from this time point and another time point further in time. Field dimensions: 51.2 x 51.2 µm. 100 planes were acquired with a ~ 200 nm step.

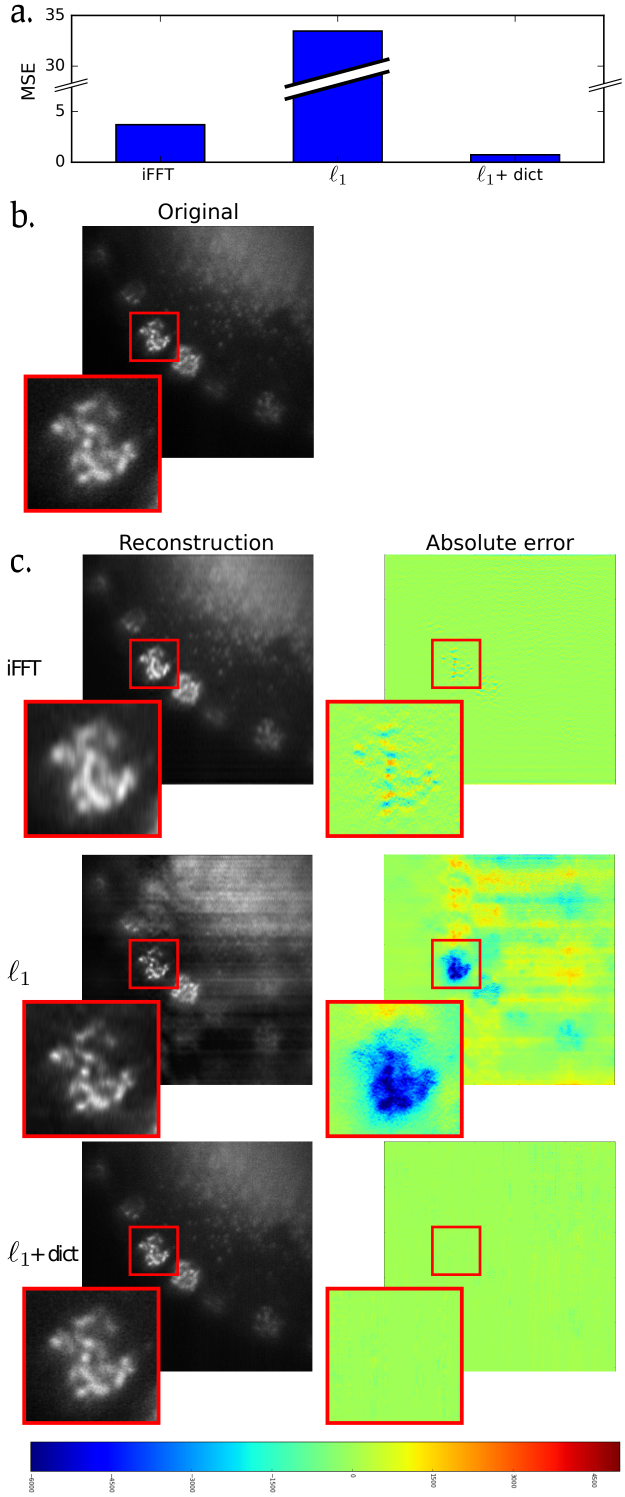

Figure 16: Assessment of the benefits of dictionary learning. **(a).** Mean square error of three different reconstruction techniques, following the protocol detailed in Figure 15, **(b).** sample image used for this figure and zoom on a features-rich region (nucleus), **(c)-left.** sample reconstructions using the three methods and **(right).** absolute error. The colorbar has the same amplitude as the input image. The compressed sensing images were obtained using ℓ_1-magic. The dictionary used for reconstruction had size `1024x512` (two times overcomplete dictionary). Dictionary learning algorithm: `NMF`.

Finally, the reconstruction with a dictionary achieves by far the best result, with a MSE dramatically lower than the other version (corresponding to several dB), capturing most of the features of the original signal. Furthermore, in contrast to the two other reconstructions, dictionary learning show little or no detectable pattern in the absolute error (the 2D image $X - \hat{X}$), which indicates that little artifacts are introduced by the method, despite the fact that its MSE is usually considered as high.

In this first simulation, we did not investigate the effects of tuning the dictionary learning parameters. In the next section, we evaluate both the impact of the parameters for dictionary learning and the robustness of the outputs.

2.3.2 Robustness of reconstructions performed with a learned dictionary

A justified (and inavoidable) drawback of dictionary learning is that it is a data-dependent procedure, thus subject to overfitting. Intuitively, features that were part of the training set are much more likely to be accurately reconstructed than features resembling nothing like the training set. Then, one wants to assess the robustness of the reconstruction when dictionary learning is applied.

Here, we evaluate the evolution of the MSE for reconstructions performed along several axes of a 3D+time movie, as emphasized in Figure 17. First, a dictionary is learned by picking sample images from one particular 3D frame. We then evaluate the evolution of the MSE as we move farther from the sample images, either in time (reconstruct 3D volumes farther in time), in space (pick planes on the same stack but farther in z), or in a combination of both.

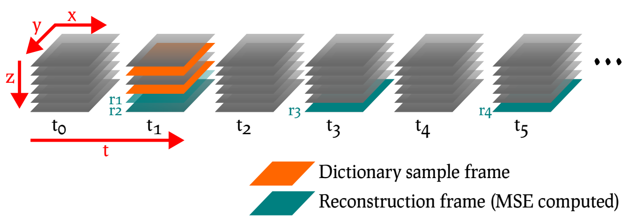

Figure 17: Procedure to assess the robustness of reconstructions performed using dictionary learning. For 3D+time acquisitions (represented as a series of 3D stacks), a specific timepoint is chosen to learn a dictionary (here, t_1) and some frames are used to learn the dictionary (pictured in *orange*). Then, several frames (pictured in *blue-green*) are used to assess the quality of the dictionary: the chosen frame is compressed 10 times and reconstructed using a Fourier basis using the ℓ_1-magic algorithm and the dictionary. A MSE is computed. Then, the algorithm moves to a frame farther from the sample frames used for the dictionary learning. Move can be either in space (from r_1 to r_2, switching z position), in time (from r_2 to r_3, changing timepoint) or both (moving from r_1 to r_3). The procedure is repeated for several dictionaries.

Results of this simulation are summarized Figure 18. Several points are worth noting:

- As the size of the training set increases, the overall MSE tends to decrease, but seems to reach a plateau for high numbers of samples. Similarly, MSE tends to decrease as the number of atoms increases.
- At the timepoint $t = 0s$, that is the timepoint corresponding to the training set for the dictionary learning, one can see a drop in the MSE at some of the frames used for the training.
- However, this effect remains highly localized (it spans very few frames, especially for dictionaries learned from a high number of samples) and it does not happen at every frame of the training set.
- Furthermore, for dictionaries built from ≤ 10 samples, the MSE is dominated by an increasing trend over the stack (reflecting the fact that the bottom of the stack is harder to reconstruct) rather than by dictionary-specific artifacts.
- Moreover, as one moves further in time ($t = 13s$, 5 3D stacks later), the drops corresponding to the training set are no longer visible, confirming the fact that the artifacts due to the closeness from the training set also quickly vanish in time.
- Finally, as one moves either in space of time, the worst MSE values are still in the acceptable range, and do not represent a dramatic deterioration compared to the frames close to the training set.

Conclusion In this subsection, we demonstrated that the implementation of dictionary learning at the same time robustly improves the reconstruction quality without adding too many artifacts. On the downside, the movie considered for reconstruction still shows similar features (mitotic chromosomes), and thus the robustness test do not assess for totally different features.

2.4 A pipeline for full 3D reconstruction

Once all these small scale simulations have been achieved, one can then scale the analysis in order to process acquired movies, which is a much higher amount of data than what has been processed for the previous

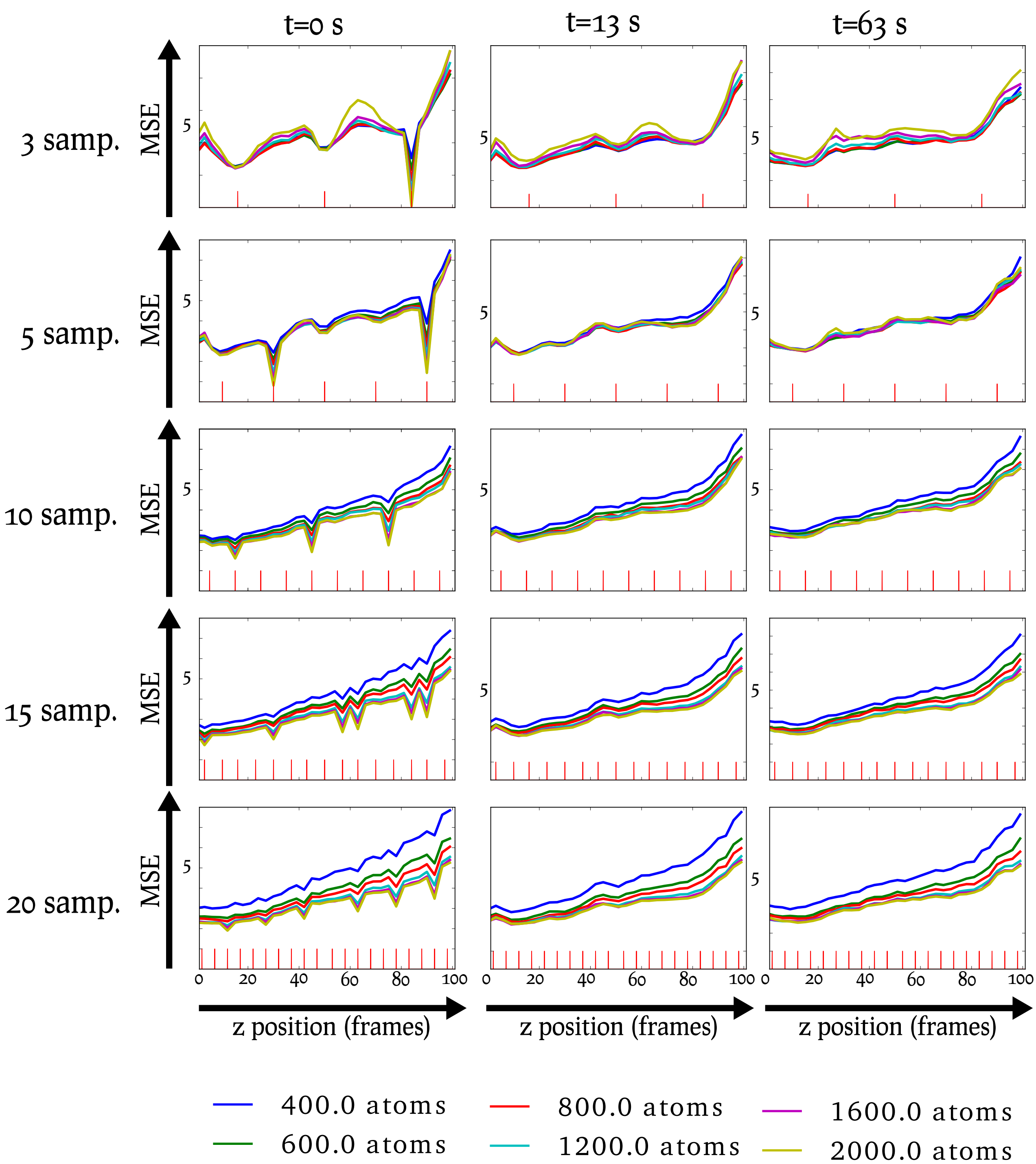

Figure 18: Robustness of the reconstruction when using dictionary learning. Each panel represents the mean square error (MSE) of a reconstructed frame at a specific timepoint **(columns)**, z position in the stack **(x coordinate)** and using a dictionary built from a given number of samples **(rows)** and with a given number of atoms **(curves)** . Dictionary is always built by taking frames from timepoint $t = 0s$. The frames used for dictionary learning are evenly split across the whole z stack, their position is marked by the red ticks. See Figure 17 for more details about the reconstruction and Figure 31 for sample images.

simulations. In order to deal with such scaling up, we implemented a simple processing pipeline. The final aim is to be able to reconstruct 3D+time movies in a reasonable time, with minimum user input.

2.4.1 Pipeline

To adress this need, we designed a simple pipeline. This pipeline can either run on a standard machine or on a HPC cluster powered with the *Simple Linux Utility for Resource Management* scheduler (SLURM, Yoo, Jette, and Grondona, 2003). It mixes `bash`, `Python` and `IPython/Jupyter` [2] scripts and achieves dictionary learning and reconstruction. An outline of the whole pipeline is depicted in Figure 19.

[2] IPython/Jupyter is a popular web-based frontend for `Python` (and many other languages). It can be downloaded from `http://jupyter.org`.

Due to the structure of the problem, it is possible that a parallelization on GPU units/cluster be beneficial. Indeed, since each small scale, 1D z-section is reconstructed independently of the others, and that the measurement matrix is the same for each 1D section, for a `512x512x201` pixels image, >260 k parallel reconstructions are launched. This architecture seems particularly suitable to be ported on a GPU. However, we didn't have time to investigate such approach.

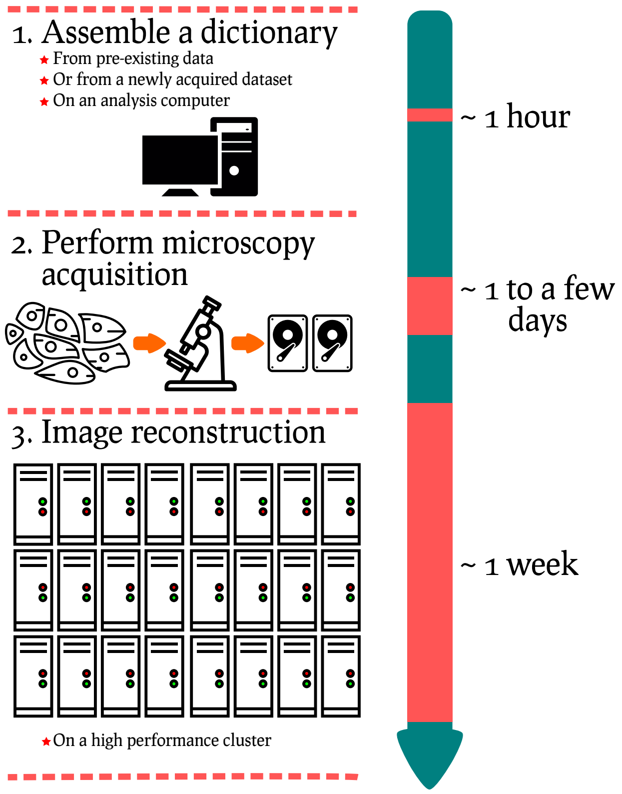

Figure 19: Pipeline to acquire and reconstruct compressed sensing images/movies. Dictionary can be assembled either from already available datasets or from a preliminary acquisition obtained in the direct, uncompressed domain. The whole pipeline can be run in one to two weeks.

2.4.2 Reconstructions

As a proof of concept of the pipeline, we artificially compressed and reconstruct a 50 3D volumes movie spanning one hundred and thirty seconds of the development of a fly embryo. Each 3D volume contains `512x512x101` voxels. Each 3D stack was compressed along the z axis by a factor of ten using a positive Fourier basis, leading to 50 3D volumes of dimensions `512x512x10` voxels. A 1200 atoms dictionary was learned from the 20 frames taken from the first 3D volume of the uncompressed acquisition. Reconstruction was eventually performed on a cluster located in Pasteur Institute, Paris, France[3] and took approximately four days, leading to a ~ 5 GB 3D+time `TIFF` movie[4].

Sample slices are presented in Figure 20 and some important comments can be made:

1. the reconstruction resolution is fairly consistent across the whole movie, further validating the use of the dictionary learning algorithm,
2. Feature-rich regions such as nuclei are accurately reconstructed. Moreover, their reconstruction is consistent both across space (moving in x or y) and time, which indicates that the reconstruction is robust.
3. In the (x, y) plane, the resolution seems to match the resolution of the input movie
4. On the downsides, in the z dimension, there seem to be a non-isotropic degradation in the resolution, especially in feature-rich regions (discussed further in Section 4.3)
5. In addition, the reconstruction exhibits in z some periodic components or *stripes* that contaminate the signal.

Conclusion In this subsection, we demonstrated the ability of a compressed sensing-based approach to recover 3D+time movies from a ten fold compression ratio using an artificially (software) compressed movie initially acquired with the lattice light sheet during a traditional microscopy experiment.

3 Results – Hardware implementation

Althoug the reconstructions presented above seem promising, the generated images were compressed post acquisition. Indeed, to our knowledge, there is no hardware that allows for compressed sensing microscopy, and a hardware implementation has to be developed to actually perform compressed sensing acquisitions.

[3]We used the computational and storage services (`TARS` cluster) provided by the IT department.

[4]Low resolution visualizations can be found on this report webpage.

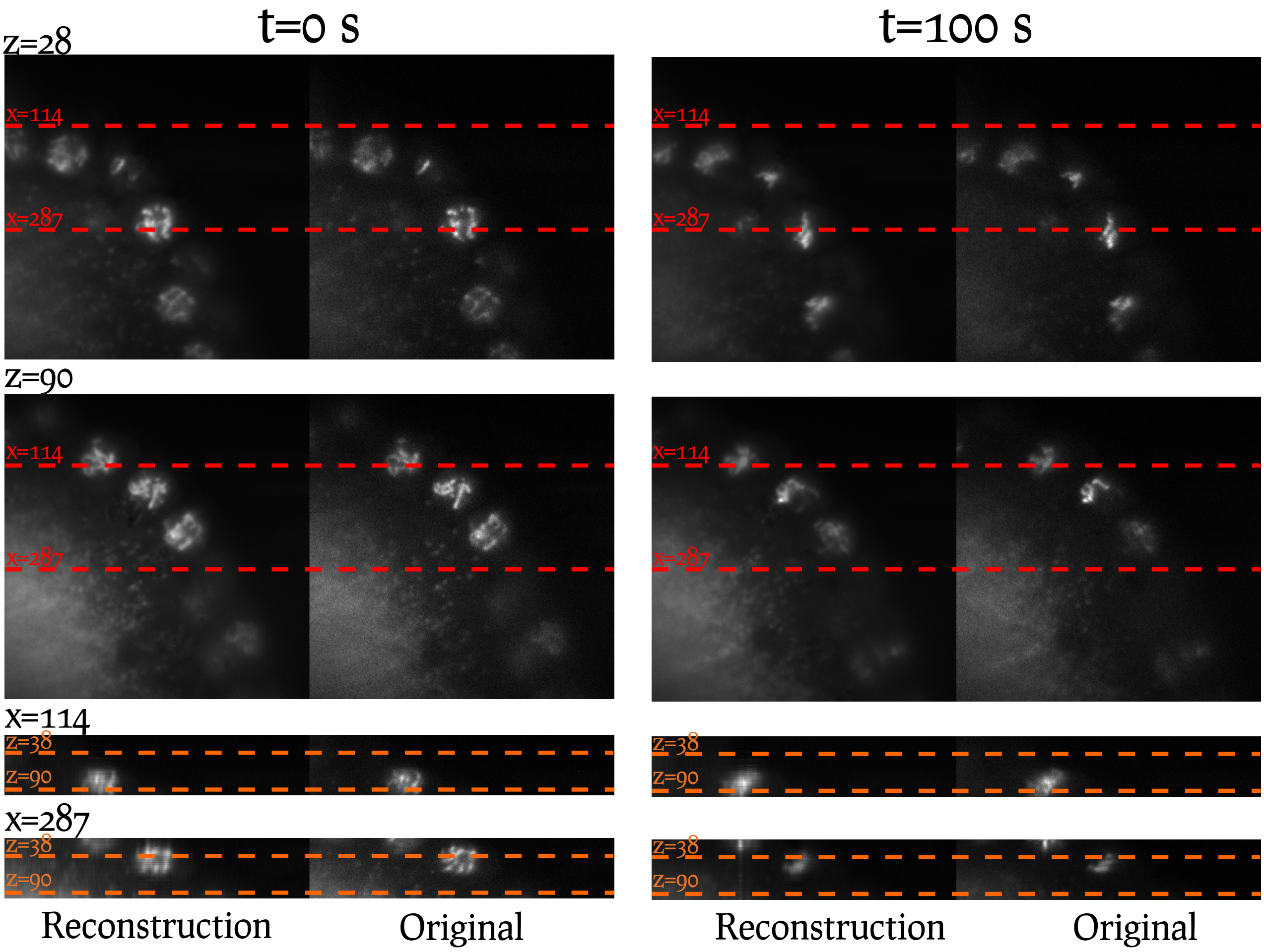

Figure 20: Sample planes from a 120 s reconstructed 3D+time movie of a dividing fly embryo. Two timepoints are presented, $t = 0$s (left column) and at $t = 100$s, which correspond to a 40 3D volume time gap. For each plane the left picture is the reconstructed image whereas the right one is the original. Two top panels: slices in the (x, y) plane, two bottom panels: slices in the (x, z) plane. Red (resp. orange) dashed lines indicate the positions of the reslicing in the (x, z) plane (resp. (x, y)). The dictionary has been learned on a set of $\sim$ 10 planes coming from this time point and another time point further in time. Field dimensions: 51.2 x 51.2 µm. 100 planes were acquired with a $\sim$ 200 nm step.

To begin with, the effectiveness of compressed sensing microscopy by compressing the z dimension can only be assessed when all the variability of the process has been experienced. For instance, our simulations, even in a noisy setting, assumed a predefined noise model, that is only an approximation of the real-life noise. Also, the influence of structural error was not assessed.

In this section, we review the schemes implemented in the lattice light sheet and present some first results. Then, we present an implementation in a widespread, traditional epifluorescence microscope.

Also, at the time of writing, this implementation is still in progress, and these results are only preliminary. In the remaining two months of the internship, we will build on top of these first results.

3.1 Lattice light sheet implementation

From a software point of view, the lattice light sheet microscope is operated by a Labview®[5] code that allows to program the FPGA (a real-time, programmable device: a field-programmable gate array), ensures the synchronization between all the devices and ultimately permits acquisitions.

An actual implementation of a compressed sensing mode running at a reasonable speed requires real-time capabilities that can only be reached in embedded devices such as the FPGA. Indeed, the compressed sensing mode was implemented as an addon sending specific AOTF waveforms to the FPGA. Due to the high complexity of the existing Labview® code, this implementation was realized by the company that developed the software[6]. Figure 21 presents an overview of the addon,

[5] `http://www.ni.com/labview/`

[6] Colemantech, `https://www.colemantech.com`

as visible by the end user.

3.1.1 Acquisition of reference images

One first step of high importance is to make sure that one can acquire accurate reference images in the direct domain, to compare with the output of the reconstruction algorithm. Two options are available:

- Use the *standard 3D stack* mode of the control software,
- Perform an acquisition in the *compressed sensing* mode and load an identity matrix as the measurement matrix, thus mimicking the acquisition in the traditional mode.

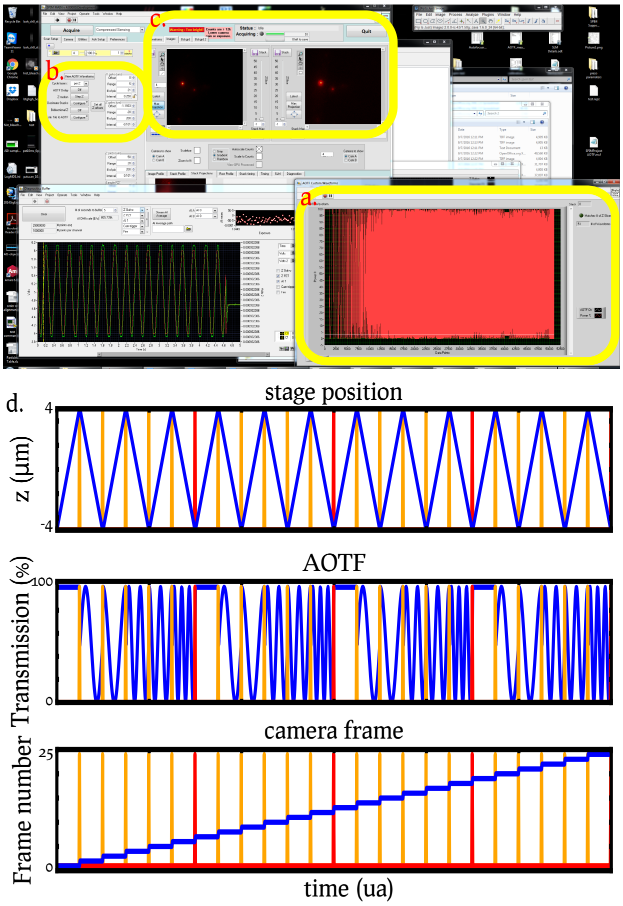

Figure 21: Graphical User Interface (GUI) part of the LabView® software used to control the compressed sensing mode. **(a).** The user can load a custom measurement matrix from a CSV file. **(b).** 3D acquisitions parameters can then be specified. These corresponds to formalism used in the traditional 3D stack formalism. Translation into the waveform modulation is performed internally. **(c).** the acquired compressed frames can then be vizualized, **(d).** Command sequence sent to the FPGA.

We opted for the second options for several reasons. The first one is that the gain (and then the output power) of the AOTF is not the same in the traditional *3D stack* mode and in the *compressed sensing* mode. In the former mode, the AOTF is controlled through the digital output of the AOTF and through a DC-DC converter whereas in the latter, one has to use the analog mode, since the intensity of the laser is modulated over time. This mode does not make use of the DC-DC converter.

Finally, the use of the identity matrix as a measurement matrix allowed us to acquire reference samples that can be compared one-to-one with the reconstructions.

3.1.2 Post-calibration of the measurement matrix

A second calibration step is to ensure that the intensity at the sample matches the measurement matrix. Discrepancies in this process result in structural error and can greatly impact the reconstruction quality.

Analysis of the reference documentations of the hardware and software elements of the microscope helped us to pinpoint potentially problematic devices. In the end, it seems that only the z-piezo that moves the objective in order to keep it confocal with the light sheet could be a source of structural errors.

Indeed, to operate the compressed sensing mode as described above requires to continuously scan the sample, and thus oscillate at a ~ 50 Hz frequency for a 10 ms acquisition time per measurement (one measurement can be made when the piezo stage goes up, the other when the stage is going down). This frequency is close to the maximal operating frequency of the piezo. Furthermore, the objective is a little bit too heavy for the piezo, which is likely to cause degraded performances.

Fortunately, the z-piezo provides an analog feedback that allows to measure the error/delay to reproduce the triangle waveform. The z-piezo voltage command and its voltage response are shown in Figure 22a for various acquisition times. One can see that for short acquisition times, there is a clear delay betweent the command and the response. As a result, the intensity at sample over the z dimension highly differs from the measurement matrix, which in theory should prevent any compressed sensing reconstruction.

Furthermore, the feedback from the z-piezo allows us to derive a corrected measurement matrix that matches the intensity at sample over time. Indeed, one can derive a model to correct the sensing matrix:

Consider what happens during one exposure frame. For the sake of simplicity, and without loss of generality, assume that the camera continually exposes between $t_0 = 0$ and $t_1 = 1$. Also, assume that we are given m measurement functions: $(m_i(t))_{1\dots m} : [0,1] \to [0,1]$ representing the intensity applied at sample for measurement i.

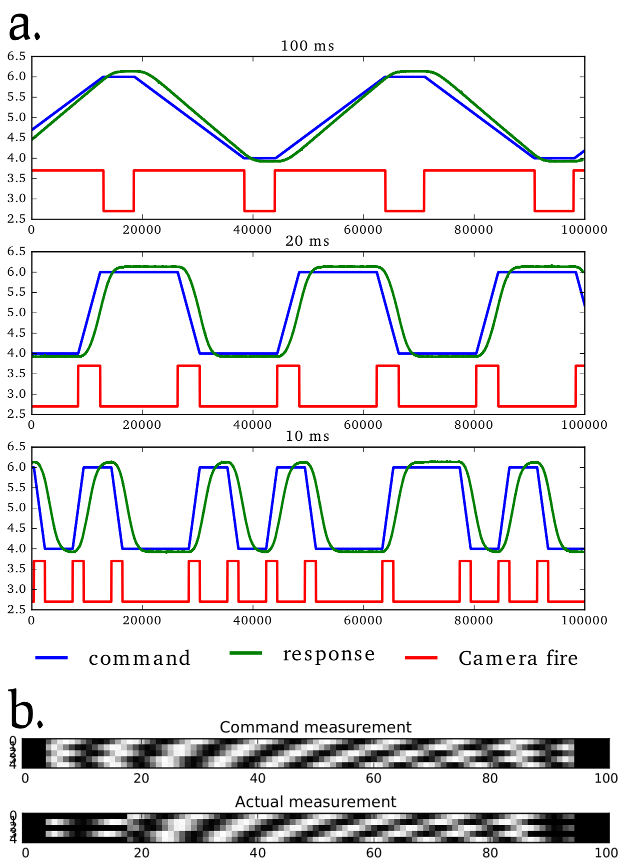

Figure 22: Correction of the z-piezo latency. **(a).** The z-piezo exhibit significan delay and distortion for short acquisition times: (top) 100 ms/measurement, (middle) 20 ms/measurement (bottom) 10 ms/measurement, the reason why the camera fires at irregular times yet has tobe investigated **(b).** *(top)* sample Fourier measurement matrix and *(bottom)* its calibrated counterpart for a 500ms/measurement acquisition. Note that the shift of each measurement vector reflect the fact that the raw z-piezo feedback is not exactly the same as the command, and such scaling error is not accounted for in the present method.

Note that these functions can easily be discretized to give back the traditional *measurement matrix* $A = (a_{ij})_{i\in 1\dots m, j\in 1\dots n} \in \mathbb{R}^{m\times n}$ used in the compressed sensing setting:

$$a_{ij} = \int_{\frac{j}{n}}^{\frac{j+1}{n}} m_i(t)dt$$

Here, the corrected measurement matrix $\tilde{A}$ can be inferred as a correction of the initial measurement matrix A and the record of the z-position of the piezo (denoted as $z(t), t \in [0,1]$). Thus, it can be easily seen that the corrected measurement matrix can be derived as follows:

$$\tilde{a}_{i,j} = \int_{t=t_i}^{t_{i+1}} \mathbf{1}_{z(t)\in[\frac{j}{n},\frac{j+1}{n}]} m_i(t)dt$$

Note that the presence of the indicator function $\mathbf{1}$ denotes the fact that we assume absolute sectioning. There is no PSF model in here.

This method can the be implemented to post-calibrate measurement matrices (as shown in Figure 22b) and to derive corrected measurement matrices (see Figure 22c).

So far, due to offset problem between the command and the feedback (visible in Figure 22a), this method was not used for the imaging presented in the next section.

3.2 Imaging results

3.2.1 Beads

To make a first demonstration of the method, we image fluorescent beads on a coverslip.

First images Beads on a coverslip were imaged using both the compressed sensing mode and the direct mode for comparison. Due to the loss of synchronization of the z-piezo at high imaging speed, imaging was performed at long exposure time in order to allow the piezo to settle (500 ms), and thus the results should only be seen as a proof of concept rather than a final result. Furthermore, this acquistion was realized using a square Fourier measurement matrix over 21 planes (no compression was applied), using a dictionary derived from a different acquisition. Sample reconstructed images are shown in Figure 23.

On the positive side, the reconstruction algorithm converges and produce a reconstructed 3D image. This image is relatively accurate in the (x,y) plane, with both large scale (halos from off-focus beads) and fine scale details (in-focus beads) reconstructed in most planes.

On the downsides, when sections are visualized in the (y,z) plane, there is almost no resolution in z, and it is almost impossible to reconstruct an image in this plane.

Influence of noise A major difference between beads and real-life samples is the brightness of the observed objects: fluorescent beads are much brighter, and thus the acquisitions are less noisy. To assess the influence of noise during the acquisitions, we varied the laser power by a factor of approximately 20, in order to obtain noisier acquisitions, closer from the SNR of real life samples. Slices of resulting 3D stacks are presented in Figure 24.

Obviously, the reconstructions degrade in the noisy setting. However, the degradation in the signal quality seems to provide similar artifacts in the direct and in the compressed domain. Also, the quality seems do not drop dramatically as the noise increases.

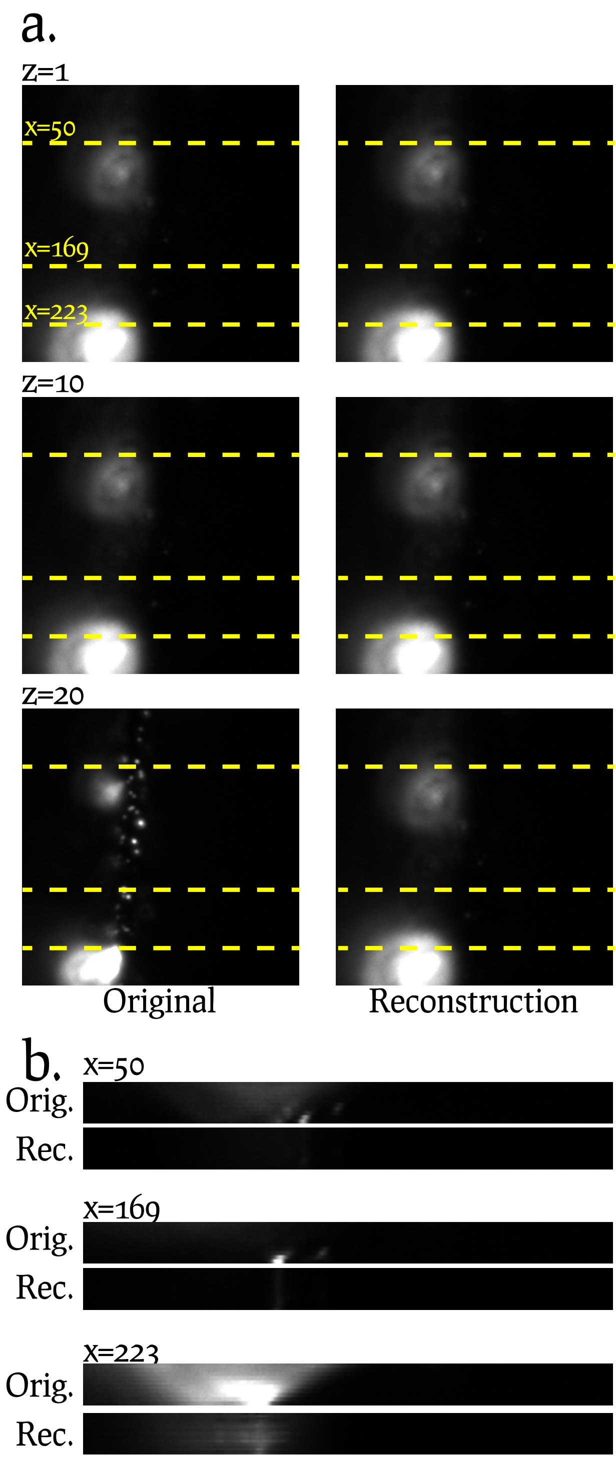

Figure 23: Comparison between a compressed sensing imaging reconstruction and acquisitions in the direct domain. **(a).** Comparisons in the (x, y) plane, *(left)* original plane, *(right)* compressed sensing reconstruction, yellow dashed lines mark the position of the (x, y) slices depicted in (b), **(b).** comparison in the (y, z) plane. Reconstruction from 20 planes acquired at a 500 ms framerate. Physical dimensions: 51.2 x 51.2x 3 µm.

This experiment can be seen as a first argument to prove that acquisitions and compressed sensing reconstructions are achievable in the presence of real-life imaging noise.

3.2.2 Live samples

Although compressed imaging of live samples is the major goal of this project, its realization is still in progress.

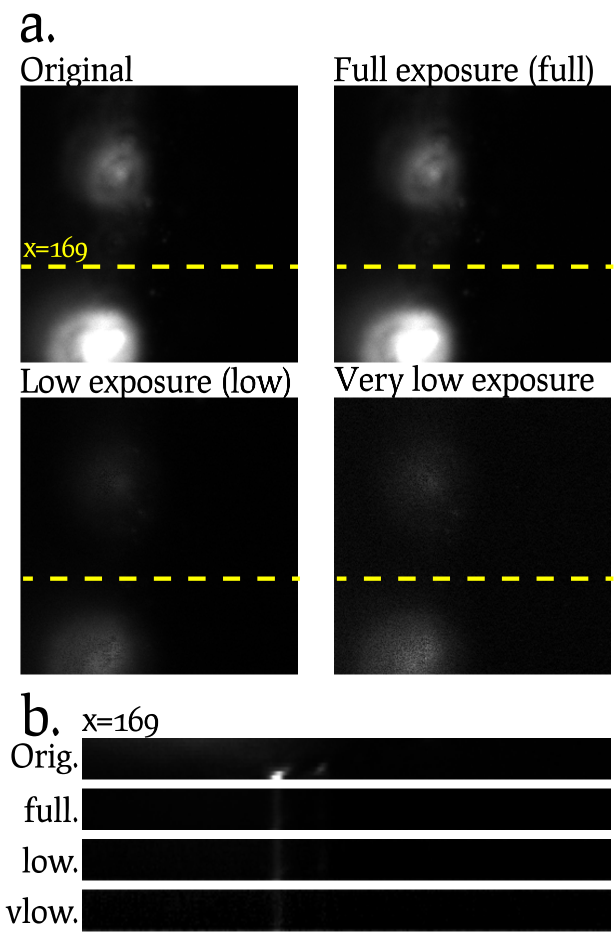

Figure 24: Influence of noise when reconstructed 3D images of beads. Images acquired with various laser powers, thus inducing various SNR. **(a).** slices in the (x, y) plane, **(b).** slices in the (y, z) plane. Reconstruction from 20 planes acquired at a 500 ms framerate. Physical dimensions: 51.2 x 51.2x 3 µm.

3.3 Towards a widely applicable setup

Due to the high sectioning ability of the lattice light sheet microscope, acquisition schemes performing a compression in the z dimension is of particular interest, such a compressed sensing scheme can be transposed to any microscope used to realize 3D acquisitions.

Indeed, the acquisition scheme presented above can be applied to any microscope as as soon as one can synchronize the z position with the illumination power. This is achievable either with lasers and an AOTF or a LED system (to get fast control of the LED, see Bosse et al., 2015).

Here, we decided to implement a free/open source and open hardware when possible compressed sensing module for a traditional epifluorescence microscope.

3.3.1 Main components

We rely on the following components to implement the compressed sensing imaging scheme.

`µ-Manager` (micro-Manager[7]) is a free/open source software to command microscopes. It is written in Java and builds on top of ImageJ. A major asset of `µ-Manager` is its correctly documented API that easily allow for the development of plugins.

`µ-Manager` plugins can pilot the hardware through generic drivers and adapters. Then, very generic code can be written to control a high diversity of hardware configurations.

The aim of the project is to implement a compressed sensing mode as a `µ-Manager` plugin.

Arduino

3.3.2 Software architecture

At the time of writing, the development of the plugin is still ongoing. The code is divided into several parts:

Real-time code The code to be loaded on the Arduino. It will ensure the command of the AOTF and the z-piezo but also the communication with the computer.

Arduino adapter for `µ-Manager` To implement compressed sensing acquisitions, one need to load the sensing matrix on the Arduino. This requires to modify the Arduino driver/adapter to send and receive additional commands to the Arduino.

`µ-Manager` plugin Communicates with the Arduino adapter and provides an easy-to-use (graphical) control over the compressed sensing mode, allowing to load sensing matrices, perform acquisitions, etc. A first draft can be seen on Figure 25.

[7] `https://micro-manager.org`, Edelstein et al., 2014

3.3.3 Results

As of the time of writing, some parts are still missing on the microscope, and thus no experiment has been performed yet.

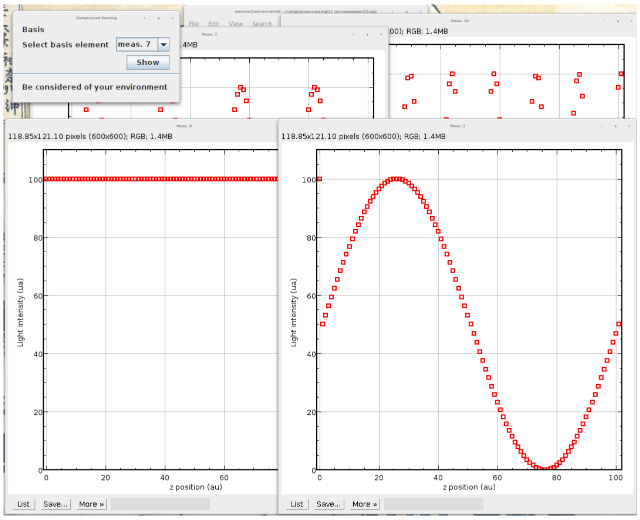

Figure 25: First version of the control plugin for micromanager: a measurement matrix has been loaded and can be displayed.

4 Discussion

Compressed sensing has been the subject of hectic developments in the past ten years and unraveled seminal ideas that broadly affected the signal processing community. Indeed, compressed sensing setups have revolutionized the field of MRI (Lustig et al., 2008), are regularly used in astronomy and have seeded key concepts in the machine learning community (Mairal, 2014).

Surprisingly, to our knowledge, compressed sensing-based setups have never made it to a production stage in the live microscopy imaging field, although many proof of concepts were developed (Marim, Angelini, and Olivo-Marin, 2009, Ye et al., 2009, Wu et al., 2010, Studer et al., 2012, Schwartz, Wong, and Clausi, 2012, Zhu et al., 2012).

Many reasons can be inovked:

1. **Interdisciplinarity**: designing a performant microscope is traditionally challenging, as it requires at the same time a significant theoretical background and highly technical practical skills, and it takes significant efforts for a mathematician to understand such a setup.
2. **Interdisciplinarity (again)**: the traditional reasoning scheme in the compressed sensing field is

nearly orthogonal to the ones in the microscopy field. Indeed, the effects of common imaging artifact (such as Poisson noise) is highly counter-intuitive.

3. **Ease of use**: When imaging in a compressed sensing mode, the feedback to the user is the compressed image, which is usually useless before decompression. With only the compressed image in hand during the acquisition, the biologist has no clue of how the sample is evolving. Even utterly simple parameters such as assessing the focus cannot be performed without a reconstruction.

4. **Ease of use (again)**: most of the setups require day-long reconstruction times, which makes the time between the acquisition and the analysis even longer. Note that in the MRI field the use of GPU allows some 3D reconstructions to be performed nearly real-time (Smith et al., 2012, Nett, Tang, and G.-H. Chen, 2010).

5. **Low SNR**: to gain either time or signal, the user can always play on the exposure time of the camera. which might have prevented extended studies on how to speed-up the acquisitions.

In the above sections, we tried to demonstrate the applicability of a compressed sensing approach to microscopy. To do so, we showed on simulations that compressed sensing algorithms can efficiently recover 3D images, with a compression factor up to ten times in the z dimension.

An interesting feature of compression in z with a Fourier basis is that each first measurement of a 3D stack is just the mean intensity over the stack (as the first Fourier coefficient is the mean value of the signal), allowing the user to get a 2D live feedback. Furthermore, fast, low quality reconstructions can be applied to reconstruct a preview signal using inverse fast Fourier transform.

However, the proof of concept on an actual hardware setup is still ongoing, although we obtained encouraging preliminary results.

4.1 Can this setup allow increased imaging speed?

One of the main promises of compressed sensing imaging is to reduce the time to acquire one 3D stack, thus allowing faster acquisitions at a given SNR.

Indeed, if in the traditional acquisition mode a plane is acquired in 10 ms and 200 planes make a 3D volume, the capture of each 3D volume takes 2000 ms. If by the means of compressed sensing one can achieve a 10 times compression factor, thus leading to 20 measurements of 10 ms, the acquisition of a 3D volume is now as low as 200 ms, indeed allowing imaging ten times faster.

However, there might be several serious issues with this:

1. **Hardware limitation**: in the actual setup, part of the hardware is pushed to its physical limit (the z-piezo scans at maximum velocity), and reconstructions require careful software correction (the post-calibration of the sensing matrix). Due to this structural error, it is hard to master the actual sensing matrix properties. Several solutions could nonetheless be envisioned: first, the hardware could be upgraded to include a more powerful z-piezo. Otherwise, one can imagine design sensing matrices accounting for the z-piezo distortion, ensuring that the theoretical measurement matrix matches the empirical one (and not the opposite as in the current setup).

2. **Reconstruction times**: if an acquisition at maximum speed were possible and a movie of 10 minutes realized (to cover several mitosis in a fly embryo for instance), this would result in 5x60x10 = 3000 3D volumes. Reconstruction of such a movie would take several weeks on a HPC cluster. This can be regarded as a poor compromise to get a ten times speed-up in the acquisition time.

3. **Structural stability**: biologists usually perform fast acquisitions to image fast moving structures. Conversely, the compressed sensing setting assumes that the sample reamins idle during the acquisition of a 3D stack, that is the first and the last projection of the sensing matrix indeed image the same thing. Whereas fast motion in the traditional acquisition mode results in motion blur and possibly duplicated objects across planes, the effect of such movements when imaging is compressed have not been assessed in this work, but are likely to at least highly disturb the reconstructed image, or, worst, prevent the algorithms from converging. Similar problems have

been addressed in the MRI field (Bilen, Wang, and Selesnick, 2012).

All these considerations point to the fact that fast imaging with compressed sensing remains challenging. It should be noted that none of these hurdles are blocking. They should be assessed carefully and the best trade-off should be selected.

4.2 Can compressed sensing decrease phototoxicity?

Another promising application of a compressed sensing microscopy setup is to reduce phototoxicity. Indeed, light imaging causes several damages to the sample. These goes from photobleahing of the fluorescent probe due to the formation of highly oxidizing, lihght-induced free radicals to various metabolic and developmental defects.

To this respect, many experiments are conducted under a time lapse imaging mode: one 3D image is taken at fixed time interal and the sample reamins in the dark between them.

A compressed sensing scheme coupled to the lattice light sheet microscope could bring a significant improvement with respect to phototoxicity, indeed:

- The lattice light sheet already has a very low phototoxicity because light is closely focused on the objective focal plane, limiting off-focus illumination
- In the compressed sensing mode, a 10 times compression is equivalent to ten times less exposure.
- Furthermore, in the compressed sensing mode, since the z-piezo is continuously swapping through the sample, each region is illuminated for a very short amount of time (of the order of less than 200 µs) for every measurement. Thus, the sample is given time to recover between two exposures. Such a feature has been shown to reduce phototoxicity.

Although we didn't have time to assess phototoxicity *in vivo*, it is likely that this parameter will be improved thanks to compressed sensing.

4.3 Reconstruction artifacts can be minimized

Across the experiments, several major artifacts appeared repeatedly: stripes in the reconstruction and low z resolution (see Figure 26). A major question is whether these obstacles can be overcomed or do they represent hard limitations of the method.

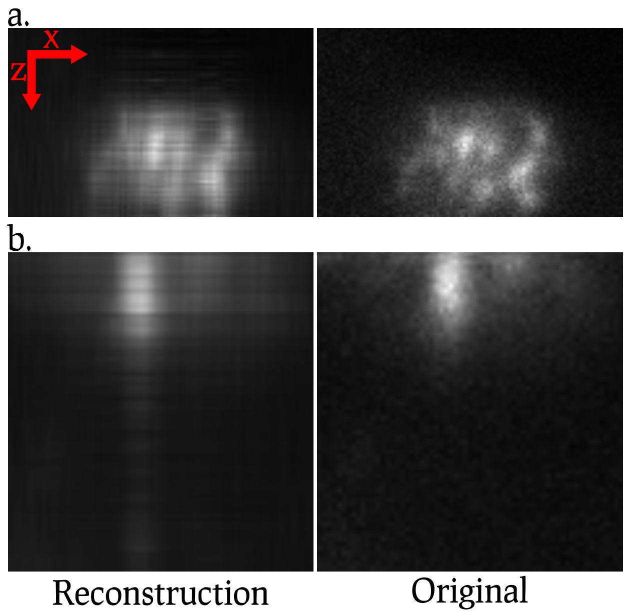

Figure 26: Main reconstruction artifacts encountered in compressed sensing reconstructions with a dictionary: *stripes* in the z dimension and a *degraded z resolution*, as can be seen in both panels (a) and (b). The two panels are zoom over a nucleus (~ 15 µm).

Stripes Some reconstructions exhibit superimposed stripes in the z dimension (visible in both panels of Figure 26. Such artifact turns out to be rather common and is related to the size of the dictionary. In our simulations, the use of a dictionary with a wider training set exhibited less stripes than the ones with smaller training sets. Thus, it seems likely that this artifact can be overcomed.

Low z resolution In many of the reconstructions, the z resolution appears limited (the reconstructed objects have a poor localization in z). One possible hypothesis is that not enough samples have been collected during the acquisition. A lower compression rate should then solve the problem.

Another hypothesis lies in the nature of the measured vectors. Indeed, it is not the same to perform a 10 times compression on a 300 element long vector than in a 20 element long vector. In the former case reconstruction is made from a 30 element compressed vector whereas in the latter the vector only has length two, which obviously encodes a very very limited amount of information. Indeed, a general rule in machine learning is that the higher the number of dimensions the sparser the signals tend to be.

The images depicted in Figure 26 show limited resolution in z whereas our first simulations did not had those artifacts. The major (if not only) difference between those two simulations is the size of the input vector, which was > 200 for the former and 100 for the latter. We hypothesize that getting longer z vectors would solve this problem.

This hypothesis is further strengthened by the acquisition performed on beads with only 20 positions in z: it shows almost no z sectioning.

4.4 Compressed sensing as a generic upgrade of an epifluorescence microscope

One of the main shortcomings of previously proposed compressed sensing settings for microscopy (recall Marim, Angelini, and Olivo-Marin, 2009, Ye et al., 2009, Wu et al., 2010, Studer et al., 2012, Schwartz, Wong, and Clausi, 2012, Zhu et al., 2012) is that they require strong modifications of an existing microscope.

Indeed, part of the setups implement a compressed confocal scheme that require to add a DMD (digital micromirror device) into the light path. Another part of the setups require to move the camera from a plane conjugated with the pupil plane to a plane conjugated with the back pupil plane to directly image in the Fourier space. Again, this requires a massive modification of a traditional setup.

On the other hands, the setup we propose can be fit to any epifluorescence microscope whose stage is controlled by a z-piezo with no optical modification of the setup, and, and likely little hardware modification. In case a hardware modification is required, we propose an Arduino-based replacement, which is both cost effective and highly versatile (due to the high-level API provided by the device and the numerous free/open-source code published to work with microscope parts).

5 Conclusion

During the course of this six month internship we partially developed a widely applicable compressed sensing scheme that can be fitted with little or no hardware modification to a wide range of microscopes, from a high-performance lattice light sheet microscope to a standard epifluorescence microscope.

To do so, we first demonstrated on simulations the feasability of compressed sensing techniques under noisy conditions and designed a scheme where the signal of a 3D acquisition is compressed in z during the acquisition. We demonstrated the applicability of this scheme both on extensive simulations and on preliminary implementation on a lattice light sheet microscope.

Together with the hardware implementation, we developed a free (GPLv3) analysis and reconstruction pipeline that can run both on an analysis computer or on a computing cluster (HPC).

An implementation of the scheme on a generic epifluorescence microscope, together with better quality acquisitions are still in progress.

> *A pile of rubble randomly heaped: the most beautiful order in the world.*
>
> — Heraclites

Acknowledgements

I want to thank Mustafa its mentoring and its patience when showing me the lattice light sheet, Astou for teaching me how to align lasers of a microscope, the whole IMOD lab in Institute Pasteur for hosting me for the first three months of my internship and the Tjian-Darzacq lab for the rest of it. Obviously, thank you to Christophe and Xavier for giving me the opportunity to work in such exciting environments. ♡ ♡ ♡

Finally, I have to thank Wei Ouyang (IMOD lab), Moran Mordechay (Schechner group, Technion, Israel) and Zac Harmany (Levenson lab, UC Davis, USA) for sharing code with me and answering my naive questions about compressed sensing.

References

[1] Ery Arias-Castro and Yonina C. Eldar. “Noise Folding in Compressed Sensing”. In: *IEEE Signal Processing Letters* 18.8 (Aug. 2011), pp. 478–481. ISSN: 1070-9908, 1558-2361. DOI: 10.1109/LSP.2011.2159837.

[2] Richard Baraniuk. “Compressive Sensing”. In: *IEEE Signal Processing Magazine* (2007). 00000.

[3] Stephen Becker, Jérôme Bobin, and Emmanuel Candès. “NESTA: A Fast and Accurate First-Order Method for Sparse Recovery”. In: (2009). 00000.

[4] Cagdas Bilen, Yao Wang, and Ivan Selesnick. “Compressed Sensing for Moving Imagery in Medical Imaging”. In: *arXiv preprint arXiv:1203.5772* (2012). 00003.

[5] Jens B. Bosse et al. "Open LED Illuminator: A Simple and Inexpensive LED Illuminator for Fast Multicolor Particle Tracking in Neurons". In: *PLOS ONE* 10.11 (Nov. 23, 2015). Ed. by Kurt I. Anderson. 00000, e0143547. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0143547.

[6] E.J. Candes, J. Romberg, and T. Tao. "Robust Uncertainty Principles: Exact Signal Reconstruction from Highly Incomplete Frequency Information". In: *IEEE Transactions on Information Theory* 52.2 (Feb. 2006). 10205, pp. 489–509. ISSN: 0018-9448. DOI: 10.1109/TIT.2005.862083.

[7] E.J. Candes and M.B. Wakin. "An Introduction To Compressive Sampling". In: *IEEE Signal Processing Magazine* 25.2 (Mar. 2008). 05537, pp. 21–30. ISSN: 1053-5888. DOI: 10.1109/MSP.2007.914731.

[8] Emmanuel Candes, Yonina C. Eldar, et al. "Compressed Sensing with Coherent and Redundant Dictionaries". In: (2010). 00010.

[9] Emmanuel Candes and Justin Romberg. "l1-Magic: Recovery of Sparse Signals via Convex Programming". In: *URL: www. acm. caltech. edu/l1magic/downloads/l1magic. pdf* 4 (2005). 00608, p. 14.

[10] Emmanuel Candès and Terence Tao. "Near Optimal Signal Recovery From Random Projections: Universal Encoding Strategies?" In: (2004). 00000.

[11] B.-C. Chen et al. "Lattice Light-Sheet Microscopy: Imaging Molecules to Embryos at High Spatiotemporal Resolution". In: *Science* 346.6208 (Oct. 24, 2014). 00184, pp. 1257998–1257998. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.1257998.

[12] Xuefeng Chen et al. "Compressed Sensing Based on Dictionary Learning for Extracting Impulse Components". In: *Signal Processing* 96 (Mar. 2014). 00034, pp. 94–109. ISSN: 01651684. DOI: 10.1016/j.sigpro.2013.04.018.

[13] Mark A. Davenport, Deanna Needell, and Michael B. Wakin. "Signal Space CoSaMP for Sparse Recovery With Redundant Dictionaries". In: *IEEE Transactions on Information Theory* 59.10 (Oct. 2013). 00047, pp. 6820–6829. ISSN: 0018-9448, 1557-9654. DOI: 10.1109/TIT.2013.2273491.

[14] Mark A. Davenport and Michael B. Wakin. "Analysis of Orthogonal Matching Pursuit Using the Restricted Isometry Property". In: *IEEE Transactions on Information Theory* 56.9 (Sept. 2010). 00319, pp. 4395–4401. ISSN: 0018-9448, 1557-9654. DOI: 10.1109/TIT.2010.2054653.

[15] David L. Donoho and Michael Elad. "Optimally Sparse Representation in General (Nonorthogonal) Dictionaries via l1 Minimization". In: *Proceedings of the National Academy of Sciences* 100.5 (Apr. 3, 2003). 02103, pp. 2197–2202. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.0437847100.

[16] David L. Donoho and Xiaoming Huo. "Uncertainty Principles and Ideal Atomic Decomposition". In: *IEEE Transactions on Information Theory* 47.7 (2001). 01597, pp. 2845–2862.

[17] Arthur D Edelstein et al. "Advanced Methods of Microscope Control Using Micro-Manager Software". In: *Journal of Biological Methods* 1.2 (Nov. 7, 2014). 00074, p. 10. ISSN: 2326-9901. DOI: 10.14440/jbm.2014.36.

[18] Michael Elad. "Optimized Projections for Compressed Sensing". In: (2006). 00000.

[19] Michael Elad. *Sparse and Redundant Representations*. New York, NY: Springer New York, 2010. ISBN: 978-1-4419-7010-7 978-1-4419-7011-4.

[20] Yonina C. Eldar and Gitta Kutyniok, eds. *Compressed Sensing: Theory and Applications*. 00000. Cambridge ; New York: Cambridge University Press, 2012. 544 pp. ISBN: 978-1-107-00558-7.

[21] Hernan G. Garcia et al. "Quantitative Imaging of Transcription in Living Drosophila Embryos Links Polymerase Activity to Patterning". In: *Current Biology* 23.21 (Nov. 2013). 00044, pp. 2140–2145. ISSN: 09609822. DOI: 10.1016/j.cub.2013.08.054.

[22] Jonathan B Grimm et al. "A General Method to Improve Fluorophores for Live-Cell and Single-Molecule Microscopy". In: *Nature Methods* 12.3 (Jan. 19, 2015). 00066, pp. 244–250. ISSN: 1548-7091, 1548-7105. DOI: 10.1038/nmeth.3256.

[23] Zachary T. Harmany, Roummel F. Marcia, and Rebecca M. Willett. "This Is SPIRAL-TAP: Sparse Poisson Intensity Reconstruction Algorithms—theory and Practice". In: *Image Processing, IEEE Transactions on* 21.3 (2012), pp. 1084–1096.

[24] Matthew A. Herman and Thomas Strohmer. "General Perturbations in Compressed Sensing". In: *SPARS'09-Signal Processing with Adaptive Sparse Structured Representations*. 00003. 2009.

[25] Jan Huisken et al. "Optical Sectioning Deep inside Live Embryos by Selective Plane Illumination Microscopy". In: *Science* 305.5686 (2004). 00963, pp. 1007–1009.

[26] Myung K. Kim. "Diffraction and Fourier Optics". In: *Digital Holographic Microscopy*. Vol. 162. 00000. New York, NY: Springer New York, 2011, pp. 11–28. ISBN: 978-1-4419-7792-2 978-1-4419-7793-9.

[27] Daniel D. Lee and H. Sebastian Seung. "Algorithms for Non-Negative Matrix Factorization". In: *Advances in Neural Information Processing Systems*. 04984. 2001, pp. 556–562.

[28] Jonathan Le Roux, Felix Weninger, and J. Hershey. "Sparse NMF–half-Baked or Well Done". In: *Mitsubishi Electric Research Labs (MERL), Cambridge, MA, USA, Tech. Rep., no. TR2015-023* (2015). 00007.

[29] Georgyi V. Los et al. "HaloTag: A Novel Protein Labeling Technology for Cell Imaging and Protein Analysis". In: *ACS Chemical Biology* 3.6 (June 2008). 00593, pp. 373–382. ISSN: 1554-8929, 1554-8937. DOI: 10.1021/cb800025k.

[30] M. Lustig et al. "Compressed Sensing MRI". In: *IEEE Signal Processing Magazine* 25.2 (Mar. 2008). 00834, pp. 72–82. ISSN: 1053-5888. DOI: 10.1109/MSP.2007.914728.

[31] Julien Mairal. "Sparse Modeling for Image and Vision Processing". In: *Foundations and Trends® in Computer Graphics and Vision* 8 (2-3 2014). 00039, pp. 85–283. ISSN: 1572-2740, 1572-2759. DOI: 10.1561/0600000058.

[32] Marcio M. Marim, Elsa D. Angelini, and J.-C. Olivo-Marin. "A Compressed Sensing Approach for Biological Microscopic Image Processing". In: *2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*. 00010. IEEE, 2009, pp. 1374–1377.

[33] D McGloin and K Dholakia. "Bessel Beams: Diffraction in a New Light". In: *Contemporary Physics* 46.1 (Jan. 2005). 00554, pp. 15–28. ISSN: 0010-7514, 1366-5812. DOI: 10.1080/0010751042000275259.

[34] Moran Mordechay and Yoav Y. Schechner. "Matrix Optimization for Poisson Compressed Sensing". In: *Signal and Information Processing (GlobalSIP), 2014 IEEE Global Conference on*. 00000. IEEE, 2014, pp. 684–688.

[35] D. Needell and J. A. Tropp. "CoSaMP: Iterative Signal Recovery from Incomplete and Inaccurate Samples". In: (Mar. 17, 2008). 02414. arXiv:0803.2392 [cs, math].

[36] Brian E. Nett, Jie Tang, and Guang-Hong Chen. "GPU Implementation of prior Image Constrained Compressed Sensing (PICCS)". In: ed. by Ehsan Samei and Norbert J. Pelc. 00005. Mar. 4, 2010, p. 762239. DOI: 10.1117/12.844578.

[37] Thomas A Planchon et al. “Rapid Three-Dimensional Isotropic Imaging of Living Cells Using Bessel Beam Plane Illumination”. In: *Nature Methods* 8.5 (May 2011). 00392, pp. 417–423. ISSN: 1548-7091, 1548-7105. DOI: 10.1038/nmeth.1586.

[38] Maxim Raginsky, Sina Jafarpour, et al. “Performance Bounds for Expander-Based Compressed Sensing in Poisson Noise”. In: *Signal Processing, IEEE Transactions on* 59.9 (2011), pp. 4139–4153.

[39] Maxim Raginsky, Rebecca M. Willett, et al. “Compressed Sensing Performance Bounds under Poisson Noise”. In: *Signal Processing, IEEE Transactions on* 58.8 (2010), pp. 3990–4002.

[40] J. C. Ramirez-Giraldo et al. “Nonconvex prior Image Constrained Compressed Sensing (NCPICCS): Theory and Simulations on Perfusion CT”. In: *Medical Physics* 38.4 (2011). 00050, p. 2157. ISSN: 00942405. DOI: 10.1118/1.3560878.

[41] Shimon Schwartz, Alexander Wong, and David A. Clausi. “Compressive Fluorescence Microscopy Using Saliency-Guided Sparse Reconstruction Ensemble Fusion”. In: *Optics Express* 20.16 (July 16, 2012). 00006, pp. 17281–17296.

[42] David S. Smith et al. “Real-Time Compressive Sensing MRI Reconstruction Using GPU Computing and Split Bregman Methods”. In: *International Journal of Biomedical Imaging* 2012 (2012). 00000, pp. 1–6. ISSN: 1687-4188, 1687-4196. DOI: 10.1155/2012/864827.

[43] V. Studer et al. “Compressive Fluorescence Microscopy for Biological and Hyperspectral Imaging”. In: *Proceedings of the National Academy of Sciences* 109.26 (June 26, 2012). 00110, E1679–E1687. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.1119511109.

[44] Terence Tao and Emmanuel Candès. “Decoding by Linear Programming”. In: *arXiv preprint arXiv:math/0502327* (2004). 00000.

[45] Andreas M. Tillmann and Marc E. Pfetsch. “The Computational Complexity of the Restricted Isometry Property, the Nullspace Property, and Related Concepts in Compressed Sensing”. In: *IEEE Transactions on Information Theory* 60.2 (2014). 00000, pp. 1248–1259.

[46] Michael B. Wakin et al. “Compressive Imaging for Video Representation and Coding”. In: *Image Processing, 2006 IEEE International Conference on*. 00286. IEEE, 2006, pp. 1273–1276.

[47] Tim Wescott. “Sampling: What Nyquist Didn’t Say, and What to Do About It”. In: *Wescott Design Services, Oregon City, OR* (2010).

[48] Rebecca M. Willett, Roummel F. Marcia, and Jonathan M. Nichols. “Compressed Sensing for Practical Optical Imaging Systems: A Tutorial”. In: *Optical Engineering* 50.7 (2011). ISSN: 0091-3286. DOI: 10.1117/1.3596602.

[49] Rebecca M. Willett and Maxim Raginsky. “Performance Bounds on Compressed Sensing with Poisson Noise”. In: *Information Theory, 2009. ISIT 2009. IEEE International Symposium on*. IEEE, 2009, pp. 174–178.

[50] Yuehao Wu et al. “Experimental Demonstration of an Optical-Sectioning Compressive Sensing Microscope (CSM)”. In: *Optics express* 18.24 (2010). 00020, pp. 24565–24578.

[51] P. Ye et al. “Compressive Confocal Microscopy: 3D Reconstruction Algorithms”. In: *SPIE MOEMS-MEMS: Micro-and Nanofabrication*. 00013. International Society for Optics and Photonics, 2009, 72100G–72100G.

[52] Andy B. Yoo, Morris A. Jette, and Mark Grondona. “Slurm: Simple Linux Utility for Resource Management”. In: *Workshop on Job Scheduling Strategies for Parallel Processing*. 00366. Springer, 2003, pp. 44–60.

[53] Zhou Wang and A.C. Bovik. “Mean Squared Error: Love It or Leave It? A New Look at Signal Fidelity Measures”. In: *IEEE Signal Processing Magazine* 26.1 (Jan. 2009). 01234, pp. 98–117. ISSN: 1053-5888. DOI: 10.1109/MSP.2008.930649.

[54] Lei Zhu et al. “Faster STORM Using Compressed Sensing”. In: *Nature Methods* 9.7 (Apr. 22, 2012), pp. 721–723. ISSN: 1548-7091, 1548-7105. DOI: 10.1038/nmeth.1978.

6 Supplementary materials

6.1 Lattice lightsheet microscope

6.1.1 Principle of light sheet microscope

In standard *epifluorescence* microscopy (Figure 1a), the same objective is used to illuminate the sample and excite the fluorophores and to collect the emitted light. Thus, while the objective focal plane is relatively narrow (in the order of micrometers), the whole sample is illuminated (it is easy to show that the integrated laser power over one z position is a constant: each plane receives the same total power). Since the objective is focused on a definite focal plane, this off-focus illumination should not be an issue. However, this excites out-of-focus fluorophores, increasing the background noise and reducing the SNR.

Light sheet microscopes (Figure 1b) benefit from an illumination confined to the depth of field of the objective. Indeed, a second objective (called excitation objective) is used to illuminate the sample. A clever optical path upstream of the excitation objective patterns the radial, Gaussian laser beam into a structured "sheet" that in theory can be as thin as allowed by the numerical aperture of the excitation objective. The detection objective is placed orthogonally and kept confocal to the illumination sheet. Table 1 summarizes the differences between the two techniques.

Table 1: Comparison between epifluorescence and light sheet microscopy

Criterion	Epifluorescence	Lattice Light sheet
off-focus illum.	yes	no
photodamage	high	low
Ease of use	easy	more challenging
Max. z resolution	$\sim 1\ \mu$ m	$\sim$ 300 nm

One can think that the illumination sheet can be made as thin as possible. Indeed, techniques achieved various sheet depth (Huisken et al., 2004, B.-C. Chen et al., 2014). However, the rules of optics put a hard limit on how focused a beam can be: the image of a point source through an optical system, called the *point spread function* (PSF) determines the resolution of the system. This resolution depends on the wavelength of the laser (λ) and of the numerical aperture (NA) of the excitation objective through and scales as: $\lambda/(2NA)$. Kim, 2011 details and examplifies the image formation process.

The first light sheet microscopes were based on cylindrical lenses to produce an elongated beam, or light sheet (Huisken et al., 2004). More recent techniques rely on Bessel beams and lattice-derived interference patterns.

6.1.2 Theory of lattice light sheet microscopy

The concept of the lattice light sheet can be summarized as follows:

1. An array of radially symmetric, non-diffracting beam of light is created: a lattice of *Bessel beams*
2. The spacing between the Bessel beams is optimized in order to ensure destructive interference betwee the side lobes of the Bessel beams, thus "confining" the beam.
3. This array of beams is then swept continually back and forth across the sample at high speed, thus creating a uniform "sheet".

Part of this introduction follows closely B.-C. Chen et al., 2014 (Supplementary Materials).

Bessel beams Light beams, as traveling waves, can be described with an electric field $\mathbf{e}(\mathbf{x}, t)$, where $\mathbf{x}$ stands for the 3D coordinates. Under some asumptions (homogeneous medium), the electrical field can be decomposed into into the superposition of N propagating plane waves:

$$\mathbf{e}(\mathbf{x}, t) = \sum_{k=1}^{N} \mathbf{e_n} \exp(i(\mathbf{k_n} \cdot \mathbf{x} - \omega t)) \tag{3}$$

where $\mathbf{k_n}$ are the *wavevectors* characterizing each wave plane, $\omega = 2\pi c/\lambda$ with λ the wavelength in the medium and c the speed of light.

Now assume that one manages to tweak the beam such as all the N wavevectors $\mathbf{k_n}$ lie on the surface of a cone of half-angle θ, that is: $\mathbf{k_n} \cdot \mathbf{e_y} = kcos\theta$. Eq. 3 can then be factorized (McGloin and Dholakia, 2005 and assuming that the wavevectors fully cover the cone):

$$\mathbf{e}(\mathbf{x}, t) = \frac{A}{2\pi} \mathbf{e} J_0(k\sqrt{x^2+z^2} \sin\theta) \exp\left[i(ky\cos\theta - \omega t)\right] \tag{4}$$

Where $J_0(\alpha) = \frac{1}{2\pi}\int_0^{2\pi} \exp(i\alpha\cos\phi)d\phi$. Such beams where the wavevectors lie on a cone are called *Bessel beams*. Notice that J_0 is a zeroth-order Bessel function of the first kind. One can then notice a very interesting property of such beams: they are *non-diffracting*. That is, for the intensity I of a beam propagating in the z direction: $e(x, y, z) = I(x, y)$.

Such beams are solutions of the Helmholtz equation and quasi-Bessel beams (of finite spatial extension) can be experimentally generated (McGloin and Dholakia, 2005). Figure 27 shows a representation of the pattern of a Bessel beam.

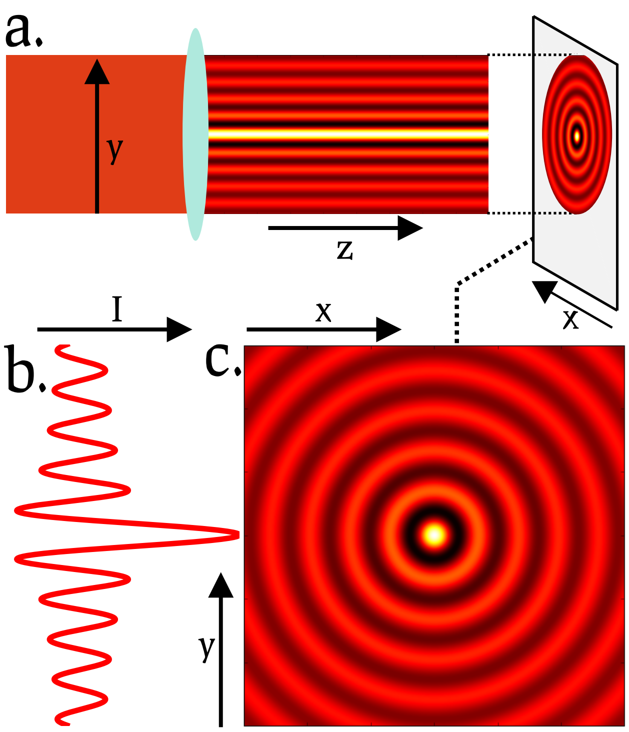

Figure 27: **(a).** Bessel beam in the yz plane: the intensity pattern is constant over z at $y = cst$, **(b).** intensity profile for $x = 0$, **(c).** bessel beam in the xy plane: bessel function of the first kind of order zero.

A Bessel beam is an ideal candidate to be swept across the sample and thus form a light sheet (see Planchon et al., 2011). Furthermore, theoretical and experimental studies have showed that the central peak of the Bessel function can be made diffraction-limited. However, a characteristics of the Bessel functions is that each ring carries the same energy, which would ultimately results in out-of-focus illumination.

Lattice of interfering Bessel beams One way to partly suppress outer rings of the Bessel function is to trigger destructive interferences of the outer rings of adjacent Bessel beams. Such interference can be achieved by designing an *array* of Bessel beams, where the characteristic lengths of the lattice are fine-tuned in order to exhibit negative interferences (Figure 28a and b).

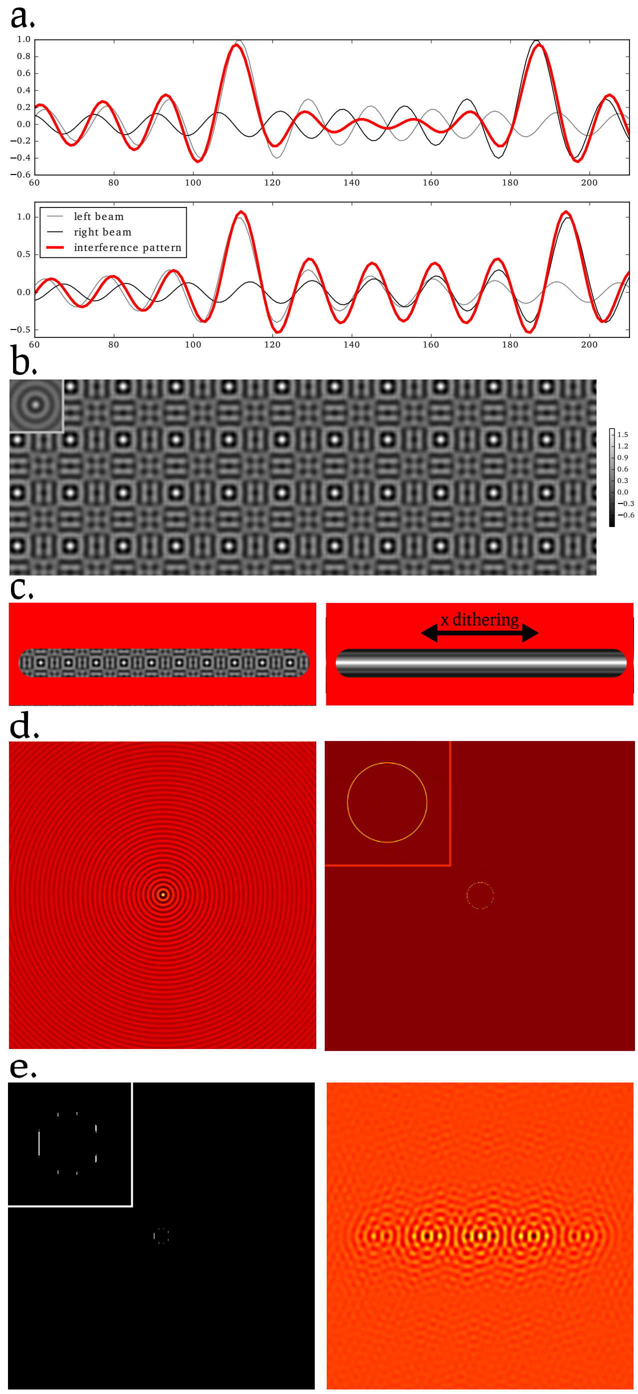

Figure 28: Lattice of destructively interfering Bessel beams. **(a).** schematic of destructive (top) and constructive (bottom) interference patterns between two Bessel beams. Note how the rings between the two centers get attenuated/amplified. **(b).** interference pattern on a 2D optical lattice created by the interference of Bessel beams. **inset:** one Bessel function before interference, **(c)**-left: SLM-mediated selection of a subpattern of the lattice, red: SLM pixels off, -right: view of the incident beam at sample (after dithering the beam in the x direction), **(d).** Bessel function (left) and its intensity in the Fourier domain (right), and a zoom on the annulus (inset), **(e)** intensity at the Fourier plane of an array of Bessel beams (left) and its intensity at the object plane (right).

Once the beam has been confined thanks to the interfering array of Bessel beams, a moving mirror continually sweeps the pattern in order to "blur" it in the x direction (Figure 28c-right) and ensure a uniform light intensity at the sample. This allows to expose the sample with a very thin, possibly diffraction-limited light sheet.

Generation of the beam A last remaining question is how to actually generate the array. Two non-exclusive ways to generate the array of Bessel beams are:

- To engineer it in the direct domain using a spatial light modulator (SLM, a liquid crystal device)
- To engineer it in the Fourier space (that corresponds to the back focal plane of the objective, also called "Fourier domain"). Indeed, a Bessel beam can be generated by putting an annulus mask in the back focal plane of the objective[8] : the inverse Fourier transform of an annulus is a Bessel function. The array can be generated by restricting the illumination of the ring to regularly spaced spots (Figure 28e). To get an ideal Bessel lattice, the spots should be infinitesimally small. Since this is physically impossible and impractical (because as the size of the spots decreases, so does the total incident light at sample), spots are kept as a finite size. This leads to the formation of a Gauss-Bessel beam, which exhibits similar focusing properties as the Bessel beam.

As detailed below, a combination of the two approaches can be implemented.

6.1.3 Optical implementation

In practice, a complex optical path is required to generate the light sheet (Figure 29 and 2). First, an elongated beam is created by a pair of cylindrical lenses. Then, a spatial light modulator applies a binary intensity mask conjugated with the object plane, in order to confine the light sheet in z and to produce the interference pattern. An annular mask is placed in the Fourier plane to filter the pattern and reject unwanted frequencies.

At that stage, the lattice light sheet has been created. It is then dithered along the x axis in order to produce a homogeneous illumination pattern. This is done by the x galvo (a digitally-controlled orientable micromirror). The dithered sheet is then projected onto the sample by the excitation objective, and the emitted light is collected with the observation objective at a 90° angle and projected onto the sensor of a electron-multiplying CCD camera (EMCCD), providing a very high sensitivity.

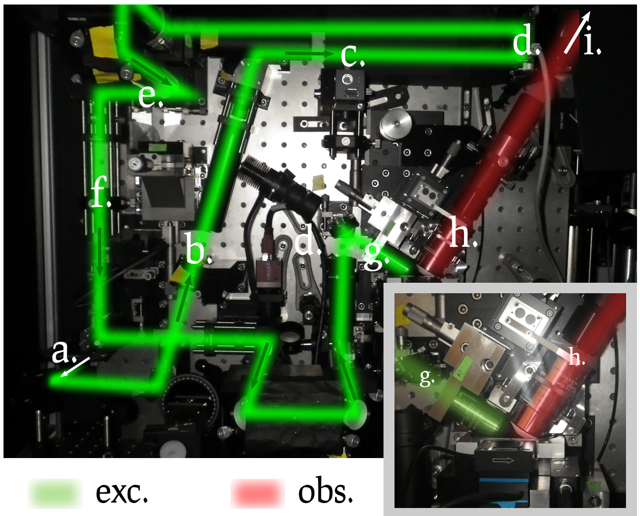

Figure 29: Picture of the lattice light sheet microscope. **(a).** from the laser lines and the AOTF, **(b).** x cylindrical lens, **(c).** z cylindrical lens, **(d).** SLM, **(e).** annulus mask, **(f).** z and x galvos., **(g).** excitation objective, **(h).** observation objective, **(i).** to EMCCD camera. **inset**: zoom on the objectives.

When focusing through the sample, the observation objective moves with respect to the sample in the z plane. To keep the sheet in focus, the z galvo adjusts the position of the sheet, so that the sheet and the focal plane of the observation objective remain confocal.

Finally, the laser intensity at the sample can be adjusted by an acousto-optical tunable filter (AOTF), a piezo-electric device that can be used to precisely and rapidly modulate the light intensity. A picture of the actual light path is shown Figure 2.

To achieve successful imaging, several pieces of hardware have to be precisely synchronized:

- The dithering in x has to be synchronized with the camera exposure in order to avoid aliasing effects and to ensure even illumination,
- The light sheet and the focal plane of the observation objectives have to remain confocal,
- Since the SLM has has a short periodic off time, this also has to be synchronized with the camera.

Synchronization is achieved by using a real-time, programmable device: a field-programmable gate array (FPGA). The controllable parts of the microscope are connected to the FPGA through either digital or analog channels. The FPGA runs a specific execution sequence to acquire an image (such as: open AOTF, start

[8]Derivation can be found on this page: `http://math.stackexchange.com/questions/78316/fourier-transform-of-bessel-functions`

the SLM, trigger the camera, etc). A custom-built GUI helps to design control sequences that are compiled and sent to the FPGA. Thanks to the programmable nature of the FPGA, custom imaging sequences can be implemented, allowing some tweaking from the user.

In order to implement the specific, user-defined imaging schemes required to do compressed sensing, some parameters of the microscope will have to be tuned. It is thus useful to review a set of characteristics and limitations of the microscope, in order to design a scheme that fits those requirements. A few characteristics are summarized in Table 2.

Table 2: Characteristics and limitation of the lattice light sheet microscope. **(*)** depending on the size of the region of interest imaged, **(**)** valid for an unloaded piezo. When the piezo carries the detection objective, this frequency is much lower (see Section 3.1.2 for a more detailed discussion). Values come from reference sheets, user manuals or effective measured characteristics.

Parameter	Value
Thickness of the sheet	down to $\sim$ 400 nm
Laser power (input/at sample)	1 W / 5 mW
Camera framerate (512x512 px)	100 Hz*
Camera framerate (128x128 px)	1000 Hz*
x or z dithering	15 kHz
AOTF off $\rightarrow$ on time	$\approx$ 10 μ s
Piezo stage speed	up to 250 Hz**

6.2 Theory of compressed sensing

Rather than presenting an exhaustive view of the compressed sensing field, this section focuses on key concepts and theorem that are relevant for the implementation of compressed sensing schemes in microscopes.

6.2.1 Statement of the problem

The compressed sensing problem can be stated as an optimization problem such that standard optimization techniques can be applied.

We have introduced in Section 1.2 the connections between compressible signals and sparsity. Indeed, assuming a sparse input signal x of length n, we aim at showing that it can be recovered from $m \ll n$ non-adaptive measurements y performed with a measurement matrix $A \in \mathbb{R}^{m \times n}$, which can be formulated as:

$$\hat{x} = x \text{ sparse, s.t. } y = Ax$$

Now, when solving a linear problem such as $y = Ax$, one can see (Figure 30) that some optimization criteria such as $\ell_p, p \leq 1$ norms tend to produce sparse solutions. Indeed, assume that one solves the optimization problem $\min_x ||x||_{\ell_p}$ s.t. $y = Ax$ and that the ℓ_p cost is given (the problem can be seen as finding the intersection between the linear constraint and the ℓ_p norm polyhedron. For $p > 1$, the polyhedron is convex, and the intersection (the solution to the problem) is likely to give a point with many non-zero coordinates. On the other hand, when $p \leq 1$, the polyhedron is not convex, and the intersection with the linear constraint is likely to fall on the axis, ie. the other coordinates are zero: the solution is sparse.

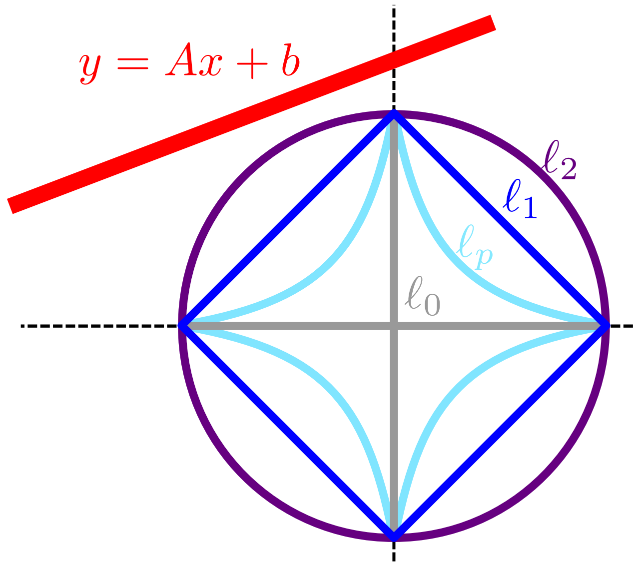

Figure 30: ℓ_p norms induce sparse solutions when $p \leq 1$. Solving the problem $y = Ax$ under various ℓ_p norms. Here the polyhedra represent fixed cost functions. When the solution lies on the axis (ℓ_p, p $\leq$ 1), it has some zero coordinates (an increasing amount with the dimensionality of the problem) and thus is sparse.

An interesting limit case arises when p tends to zero. One then finds the so-called ℓ_0 pseudo-norm, that is simply the number of non-zero coordinates. Thus, the compressed sensing problem is canonically stated as (E.J. Candes and M. Wakin, 2008, Baraniuk, 2007):

$$\min_x ||x||_{\ell_0} \text{ s.t. } y - Ax = 0 \tag{5}$$

That said, such formulation of the problem leads to several questions. One of the firsts is related to the fact that the ℓ_0 pseudo-norm is non-convex, and thus one has no guarantee about the uniqueness of the solution of the optimization problem and the existence of local minima.

6.2.2 Uniqueness of the solution

Uniqueness via the *spark* Although non-convex, one can show that this problem has a unique solution.

Following Elad, 2010, let us first define the *spark* of a matrix:

Definition 1 (Spark, from Donoho and Elad, 2003)**.** *The spark of a given matrix A, spark(A), is the smallest number of columns from A that are linearly dependent, that is:*

$$spark(A) = \min_{d \neq 0} ||d||_0 \ s.t. \ Ad = 0$$

According to Donoho and Elad, 2003: "note that, although spark and rank are in some ways similar, they are totally different". The spark characterizes the null-space of a matrix with ℓ_0 norm. For instance, if A is full rank, then $spark(A) = 2$.

In contrast with the rank, computing the spark of a matrix is hard, because it requires to enumerate over all the combinations of columns. Indeed, it turns out that the computation of the spark is a `NP-hard` problem (Tillmann and Pfetsch, 2014).

Furthermore, vectors x in the null-space of A have $||x||_{\ell_0} \geq spark(A)$. Indeed, since the null-space is characterized by $\{x, Ax = 0\}$, one need to select and combine linearly at least $spark(A)$ vectors from A to create the zero vector.

Theorem 2 (Uniqueness – spark, from Elad, 2010)**.** *If a system of linear equations $Ax = y$ has a solution x such that $||x||_{\ell_0} < spark(A)/2$, this solution is necessarily the sparsest.*

Proof. Assume x and x' two solutions to the undertermined linear problem $Ax = y$ and that x is such that $||x||_{\ell_0} < spark(A)/2$.

Since x and x' are two solutions, $x - x'$ is in the null-space of A: $A(x - x') = 0$. Then, by the definition of the spark, it follows that:

$$||x - x'||_{\ell_0} \geq spark(A)$$

Then, by the triangular inequality:

$$||x||_{\ell_0} + ||x'||_{\ell_0} \geq spark(A)$$

Then, since $||x||_{\ell_0} < spark(A)/2$, it follows that $||x'||_{\ell_0} \geq spark(A)/2$, which concludes the proof. □

Theorem 2 provides a way to characterize a solution to the compressed sensing problem, and a criterion to assess its optimality. It must be stressed how beautiful and inexpected is this result: there exists a criterion that allows to determine whether the solution is optimal or not and this relaxes the need to worry about local minima in the cost function.

Nonetheless, one has to remember that computing the spark of a matrix is a difficult problem (actually, it is at least as difficult as solving the optimization problem). Computable bounds on the spark are then needed.

Uniqueness via the *mutual coherence* one traditional bound on the spark is called the *mutual coherence* and measures the degree of redundancy between the columns of a measurement matrix:

Definition 2 (Mutual coherence)**.** *The mutual coherence of a matrix A, $\mu(A)$ is the maximal inner product between two columns of A (assuming that the columns are ℓ_2 normalized), that is:*

$$\mu(A) = \max_{1 \leq i,j < m, i \neq j} \frac{|a_i^t a_j|}{||a_i||_{\ell_2} ||a_j||_{\ell_2}}$$

For a matrix where the columns are pairwise orthogonal (for instance an identity matrix), the mutual coherence is zero. In the compressed sensing setting, A has more columns than rows, and thus $\mu(A) > 0$. It has been shown (Donoho and Huo, 2001) that random matrices tend to have low mutual coherence. Such a property is very interesting for comprxessed sensing applications, as emphasized by Theorem 3. It can be noted that $\mu(A)$ can be computed in polynomial time.

Theorem 3 (Uniqueness – mutual coherence, from Elad, 2010)**.**

$$spark(A) \geq 1 + \frac{1}{\mu(A)}$$

Proof. see Donoho and Elad, 2003 or Elad, 2010. □

Finally, it has to be noted that mutual coherence is really a rough estimate of the spark, and not an equivalent property. Nonetheless, this criterion will be used to design good sensing matrices.

Conclusion Although the compressed sensing measurement problem leads to a (sometimes highly) underparametrized problem, one can exhibit a computable criterion (the *mutual coherence*) that ensure the uniqueness and the characterization of solutions to the problem.

6.2.3 Theoretical guarantees

Once one has shown the existence of a unique solution to the original problem, one might wonder: how close is this solution from the original signal? This closeness can be expressed in terms of mean squared error between the original signal $\tilde{x}$ and x the output of an algorithm solving the optimization problem ($MSE(x, \tilde{x}) = \frac{||x-\tilde{x}||_{\ell_2}}{||\tilde{x}||_{\ell_2}}$), and we are particularly interested in how this error scales with the number of measurements (inversely proportional to the compression ratio). Otherwise stated, one seeks the relation between the error and the compression rate, i.e. what do we lose when we choose to collect, let's say, ten times less data than in the traditional acquisition scheme?

Restricted isometry property First, it is obvious that the efficiency of the reconstruction strongly depends on the ability of the sensing matrix to capture the features of the signal. For instance, if one uses a matrix whose coefficients are all equals, then the resolution will be very low. On the other hand, a measurement matrix where all the columns are orthogonal will have the ability to capture more information with less coefficients.

Remarkably, this property can be partly summarized in one figure: the order of the *restricted isometry property* or RIP, introduced in Terence Tao and Candès, 2004. This constant characterizes a sensing matrix and is defined as follows:

Definition 3 (Restricted isometry property, from E.J. Candes and M. Wakin, 2008)**.** *Consider a measurement matrix $A \in \mathbb{R}^{m \times n}, n > m$ where the columns are ℓ_2 normalized and A_s a subset of $s < n$ columns of A. A is said to satisfy a s-restricted isometry property with constant δ_s if for any $m \times s$ submatrix and for every vector y:*

$$(1-\delta_s)||y||_2^2 \leq ||A_s y||_2^2 \leq (1+\delta_s)||y||_2^2$$

The fact that one takes a subset of s columns from A can also be interpreted as imposing that y is s-sparse, and one can then define the restricted isometry constant for a given sparisty s of y: δ_s as the smallest δ such as:

$$(1-\delta_s)||y||_2^2 \leq ||Ay||_2^2 \leq (1+\delta_s)||y||_2^2, y \text{ } s\text{-sparse}$$

That is, for any subset of columns A behaves, up to a factor δ_s similarly as an orthogonal transform that lose/gain no energy: it approximately preserves the magitude of the the s-sparse vectors y.

Note that computing δ_s is a hard problem.

Link with compressed sensing When a measurement matrix A satisfies the s-restricted isometry property with constant δ_s, the distances between any two s-sparse vectors are conserved up to a factor δ_s. When δ_s is sufficiently small, "little information is lost" and (among many others), the following theorem holds:

Theorem 4 (Reconstruction error, from E.J. Candes and M. Wakin, 2008)**.** *Let $x^\star$ be the solution to the following optimisation problem:*

$$\min_x ||x||_{\ell_1} \ s.t. \ y = Ax (= A\tilde{x})$$

with $\tilde{x}$ the original vector before compression. Assume $\delta_{2s} < \sqrt{2} - 1$ and denote $\bar{x}_S$ the vector $\tilde{x}$ where all but the S largest coefficients have been set to 0. Then the reconstruction error obeys:

$$||x^\star - \tilde{x}||_{\ell_2} \leq C_0 ||\tilde{x} - x_s||_{\ell_2} / \sqrt{s}$$

Note that when the original $\tilde{x}$ vector is s-sparse, then $\tilde{x} - \bar{x}_s = 0$ and the reconstruction is exact.

This theorem states that provided a s-sparse signal measured with m measurements $m < n$ (and possibly $m \ll n$) and whose measurement matrix satisfies a s-RIP of constant δ_s, an *exact* solution can be derived from the optimization problem.

Sensing matrices satisfying the RIP Now that one has a criterion allowing to perform *exact* reconstructions in the compressed sensing setting, provided that A satisfies a given restricted property, one seeks a way to find such matrices.

A very interesting result that will conclude this introduction is the following: any random matrix (with Gaussian or so-called sub-Gaussian distribution) satisfies the RIP with high probability (exponentially small in m) provided that (Candès and Terence Tao, 2004):

$$m \geq C \cdot s \log(\frac{n}{s}) \tag{6}$$

Stated differently, it is possible to recover exactly a s-sparse signal using only m measurements (with m satisfying Eq. 6) with a very high probability using a random matrix as sensing matrix.

6.3 Algorithms for compressed sensing

6.3.1 A simple algorithm – OMP

In this subsection we present a simple (yet empirically efficient) algorithm for compressed sensing reconstructions: Orthogonal Matching Pursuit (OMP, detailed in Elad, 2010 and analyzed in Davenport and M. B. Wakin, 2010):

To recover a s-sparse signal of length n from a compressed measurement y of length $m < n$ can be envisioned as a two-step process:

1. *Estimation of the support*: the location of the s non-zero coefficients is determined,
2. *Estimation of the coefficients*: the value of these s coefficients is estimated.

In most of the cases, the sparsity s of the input vector x is unknown. Once the set of non-zero coefficients has been estimated, simple methods such as least square regression can be used to determine their magnitude.

From this intuition, it is easy to create an algorithm that will solve the compressed sensing problem. One can cycle over all the possible supports, perform a least square estimate and finally select the sparse vector x that has the lowest residual $||y - Ax||_{\ell_2}$. Although rigourous, this algorithm is obviously intractable, because it runs in non-polynomial, exponential time and requires to iterate over all the possible supports, that is, to iterate over 2^n elements, where n is the length of x.

Nonetheless, it is possible to derive an approximate algorithm from this intuition: one can imagine progressively estimating the support of x, starting with a $(n-1)$-sparse vector and adding new elements one by one until a convergence criterion is reached. Such an algorithm is called *orthogonal matching pursuit* and can be stated as in Algorithm 1 (from Elad, 2010).

Note that one of the key points on the speed of this algorithm is the fact that minimizations of the form $\arg\min_x ||A_{\mathcal{S}^k}x - y||^2_{\ell_2}$ can be solved analytically (least squares), thus, no iterative optimization subroutine is ever needed.

We present below several algorithm with good enpirical performances.

Data: compressed vector y, sensing matrix A, convergence criterion ϵ
Result: x, a sparse solution to the $||y - Ax||_{\ell_2}$ problem
Initialization
- Initialize support $\mathcal{S}^0 \leftarrow \emptyset$;
- Initialize solution $x^0 \leftarrow 0$;
- Initialize residual $r^0 \leftarrow y - Ax^0 = b$;
- $k = 0$;

while *Convergence not reached:* $||r^k||_{\ell_2} > \epsilon$ **do**
    $k \leftarrow k + 1$;
    for *each column* a_j *of* A **do**
        Compute $\epsilon_j \leftarrow \min_x ||a_j x - r^{k-1}||^2_{\ell_2}$;
        that is the value of x that minimizes the residual as much as possible when applied to column j of the matrix.
    end
    Update support
    - $j^\star \leftarrow \arg\min_{j \notin \mathcal{S}^{k-1}} \epsilon_j$;
    - $\mathcal{S}^k \leftarrow \mathcal{S}^{k-1} \cup j^\star$;
    Update solution
    - $x^k \leftarrow \arg\min_x ||A_{\mathcal{S}^k}x - y||^2_{\ell_2}$ where $A_{\mathcal{S}^k}$ stands for the restriction of A to the columns index are in $\mathcal{S}^k$.
    - Update residual: $r^k = y - Ax^k$
end

Algorithm 1: Orthogonal matching pursuit

6.3.2 Algorithms for compressed sensing

Since the compressed sensing problem can be restated in several equivalent forms, literally hundreds of algorithms have been derived to solve the problem[9]. Unfortunately, to our knowledge, no study ever tried to aggregate the performances of the algorithms in order to assess the conditions under which the algorithms performs best. A useful starting point is Becker, Bobin, and Candès, 2009.

Here, we give some key characteristics of the algorithms used in this work.

ℓ_1- `magic` (Emmanuel Candes and Justin Romberg, 2005) is presented as a "toy" library to solve several compressed sensing problem. We used the implementation that solves the so-called *basis-pursuit* version:

$$\min_x ||x||_{\ell_1} \text{ s.t. } y = Ax$$

Note that this problem actually differs from Eq. 1 only by the use of the ℓ_1-norm instead of the ℓ_0-pseudonorm.

[9] A very extensive list of compressed sensing reconstruction algorithms can be found on Igor Carron's webpage. Many of them are also reviewed on the Nuit Blanche website by the same author.

ℓ_1- `magic` solves this problem[10, 11] by using a so-called interior point method (Newton method) inside a so-called primal-dual algorithm, a second-order method.

Over the years, ℓ_1- `magic` has become a reference implementation for beginners.

`NESTA` (Nesterov's algorithm, Becker, Bobin, and Candès, 2009) was introduced as an intermediate trade-off between fast, low accuracy, first-order methods and slow, high accuracy second-order methods.

`NESTA` solves the following ℓ_1-relaxed basis pursuit problem[11, 12]:

$$\min_x ||x||_{\ell_1} \text{ s.t. } ||y - Ax|| \leq \epsilon$$

Resolution of the problem is achieved by replacing the traditional non-smooth ℓ_1-norm by a smoothed version, and then applying the so-called Nesterov's method to solve the regularized problem.

`lasso` (least absolute shrinkage and selection operator) It is interesting to see that Eq. 1 can also be equivalently reformulated as a `lasso` problem, a classical sparse regression widely used in statistics. Indeed, the problem can be restated as:

$$\min_x ||y - Ax||_{\ell_2}^2 + \lambda ||x||_{\ell_1}$$

where λ governs the trade-off between accuracy of the reconstruction (low λ) and the sparsity of reconstruction (high λ).

6.3.3 Algorithm for Poisson noise settings

The generic algorithms highlighted below usually perform well in the general case. They can, however, underperform when applied to real-life compressed sensing imaging settings. Several factors can explain this underperformance:

1. As described above, image acquisitions are contaminated with Poisson noise, a signal-dependent perturbation that is not accounted for by the traditional relaxed compressed sensing solvers.

2. Images only have positive values (no negative photon count).

Fortunately, some algorithms have been developed to deal with the specific constraints of image sensing. We present one of them:

`SPIRAL-TAP` (Harmany, Marcia, and Willett, 2012) is a a Matlab™ code that uses a penalized Poisson likelihood with positivity constraints to ensure the quality of reconstruction. Indeed, under a Poisson model as described in Section 1.3.2 the probability of observing a given compressed measurement vector y of length m from the original vector x is given by the following likelihood:

$$\mathbb{P}(y|Ax) = \prod_{i=1}^{m} \frac{((Ax)_i)^{y_i}}{y_i} \exp(-(Ax)_i)$$

where $(Ax)_i$ denotes the i th element of the vector Ax.

From this likelihood, the negative log-likelihood penalty $F(x) = -\log(\mathbb{P}(y|Ax))$ is derived, and the `SPIRAL-TAP` algorithm solves[13] the following constrained problem:

$$\min_x ||x||_{\ell_1} + \lambda F(x) \text{ s.t. } x \geq 0$$

This problem is solved by using a second-order Taylor expansion of the Poisson likelihood term. Such approximation brings separability and allows to derive subproblems that can then be solved by gradient descent.

Conclusion Although the noisy setting significantly differs from the noiseless one, there exists both theoretical grounds and efficient algorithms that allow to perform reconstructions with an error bounded by a function usually linear in the noise level.

6.4 Dictionary learning

In this section, we examine two key properties that allow the use of sparsifying transforms in the context of compressed sensing:

1. If the signal has a sparse representation in some known basis, all the compressed sensing theory applies with very little modification.

2. One can show that an uncertainty principle holds that states that a signal cannot be spread out in all of its representations.

[10] A Matlab™ implementation of ℓ_l- `magic` is available.

[11] This code also has a Python implementation.

[12] A Matlab™ implementation of `NESTA` is available.

[13] Both a Matlab™ and a Python implementation are available.

6.4.1 Existence of sparse representations: an uncertainty principle

Assume that one can apply the compressed sensing theory and algorithms to a signal that has a sparse representation in some *known* basis. Then a natural question arises: how often does a signal admit a sparse representation in some basis? How can such a basis be found? To first give an intuition of the existence of sparse representation, let us xsconsider the following example:

Consider a signal f in the direct/Dirac domain and its Fourier transform F. Then, assuming that f is ℓ_2-normalized, one can derive a so-called *uncertainty principle*:

$$\int_{-\infty}^{\infty} x^2|f(x)|^2dx \cdot \int_{-\infty}^{\infty} \omega^2|F(\omega)|^2d\omega \geq \frac{1}{2}$$

Note that the quantity $x^2|f(x)|^2$ (resp. $\omega^2|F(\omega)|^2$) expresses the idea of concentration in time (resp. frequency). Thus, one can interpret this relation by saying that a signal cannot be infinitely concentrated both in time and frequency. This uncertainty principle is similar to the Heisenberg uncertainty principle in physics. Closer to our compressed sensing setup, this principle translates into the fact that a signal cannot be sparse (or conversely, non sparse) in both the Dirac and the Fourier basis. This principle can be further generalized to orthobases, and requires a definition of the proximity of two orthobases. This definition is a small extension of the *coherence* stated above (Definition 2 :

Definition 4 (Mutual coherence for two orthobasis of size $n \times n$, from Elad, 2010)**.** *Assume Φ and Ψ two orthobases of size $n \times n$ and let A their concatenation $A = [\Phi, \Psi]$, then the mutual coherence for two orthobases $\mu(A)$ is defined as :*

$$\mu(A) = \mu([\Phi, \Psi]) = \max_{1 \leq i,j \leq n} |\Phi_i^t \Psi_j|$$

One can then consider the following setup: assume a vector b and its representation in two different orthobases Φ and Ψ of mutual coherence $\mu(A)$: $b = \Phi\alpha = \Psi\beta$. Then, the following uncertainty principle holds, similar to the one with Fourier/direct orthobases:

Theorem 5 (Uncertainty principle for orthobasis)**.** *For arbitrary orthobases Φ and Ψ of mutual coherence $\mu(A)$, a vector b and α (resp. β) its representation in the Φ basis (resp. Ψ), i.e. $b = \Phi\alpha = \Psi\beta$, then:*

$$||\alpha||_{\ell_0} + ||\beta||_{\ell_0} \leq \frac{2}{\mu(A)}$$

Indeed, recall that the mutual coherence is bounded $1/\sqrt{n} \leq \mu(A) \leq 1$, thus, provided a pair of orthobasis with mutual coherence low enough, this theorem guarantees the existence of a sparse representation.

Proof. See Elad, 2010 □

Finally, knowing the existence of a sparsifying transform doesn't tell which transform to use.

6.4.2 From sparsifying basis to dictionaries

Remarkably, sparsifying transforms do not need to be a basis. Indeed, sometimes signals can be "sparsified" using more vector elements than the ones in a basis. In this case the matrix D is called a *dictionary* and has more lines than columns (and sometimes many more lines than columns).

First, it is interesting to notice that the whole compressed setting can be applied without modification. Furthermore, the range of dictionaries that are traditionally used is wider than with basis. Examples include curvelet dictionaries, redundant Discrete Fourier Transform, Gabor frames, etc.

One of the main differences between a sparsifying basis and a dictionary is that there is no bijection anymore between the reconstructed signal x' and the original x: because the dictionary has more lines than columns, several values of x' can lead to the same vector x. This is why such dictionaries are often called *redundant* dictionaries.

Although seductive, several questions arise. First, the actual basis used for reconstruction is the matrix product AD. We can then wonder which properties both the matrix AD and D should follow in order to provide good reconstruction guarantees. For instance, what are the constraints on the RIP in this setup?

What happens to the RIP? This situation has been studied in details by Emmanuel Candes, Eldar, et al., 2010, and a general introduction can be found in X. Chen et al., 2014.

To do so, the authors introduce Σ_s the union of all subspaces created by keeping s columns of the dictionary D (of size $N \times n$) and the so-called *D-RIP*:

Definition 5 (RIP adapted to D – D-RIP, from Emmanuel Candes, Eldar, et al., 2010)**.** *The measurement matrix A obeys a restricted isometry property adapted to D (D-RIP) of constant δ_s if for any x an image by D of a s-sparse vector the following relation holds:*

$$(1-\delta_s)||x||_{\ell_2}^2 \leq ||Ax||_{\ell_2}^2 \leq (1-\delta_s)||x||_{\ell_2}^2$$

Note that Σ_s can be seen as the set of the images by D of s-sparse vectors.

In a very analogous manner as Theorem 4, the authors show that for a tight frame (that is, $\forall v \in \mathbb{R}^n, \sum_{1\leq i\leq N} |\langle v, D_i\rangle|^2 = ||v||^2$, where $(D_i)_{1\ldots N}$ denote the columns of D):

Theorem 6 (Reconstruction guarantee under D-RIP, from Emmanuel Candes, Eldar, et al., 2010)**.** *For D an arbitrary tight frame and a measurement matrix A satisfying the D-RIP of constant $\delta_s < 0.008$, then the recovered solution x compares to the original s-sparse signal $x^\star$ as follows:*

$$||x - x^\star||_{\ell_2} \leq C_0\epsilon$$

where ϵ is the error-tolerance constant in the relaxed formulation of compressed sensing (Eq. 2).

Note that the original article also provides guarantees for approximately-sparse signals. This result states that for an arbitrary tight frame, one has reconstruction guarantees provided that the sensing matrix satisfies the D-RIP, and this theoretically justifies the use of a dictionary in the compressed sensing setting. Also, assuming that A satisfies a D-RIP is a significantly weaker asumption than requiring AD to satisfy the traditional RIP.

Are we really optimizing for the right thing? As stated above, there is no bijection between the sparse, intermediate representation x' and the solution in the original domain x (or signal space). In the case of a sparsifying basis, it is equivalent to obtain a reconstruction guarantee in the intermediate space ($||x' - x'^\star||_{\ell_2}$) than the signal space ($||x - x^\star||_{\ell_2}$). However, when one moves from a basis to a dictionary/frame, those two errors can become highly decorrelated.

To circumvent this issue, Davenport, Deanna Needell, and M. B. Wakin, 2013 propose an algorithm, called `SS-CoSaMP` (standing for *signal space* compressive sampling matching pursuit). `SS-CoSaMP` adapts the widely used `CoSaMP` algorithm (D. Needell and Tropp, 2008) to a signal space paradigm. Indeed, reconstruction guarantees for this algorithm are derived in signal space rather than in the intermediate space and reads at iteration l of the algorithm:

$$||x^\star - x^l||_{\ell_2} \leq 2^{-l}||x^\star||_{\ell_2} + C_2\epsilon$$

Despite its theoretical advantage, `SS-CoSaMP` seemed to perform poorly in our setting, and it wasn't used in the further analysis.

Conclusion In this subsection, the potential of dictionaries with respect to sensing bases has been demonstrated from a theoretical standpoint.

6.4.3 How to find a dictionary?

Now that the theoretical requirements to perform compressed sensing with a sparsifying basis/dictionary are established, one seeks real-life algorithms to find such dictionaries. Two approaches can be undertaken:

1. Blind guesses: take a generic transform that usually work well with the type of features considered. For instance, wavelet basis are usually suitable for image compression, and curvelets seem to achieve one of the best performance.

2. Learn a representation from sample data. Provided that the sample data has enough diversity, such approach usually gives the best results.

Principles of dictionary learning Dictionary learning is a *data-dependent* method that can be used to learn a dictionary in which most of the observed signals have a sparse representation (X. Chen et al., 2014). Indeed, one seeks a dictionary that has the following property:

For a training set of k vectors of length n (k signals acquired in the signal space: $x_1, x_2, \ldots, x_k \in \mathbb{R}^n$), one seeks:

1. a dictionary $D \in \mathbb{R}^{n\times N}$ (where N is a free parameter characterizing the size of the dictionary, also called *number of atoms*),

2. A series of k vectors of coefficients $\gamma_1, \gamma_2, \ldots, \gamma_k \in \mathbb{R}^N$ such that $\forall i, D\gamma_i$ is close to x_i.

That is, the dictionary is able to reconstruct the training set accurately. Since the dictionary that is being built is *overcomplete* (usually $N \gg n$), it can recreate any vector of the training set. An important feature of such algorithms is thus the sparsity of the output, that is, whether the dictionary is indeed able to encode the structure of the training set into a low number of coefficients of a sparse vector.

This problem can be recast as a traditional optimization problem:

$$\min_{D,\Gamma} ||X - D\Gamma||_F^2 \text{ s.t. } ||\gamma||_0 \leq \eta, \forall i$$

where D is the dictionary and Γ is a $N \times k$ matrix of coefficients and η a parameter to enforce sparsity of the representation. $|| \cdot ||_F$ stands for the Fröbenius norm.

6.5 Effect of dictionary parameters

When performing dictionary learning with `NMF` (or equivalently, `sparse-NMF`, see Section 6.4.3), two paramenters can be adjusted, that will impact the quality of the dictionary:

1. The number of atoms in the dictionary
2. The size of the training set, that is the number of sample vectors used to train the dictionary.

Ideally, one wants those two parametrs as big as possible, but in practice, although building enormous (>20 times redundant) dictionaries is feasible, reconstructions are impractical. Indeed, since the matrix used for the compressed sensing reconstruction is the product AD of the measurement matrix A with the dictionary D, having D a very large matrix makes the product AD too large to achieve reconstructions in areasonable time.

On the other hands, the size of the training set is mostly limited by the amount of data available.

To find the best trade-off for those two parameters, we first work with a $\sim$ 2 times overcomplete dictionary and vary the size of the training set, from five to twenty frames. The sample frames are extracted from a 3D stack and are regularly spaced. The reconstructed image is not part of the training set (although arising from the same 3D stack).

Results are presented in Figure 31 for two dictionary learnign algorithms: `NMF` and `sparse-NMF`.

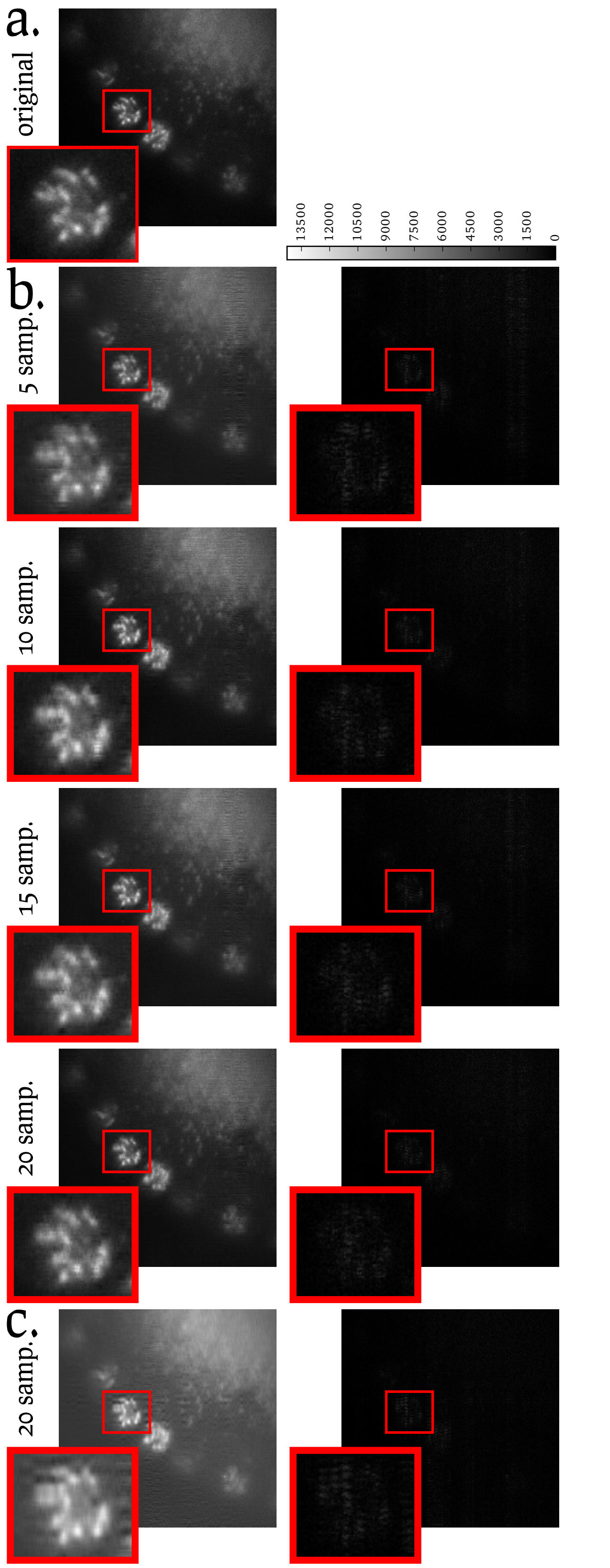

Figure 31: Compressed sensing reconstructions obtained usiong various dictionaries. **(a).** Original image, **(b)-left.** reconstructions using a dictionary obtained with `NMF` (1200 atoms, various size of the training set, from 5 sample images – *top* – to 20 – *bottom* –), and **(b)-right.** the corresponding absolute error, the colorscale is the same as the reconstructions. **(c).** reconstruction using a dictionary assembled by the `sparse-NMF` algorithm (1600 atoms, 20 sample images in the training set).

First, note that since the size of the dictionary remains fixed (only the size of the training set varies), the reconstructions take approximately the same time to run. It appears that the number of samples/size of the training set significantly impacts the quality of reconstruction. Indeed, when zooming for instance on Figure 31b-5. samp., one can spot horizontal stripes in some parts of the image (the most feature rich in general). Such artifacts are much less visible when the dictionary has been made from 20 samples.

Then, the `sparse-NMF` algorithm performs really poorly compared to traditional `NMF` (Figure 31), with very visible artifacts. Such artifacts remain striking even with a higher number of atoms and a high number of samples.

To balance the trade-off, one has to examine the time to build the dictionary, in order to keep the pipeline running time reasonable. An assessment of the dictionary learning times for the `NMF` algorithm is provided Figure 32. Note that the figures solely include the time taken to learn the dictionary, not the time to perform the reconstruction (which is usually much longer as soon as 3D volumes are concerned).

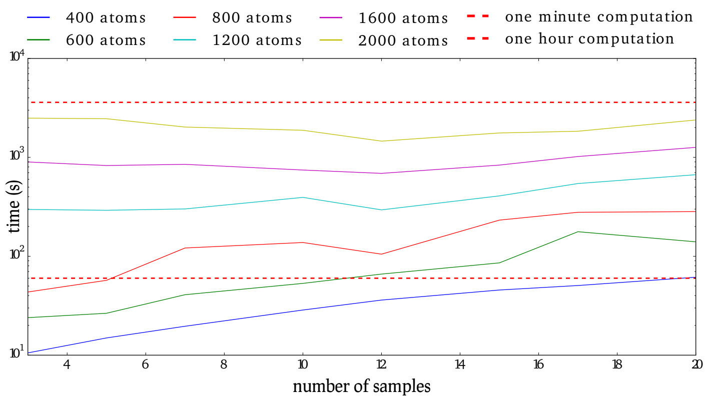

Figure 32: Benchmarking of the `NMF` algorithm. Time taken to build a dictionary with a given number of atoms from a given number of sample images. Sample images have size `512x512`. Time is expressed in logarithmic scale. Horizontal dashed line represent one minute and one hour computation. Test run on a 32-core Intel(R) Xeon(R) CPU E5-2670 0 @ 2.60GHz, 64 GB RAM running Ubuntu 14.04 64 bits (Linux 3.13)

As a result of this benchmark, it first appears that the computation time to build a dictionary remains reasonable for all the parameters tested. Indeed, since the dictionary only has to be generated once per data type (and can be precomputed), computation times less than a few hours are regarded as acceptable.

Furthermore, increasing the size of the training set only marginally increases the computation time. Thus, it is always advantageous to build dictionaries from more sample images (since it make no difference in the reconstruction time and it significantly improves the reconstruction quality).

6.6 Multidimensional case

As a first approximation, there is a one-to-one mapping between a (x, y) position on the sample and a (x, y) position on the camera, which prevents compressed sensing to be applied in more than one dimension. However, this is only a first order approximation. Indeed, due to the fact that the signal coming through the objective is bandlimited, any feature appears convolved by a point-spread function (PSF, see Kim, 2011). The PSF can be incorporated in the measurement matrix, leading to a multidimensional model.

A 3D, uncompressed volume of dimension $l \times w \times h$ can be seen as an "unfolded" 1D vector of length lwh. The measurement matrix also has to be "unfolded". Assuming true compression only in the third dimension (h) and h' the number of measurement in the compressed mode, the measurement matrix A has dimension $h' \times h$ (recall that all the (x, y) positions are independent. Then, the corresponding "unfolded" measurement matrix is a block-diagonal matrix of dimension $lwh' \times lwh$. An example on a 2D image in the (x, z) plane is shown in Figure 33.

Such an approach has been described and implemented on 2D+time acquisitions (M. B. Wakin et al., 2006, Ramirez-Giraldo et al., 2011), and GPU-optimised code has been published (Nett, Tang, and G.-H. Chen, 2010). Unfortunately, the authors did not release the code, and did not respond to our emails.

Although more realistic because it incorporates some physical information about the PSF, this model is dramatically more computationally expensive to solve because we moved from $l \times w$ problems of size h to one problem of size $l \times w \times h$. Since all the ℓ_1 solvers have a complexity always strictly worse than linear, the computational time would explode in practice.

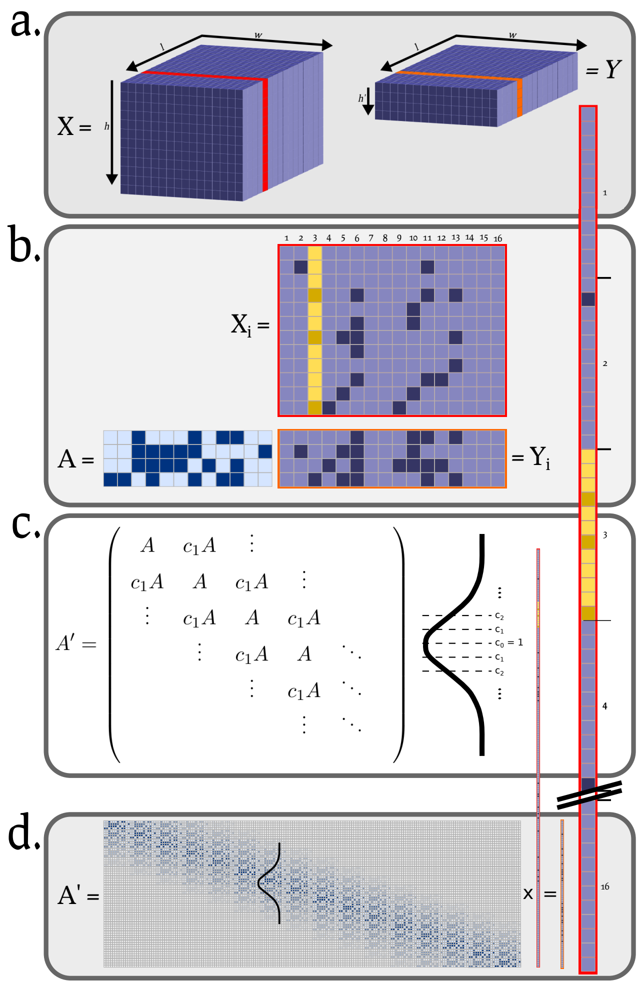

Figure 33: Towards a multidimensional model to take into account the PSF. **(a).** representation of (*left*) an uncompressed 3D acquisition X of $16 \times 16 \times 12$ voxels and (*right*) its Y counterpart compressed in z, of dimensions $16 \times 16 \times 4$ voxels. Following panels focus on the uncompressed (resp. compressed) (x, z) plane X_i (resp. Y_i) highlighted in red (resp. orange). **(b).** traditional compression scheme with a measurement matrix A. Each columns of X_i are measured independently. But the matrix X_i can also be seen as a column vector of length $16 * 12 = 192$ with all its columns vertically concatenated (long vector on the right), and a block diagonal matrix can be constructed to measure this vector and provide an equivalent output as the independently measured columns (not shown), **(c).** moreover, the new sensing matrix A' can incorporate some information about the PSF (left). A sample PSF is pictured on the right. The coefficients c_i reprensent the PSF value at the pixels coordinates, **(d).** resulting measurement matrix A' (it might be necessary to zoom to appreciate the details).

An additional hurdle is the size of the sensing matrix. Indeed, assuming a 3D stack of dimensions $512 \times 512 \times 101$ and a factor 10 compression in z, the sensing matrix would have a size of $512 * 512 * 10 \times 512 * 512 * 101 \simeq 2.6.10^6 \times 26.10^6 \simeq 69.10^{12}$ pixels, which is at least three orders of magnitude higher than the RAM of modern computers.

However, since the PSF vanishes very quickly, one can assume that it has bounded support. One can then subdivide the "unfolded" vector into subvectors that incorporate all the PSF-induced correlation, but discards pixels far apart, since they those can be processed in parallel. We implemented such a scheme when possible.

6.7 Software contributed during the internship

During the course of the internship, the following free/open source softwares were contributed (Table 3). Most of the contributions are bug reports or modifications to get the software working with `Python 3`. However, several Python ports of existing code were realized (`sparseNMF` and `pySPIRALTAP`).

Table 3: Summary of softwares contributed during the internship

Software	Description	Language	Contribution	Link
`libtiff`	Read/Write 3D TIFF files	Python	Debugging for a Python 3 version	(1)
`rwt`	Rice Wavelets Toolbox	Python	Debugging the Python 2 installation	(2)
`SPIRALTAP`	Compressed sensing solver	Matlab	Fixed several I/O bugs + Octave port	(3) and (4)
`pySPIRALTAP`	Python port of `SPIRALTAP`	Python	Initiation of the port	(5)
`sparse-NMF`	Another approach to `NMF`	Python	Initiation of the port from Matlab	(6)
`pyCSalgos`	Several CS solvers	Python	Port to Python 3	(7)

1. `http://github.com/pearu/pylibtiff/`
2. `https://github.com/ricedsp/rwt`
3. `http://drz.ac/code/spiraltap/`
4. `https://gitlab.com/padouppadoup/SPIRAL-TAP`
5. `https://gitlab.com/padouppadoup/pySPIRAL-TAP`
6. `https://gitlab.com/padouppadoup/sparseNMF`
7. `https://gitlab.com/padouppadoup/pyCSalgos`