Computational Interferometry for Hyperspectral Imaging

Amirafshar Moshtaghpour

ICTEAM Institute

Université catholique de Louvain

This dissertation is submitted for the degree of

Ph.D. in Engineering Sciences

Louvain School of Engineering

November 2019

Ph.D. Committee

Prof. Laurent Jacques, *supervisor*, Université catholique de Louvain, Belgium
Prof. José M. Bioucas-Dias, Universidade de Lisboa, Portugal
Prof. Christophe De Vleeschouwer, Université catholique de Louvain, Belgium
Prof. Nicolas Gillis, Université de Mons, Belgium
Dr. Philippe Antoine, Lambda-X SA, Belgium
Prof. David Bol, *president*, Université catholique de Louvain, Belgium

ABSTRACT

The idea of capturing hundreds to thousands images of an object, *e.g.*, biological specimen, in different wavelengths has attracted decades of research under the name of HyperSpectral (HS) imaging. Since every chemical element in the universe has a unique signature in the spectral domain, obtaining such an HS data volume allows for characterizing the object elements, their concentration, and growth. This has made HS imaging an appealing tool in ubiquitous applications, *e.g.*, chemistry, food science, agriculture, astronomy, and biology.

Among a profusion of HS imaging techniques, the Fourier Transform Interferometry (FTI) has received a renewed interest for its high spectral resolution capability; a crucial criterion for, *e.g.*, biomedical fluorescence spectroscopy. However, the effective resolution of the FTI is limited by the durability of biological elements when exposed to illuminating light, as over-exposed elements become unable to fluoresce.

The main objective of this thesis is to investigate the possibilities to reduce the light exposure received by the observed object while preserving the spectral resolution of the reconstructed HS volumes. We propose novel computational interferometric techniques for acquisition of HS volumes based on the theory of compressive sensing with variable density sampling. In particular, we introduce three variations of the FTI system, *i.e.*, Coded-Illumination FTI (CI-FTI), Structured-Illumination FTI (SI-FTI), and compressive Single-Pixel FTI (SP-FTI), based on the theory of compressive sensing. In these three contexts, our theoretical analysis addresses two main questions: *(i)* how to efficiently modulate

light exposure temporally (in CI-FTI) or spatiotemporally (in SI-FTI and compressive SP-FTI) in order to minimize the light exposure imposed on the observed object? and *(ii)* how to recover high quality HS volumes from CI-FTI, SI-FTI, and compressive SP-FTI measurements?

ACKNOWLEDGEMENTS

After more than four and a half years of reading papers, implementing numerical codes, thinking that the code is correct and then finding mistakes, countless meetings, writing manuscripts, few successes, and plenty of failures, this is time to stop, look back and thank the people who have made a difference in my personal and scientific life.

First and foremost, this thesis would not have been accomplished without the support of my advisor, Prof. Laurent Jacques. I have been extremely lucky to have such a brilliant, patient, humble, and encouraging supervisor during my Ph.D. studies. He was always available for my questions and cared so much about my work and life. I also value his personality that has enriched mine.

I would like to thank the jury members of my thesis, Prof. Nicolas Gillis, Prof. Christophe De Vleeschouwer, Dr. Philippe Antoine, Prof. José Bioucas-Dias, and Prof. David Bol for reviewing the thesis and providing valuable comments. As my initial co-supervisor and my post-doc supervisor, Prof. C. De Vleeschouwer has been of great help and influence. I appreciate his help in the first year of my Ph.D. research. I thank Dr. Philippe Antoine and Matthieu Roblin for their practical advices and helping me to acquire a set of real data used in the simulations of Chapter 4.

During my visiting period (from Jan., 2018 to Apr., 2018) at IST institute in Lisbon, Portugal, I have been kindly hosted by Prof. José Bioucas-Dias and Prof. Mario Figueiredo. It really was a magnificent experience to work with Prof. J. Bioucas-Dias whose valuable advises initiated the endeavour reported in Chapter 3 and Chapter 5 of this thesis.

I thank Valerio Cambareri for mentoring me during two years of my Ph.D. that includes a myriad of discussions and revisions of my reports and drafts. I would like to thank Gilles Puy for his enlightening advices on estimation of the power of the weighted noise embedded in Chapter 2. I thank Ben Adcock for the valuable discussions I had with him during the iTWIST'18 workshop (Marseille, France). I also thank Mehrdad Yaghoubi for his useful remarks during SPARS'19 workshop (Toulouse, France). A very special thank you to *the gifted scientist* Ulugbek Kamilov for being an inspiring example for me. I must say that the scientific events without him were boring.

My special thanks goes to *the ultimate care* Nafiseh Janatian for helping me with proofreading my manuscripts, encouraging me at all times especially in those stressful moments of my Ph.D., and making the life in abroad easier for me. I extend the special thanks to my office mates, *the absolute kindness* Adriana Gonzalez, *the silent girl* Chunlei Xu, *the cool dude* Thomas Feuillen, *the polite boy* Antoine Paris, and *the cute couple* Stéphanie Guérit and Pierre-Yves Gousenbourger for sharing incredible moments, making countless jokes, and soothing each other after receiving the manuscript reviews.

I would like to thank the present and past members of the corridor, Vincent Schellekens, Benoît Pairet, Simon Carbonnelle, Maxime Istasse, François Rottenberg, Tahani Madmad, Anne-Sophie Collin, Victor Joos, Gabriel Van Zandycke, Antoine Vanderschueren, Hussein Kassab, Mohammad Arsalan, Mohieddine El Soussi, Kévin Degraux, Cedric Verleysen, Amit Kumar, Julien Moreau, Pascaline Parisot, Arnaud Browet, Arnaud Delmotte. I thank Prof. Danielle Janvier for helping me with the administrative issues. I also thank all the technical and administrative members of the ICTEAM department, especially Brigitte Dupont, Jean Deschuyter, Isabelle Dargent, Ludzzie Ross, Patricia Focant, François

Hubin, and Souley Djadjandi. I also thank the Fonds pour la Formation à la Recherche dans l'Industrie et dans l'Agriculture (FRIA) of the FNRS for funding my Ph.D. research.

I would like to acknowledge Margarida Reis, Marina Ljubenovic, Milad Niknejad, Tenmay Verlekar, Alireza Sepas-Moghaddam, Alireza Javaheri, Falah Jabar, André Guarda, and Marco Pezzutto for creating a friendly atmosphere in the lab (during my stay in Lisbon) and playing Undercover game during coffee and lunch breaks.

Thanks are extended to my foosball comrades, Nafiseh Janatian, Ipek Akin, Ivan Stupia, Maxime Schramme, Antoine Paris, and Hamed Mirghasemi for creating great moments of victory and defeat.

Finally, my special gratitude goes to my mother (Rouhangiz Amiri), father (Ali Moshtaghpour), brother (Reza Moshtaghpour), and two sisters (Solmaz and Golnaz Moshtaghpour). They deserve enormous thanks for raising me, supporting me, teaching me, shaping my soul, and making every bit of my life unique. They have always been a supportive voice of reason for all of my decisions that remains so to this day.

List of Acronyms

1-D one-dimensional

2-D two-dimensional

3-D three-dimensional

ADHW Anisotropic Discrete Haar Wavelet

BS Beam-Splitter

BPDN Basis Pursuit DeNoise

CI-FTI Coded Illumination FTI

CS Compressive Sensing

CASSI Coded Aperture Snapshot Imager

DFT Discrete Fourier Transform

DHW Discrete Haar Wavelet

DMD Digital Micro-mirror Devices

IDHW Isotropic Discrete Haar Wavelet

fps frame-per-second

FTI Fourier Transform Interferometry

iid independent and identically distributed

HS HyperSpectral

MI Michelson Interferometry

MRI Magnetic Resonance Imaging

MDS Multilevel Density Sampling

ME Minimal Energy

MNR Measurement-to-Noise Ratio

NDS Non-uniform Density Sampling

OPD Optical Path Difference

pmf probability mass function

RGB Red-Green-Blue

r.v. random variable

RIC Restricted Isometry Constant

RIP Restricted Isometry Property

RM Robust Median

TV Total Variation

SI-FTI Structured Illumination FTI

SLM Spatial Light Modulator

SP-FTI Single Pixel FTI

SPGL1 Spectral Projected Gradient for ℓ_1 minimization

SNR Signal-to-Noise Ratio

SRE Signal-to-Reconstruction Ratio

UDS Uniform Density Sampling

VDS Variable Density Sampling

NOTATIONS

u	denotes a scalar (lowercase letters).
$\boldsymbol{u}$	denotes a vector (lowercase boldface letters).
$\boldsymbol{U}$	denotes a matrix (uppercase boldface letters).
$\boldsymbol{\mathcal{U}}$	denotes a cube (uppercase boldface calligraphic letters).
K, M, N	denotes domain dimensions (uppercase letters).
$\boldsymbol{u}^\top, \boldsymbol{U}^\top$	denotes the transpose of a vector $\boldsymbol{u}$ or a matrix $\boldsymbol{U}$.
$\boldsymbol{u}^*, \boldsymbol{U}^*$	denotes the conjugate transpose of a vector $\boldsymbol{u}$ or a matrix $\boldsymbol{U}$.
$\boldsymbol{U}^\dagger$	denotes the pseudo-inverse of a matrix $\boldsymbol{U}$. Note that the pseudo-inverse exists for any matrix. But for a full rank matrix $\boldsymbol{A} \in \mathbb{C}^{M\times N}$, the pseudo-inverse can be expressed as an algebraic formula: $\boldsymbol{U}^\dagger = (\boldsymbol{U}^*\boldsymbol{U})^{-1}\boldsymbol{U}^*$ constitutes the left pseudo-inverse if $M \leq N$ and $\boldsymbol{U}^\dagger = \boldsymbol{U}^*(\boldsymbol{U}\boldsymbol{U}^*)^{-1}$ constitutes the right pseudo-inverse if $M \geq N$.
$u_l, (\boldsymbol{u})_l$	denotes the l^{th} entry of a vector $\boldsymbol{u} = [u_1, \cdots, u_N]^\top$.
$\boldsymbol{u}_l, (\boldsymbol{U})_{:,l}$	denotes the l^{th} column of a matrix $\boldsymbol{U} = [\boldsymbol{u}_1, \cdots, \boldsymbol{u}_N]$.
$(\boldsymbol{U})_{l,:}$	denotes the l^{th} row of a matrix $\boldsymbol{U}$.
$u_{l,l'}, (\boldsymbol{U})_{l,l'}$	denotes the $(l, l')^{\text{th}}$ component of a matrix $\boldsymbol{U}$.

$\boldsymbol{U} \otimes \boldsymbol{V}$	denotes the Kronecker product of two matrices $\boldsymbol{U}$ and $\boldsymbol{V}$.
$\|u\|$	denotes the absolute value of a scalar u.
$\\|\boldsymbol{u}\\|_p$	denotes the ℓ_p-norm of a vector $\boldsymbol{u} \in \mathbb{C}^N$, for $p \geq 1$, *i.e.*, $\\|\boldsymbol{u}\\|_p := (\sum_l \|u_l\|^p)^{1/p}$, with $\\|\boldsymbol{u}\\| := \\|\boldsymbol{u}\\|_2$.
$\\|\boldsymbol{U}\\|_{p,q}$	denotes the $\ell_{p,q}$-norm of a matrix $\boldsymbol{U} \in \mathbb{C}^{M \times N}$, for $p, q \geq 1$, *i.e.*, $\\|\boldsymbol{U}\\|_{p,q} := \max_{\boldsymbol{x}} \{\\|\boldsymbol{U}\boldsymbol{x}\\|_q \text{ s.t. } \\|\boldsymbol{x}\\|_p = 1\}$.
$\mathrm{vec}(\boldsymbol{U})$	denotes the vectorized version of a matrix $\boldsymbol{U}$, *i.e.*, for $\boldsymbol{U} = [\boldsymbol{u}_1, \cdots, \boldsymbol{u}_{N_2}] \in \mathbb{C}^{N_1 \times N_2}$, its vectorized version reads $\mathrm{vec}(\boldsymbol{U}) := [\boldsymbol{u}_1^\top, \cdots, \boldsymbol{u}_{N_2}^\top]^\top \in \mathbb{C}^{N_1 N_2}$.
$\mathbf{0}_{N \times N'}$, $\mathbf{1}_{N \times N'}$	denotes an $N \times N'$ matrix with all components equal 0 (respectively, 1). We use the shorthand $\mathbf{0}_N$ and $\mathbf{1}_N$ when $N' = 1$. Moreover, we omit the values of N and N' when they are clear from the text.
$\boldsymbol{I}_N$	denotes the identity matrix matrix of dimension N.
$\boldsymbol{F}_N$	denotes the one-dimensional (1-D) discrete Fourier matrix of dimension N. The 1-D Discrete Fourier Transform (DFT) is denoted by $\boldsymbol{F}_N^*$. We omit the value of N when it is clear from the text.
$\mathcal{S}, \mathcal{T}, \Omega$	denotes sets (calligraphic letters). By an abuse of convention, unless expressed differently, we consider that a set (or a subset) of indices is actually a *tuple*, *i.e.*, the repetition and ordering of the elements are allowed.
$\|\mathcal{S}\|$	denotes the cardinality of a set $\mathcal{S}$, *i.e.*, the total number of (non-unique) multiset elements.
$\mathcal{S} \sqcup \mathcal{T}$	denotes the order-preserving concatenation of two sets $\mathcal{S} = \{s_l\}_{l=1}^M$ and $\mathcal{T} = \{t_l\}_{l=1}^N$, *i.e.*, $\mathcal{S} \sqcup \mathcal{T} := \{s_1, \cdots, s_M, t_1, \cdots, t_N\}$.

$\mathbb{N}$	denotes the set of all natural numbers.
$\mathbb{R}$	denotes the set of all real numbers.
$\mathbb{C}$	denotes the set of all complex numbers.
$\mathbb{Z}$	denotes the set of all integers.
$[\![N]\!], [\![N]\!]_0$	denotes the set of indices, *i.e.*, $[\![N]\!] := \{1, \cdots, N\}$ and $[\![N]\!]_0 := \{0\} \cup [\![N]\!]$.
$\mathcal{S} \backslash \mathcal{T}$	denotes the relative complement of $\mathcal{T}$ in $\mathcal{S}$. This set contains all those elements of $\mathcal{S}$ that are not in $\mathcal{T}$, *i.e.*, $\mathcal{S} \backslash \mathcal{T} := \{u \in \mathcal{S} \text{ s.t. } u \notin \mathcal{T}\}$.
$\boldsymbol{P}_\Omega$	denotes the restriction operator, *i.e.*, $\boldsymbol{P}_\Omega \in \{0,1\}^{M \times N}$ for a subset $\Omega = \{\omega_l\}_{l=1}^M \in [\![N]\!]$ with $(\boldsymbol{P}_\Omega \boldsymbol{x})_l = x_{\omega_l}$.
$\bar{\boldsymbol{P}}_\Omega$	denotes the mask operator, *i.e.*, $\bar{\boldsymbol{P}}_\Omega := \boldsymbol{P}_\Omega^\top \boldsymbol{P}_\Omega \in \{0,1\}^{N \times N}$ for a subset $\Omega = \{\omega_l\}_{l=1}^M \subset [\![N]\!]$ with $(\bar{\boldsymbol{P}}_\Omega \boldsymbol{x})_l = x_l$ if $l \in \Omega$ and zero otherwise.
Σ_K	denotes the set of all $K-$sparse signals, *i.e.*, $\Sigma_K := \{\boldsymbol{u} \text{ s.t. } \lvert\mathrm{supp}(\boldsymbol{u})\rvert \leq K\}$.
mod	denotes the modulo operation.
$\log_a(u)$	denotes the logarithm function in base $a > 0$. We use the shorthand $\log(u)$ when $a = e$. Moreover, we assume that $\log_a(0) = -\infty$.
$\lfloor u \rfloor$	denotes the floor function that outputs the greatest integer less than or equal to u. We assume that $\lfloor -\infty \rfloor = -\infty$.
$(u)_+$	denotes the ramp function, *i.e.*, $(u)_+ := \max(u, 0)$.
$\delta_{l,l'}$	denotes the Kronecker function, *i.e.*, $\delta_{l,l'} = 1$ if $l = l'$ and zero otherwise.

$f \lesssim g$, $f \gtrsim g$	for two functions f and g, we write $f \lesssim g$ (or $f \gtrsim g$), if $f \leq cg$ (respectively, $f \geq cg$) for some value $c > 0$ independent of their parameters.
$\mathrm{supp}(\boldsymbol{u})$	denotes the support set of a vector $\boldsymbol{u}$, *i.e.*, $\mathrm{supp}(\boldsymbol{u}) := \{l \text{ s.t. } u_l \neq 0\}$.
$l \overset{N_1,N_2}{\rightleftharpoons} (l_1, l_2)$	denotes the index conversion rule between the 1-D and two-dimensional (2-D) index representations $l \in [\![N_1 N_2]\!]$ and $(l_1, l_2) \in [\![N_1]\!] \times [\![N_2]\!]$, respectively, meaning that $l = l_1 + N_1(l_2 - 1)$, $l_1 = (l - 1 \bmod N_1) + 1$, and $l_2 = \lfloor (l-1)/N_1 \rfloor + 1$. When $N_1 = N_2 = N$ we simply write $l \overset{N}{\rightleftharpoons} (l_1, l_2)$. In a similar way, $l \overset{N_1,N_2,N_3}{\rightleftharpoons} (l_1, l_2, l_3)$ denotes the index conversion rule between the 1-D and three-dimensional (3-D) index representations $l \in [\![N_1 N_2 N_3]\!]$ and $(l_1, l_2, l_3) \in [\![N_1]\!] \times [\![N_2]\!] \times [\![N_3]\!]$, respectively, meaning that $l \overset{N_1,N_2N_3}{\rightleftharpoons} (l_1, \bar{l})$ and $\bar{l} \overset{N_2,N_3}{\rightleftharpoons} (l_2, l_3)$.
$\mathcal{S}_1 \times \mathcal{S}_2$	denotes the Cartesian product with lexicographical ordering of two sets $\mathcal{S}_1 \subset [\![N_1]\!]$ and $\mathcal{S}_2 \subset [\![N_2]\!]$.
$\overline{\mathcal{S}_1 \times \mathcal{S}_2}$	denotes the set whose elements are the 1-tuple version of the 2-tuple elements of $\mathcal{S}_1 \times \mathcal{S}_2$, *i.e.*, $\overline{\mathcal{S}_1 \times \mathcal{S}_2} := \{i \in [\![N_1 N_2]\!] : i \overset{N_1,N_2}{\rightleftharpoons} (j,k), j \in \mathcal{S}_1 \subset [\![N_1]\!],\ k \in \mathcal{S}_2 \subset [\![N_2]\!]\}$.
$\mathrm{rect}(x)$, $\mathrm{rect}(x,y)$, $\mathrm{rect}(x,y,z)$	denotes the 1-D, 2-D, and 3-D rectangular function, respectively, *e.g.*, $\mathrm{rect}(x) = 1$, if $\lvert x \rvert \leq 1/2$, and zero otherwise, which can be extended to the higher dimensions, accordingly.
$\mathcal{U}([a,b])$	denotes a uniform distribution over the interval $[a,b]$.
$\mathcal{N}(\mu, \sigma^2)$	denotes a Gaussian (or normal) distribution with mean μ and variance σ^2.

Table of contents

CHAPTER 1

INTRODUCTION

Digital images are now conquering any aspect of our lives; from a selfie taken by a cell phone to a Skype meeting, from an image on Google maps to a movie, from an X-ray image of a brain to an astronomical image of the cosmos. As humans, we use images to measure the visual world around us.

When an image of a scene is projected on the sensor-plane of a conventional camera, three main processes are necessary to capture a digital image: *(i)* spatial sampling, *i.e.*, converting a continuous image of the scene to its discrete (or pixelated) representation, *(ii)* temporal sampling, *i.e.*, integrating the amount of light received by each sensor element during regular time intervals, and *(iii)* quantization of pixel values, *i.e.*, converting the continuous values to integer scale numbers. A digital image can be thus defined by a two-dimensional (2-D) array of numbers; each number representing a light intensity. In this context, an image is directly formed on the sensor-plane.

The principles of conventional cameras have been inspired by human's visual system. Our eyes contain millions of cells, which are sensitive to either the light intensity or one of the red, green, or blue lights. Similarly, a camera can generate *(i)* a gray-scale image, which consists only of the brightness of the scene; or *(ii)* a colorful image, which is made by mixing the three (primary) Red-Green-Blue (RGB)

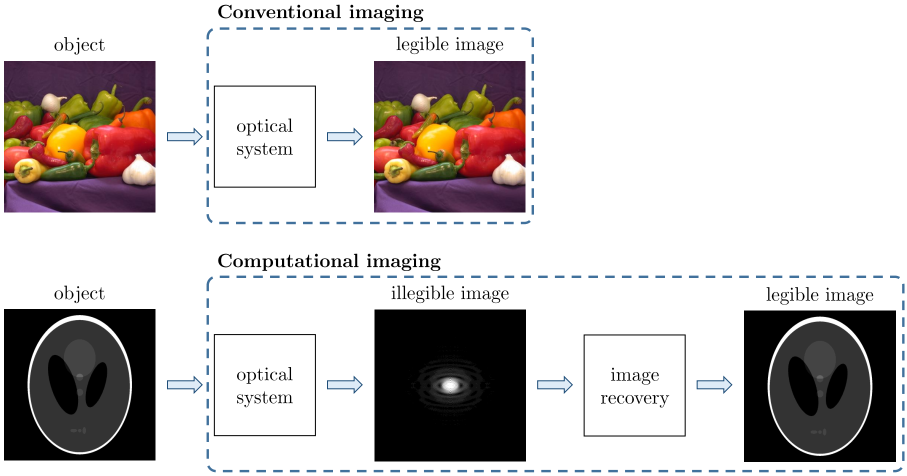

Figure 1.1: A comparison between (ideal) conventional and computational imaging systems. The illegible image for computational imaging is the 2-D Fourier transform of the object image; mimicking a measurement process in computed tomography.

colors. As an advantage of this similarity, a generated image is directly interpretable by humans; but it comes with its limitations, *i.e.*, not all the phenomena (or details) can be captured by a conventional camera. An X-ray image of a body, a height map of a terrain, and a fluorescence image of a biological cell are of few examples where the actual image is not formed on the sensor-plane. In particular, generating an X-ray image (as in computed tomography) involves recording many X-ray observations from different angles of an object. To produce a height map, a synthetic-aperture radar records the echoes of the successive transmitted radio waves. To obtain a fluorescence image, a spectrometer measures the intensity of the light illuminated by a dyed specimen over a wide range of wavelengths. A conventional camera is not able to sense an X-ray, a radio wave, nor all the colors in the range of visible light.

The three mentioned examples owe their existence to *computational imaging*. Unlike a conventional camera relying only on an optical system, computational imaging deploys a cleverly designed optical system to

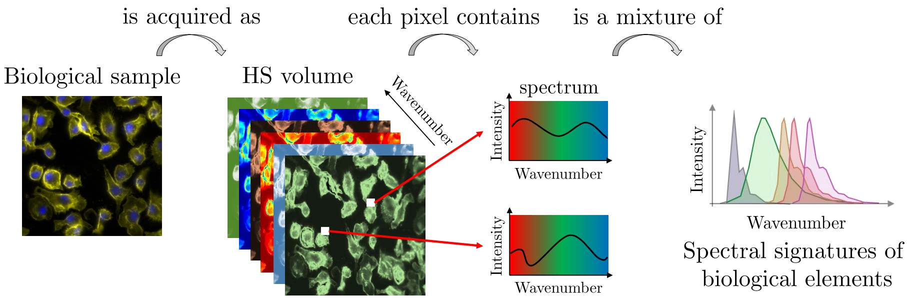

Figure 1.2: An overview of a biological HS volume. The image on the left is taken from the Broad Bioimage Benchmark Collection [68].

record an illegible phenomena and an efficient computational method or algorithm to recover a legible image (see Fig. 1.1 for an illustration).

Following the discussion above, one would agree that computational imaging has revolutionized the design of the digital imaging systems. As evident in Fig. 1.1, what a detector records may not be perceptible by our visual system; while in many cases, a meaningful image can be computationally extracted from the recorded data. Two main advantages of computational imaging are the followings: *(i)* it allows for capturing new types of information, as in the three examples above; and *(ii)* it yields an increase in the imaging performance, *e.g.*, in terms of greater robustness to noise and higher optical efficiency.

In this thesis, among a plethora of computational imaging techniques we are interested in HyperSpectral (HS) imaging modalities. An HS volume constitutes a three-dimensional (3-D) data cube associated with the collection of 2-D single-color images of the scene: it stacks one single-color (or monochromatic) image for each wavenumber. Unlike a conventional RGB image, which captures the information of a scene only in three wavenumber corresponding to RGB colors, an HS volume (fully) opens the door to another dimension, *i.e.*, spectrum. Each spatial location in an HS volume contains the recorded light intensity along a large number of spectral bands, which is essentially a mixture of the spectral signatures of the elements in the observed scene. Therefore, one may agree that if an image is worth a thousand words, an HS volume is worth a book. Fig. 1.2 depicts an example of an HS volume in biomedical application. Since every chemical elements in the universe has a unique spectral fingerprint, HS imaging allows for characterizing the constituents of a scene, their concentrations, population, and growth; and thus, it has attracted a lot of researches from agriculture, chemistry, astronomy, industry, and biology.

An HS volume contains vast amount of data; resulting in new challenges for data acquisition, storage, processing, and transmission. Imagine a typical monochromatic image of size one megabyte; the size of an HS volume stacking 1000 of those images will thus be one gigabyte. Common airborne or satellite applications require to transmit several gigabytes of data to a ground station. In biological applications, *i.e.*, the main motivation of this thesis, the acquisition time of an HS volume is translated into the light exposure. Even with very fast HS imagers, the over-exposed biological elements suffer from a photo-chemical alteration, called *photo-bleaching*: the intensity of light received from a specimen fades during the experiment.

Among different HS imaging techniques, the Fourier Transform Interferometry (FTI) [13] has shown promising achievements in the acquisition of HS volumes having high spectral resolution. Compared to the other methods, where the acquired data is an actual HS volume, the FTI captures (interferometry) indirect measurement of the HS volume, as it operates on the principles of the Michelson Interferometry (MI) [77]. Similar to the computational imaging example in Fig. 1.1, those indirect measurements need to be post-processed to obtain the HS volume. A great advantage of the FTI is that of reaching high spectral resolution by simply prolonging the acquisition process. However, the true resolution of the FTI in biomedical applications is limited by photo-bleaching, as longer acquisition implies higher light exposure. One may resolve this issue by decreasing the level of light intensity for long acquisitions; resulting in a low Signal-to-Noise Ratio (SNR).

A "computational interferometry" approach solving the problem above amounts to coding the illumination of the light source before reaching the biological specimen. Let us consider the simplest possible coding, *i.e.*, turning on and off the light source during the FTI acquisition. While this approach results in an incomplete set of FTI measurements which in turn degrades the quality of recovered HS volume, it still allows the recovery of high spectral-resolution HS volume. A more complicated coding, *i.e.*, spatiotemporal light coding, would also hold in this argument. One can, however, question *(i)* how to design an efficient light coding, as there are infinite number of ways to do it; *(ii)* how far we can reduce the light exposure received by a specimen; *(iii)* how to recover the HS volume without significant quality degradation; and *(iv)* how to mathematically support the answers to the previous three questions.

The theory of Compressive Sensing (CS) introduced by Donoho [32], Candès, and Tao [24] answers these questions. In a nutshell, CS wisely designs a sensing process to compress the available information about an unknown signal. It differs from signal compression techniques, *e.g.*, PNG and JPEG schemes for images, where the compression is done on the (known) acquired signal. Essentially, CS shows that, by a careful design of the sensing process, a low-complexity signal can be reconstructed from a few observations. A typical model for low-complexity signals is *sparsity*. A signal is called sparse if the majority of its entries are zero.

In the context of HS imaging with FTI, designing the sensing process is equivalent to coding or structuring the light illumination. Moreover, HS volumes are typically with low-complexity, although they contain a huge amount of data. Let us again consider the application of HS imaging in biomedical imaging. We notice that: *(i)* only few biological elements are present in a specimen: the spectrum at every location of the biological HS volume is the mixture of the spectral signatures of those few elements; *(ii)* the spectral signatures of the biological elements are inherently smooth: they have a sparse representation, *e.g.*, in the Fourier or a wavelet domain; *(iii)* each monochromatic image in the HS volume is actually a natural image, whose representation in a transform domain, *e.g.*, wavelet, is (approximately) sparse.

While conventional digital cameras are unable to capture beyond the human visual system, such as the HS volume of a biological specimen, computational imaging mitigates this limitation by wisely integrating the data collection process with a computational decoding step. In this thesis, among different computational techniques, we focus on the FTI modality, which records interferometric data of the observed object. When intended to yield high resolution HS volumes in biomedical applications, the efficiency of the FTI is constrained by over-exposed biological elements. Leveraging the theory of CS, the illumination of the light source in the FTI system can be purposely coded or structured, while preserving the quality and the resolution of the recovered HS volume.

1.1 Contributions

This self-contained dissertation consists of novel computational imaging strategies for acquisition of HS volumes based on interferometry method. General methodology of this thesis involves the definition of the problems, proposing a solution, developments of the theory to support the solution followed by relative numerical simulations.

In particular, we outline the main contributions of this dissertation as follows.

In Chapter 2, we present an overview of CS with orthonormal sensing and sparsity bases. In this context, our aim is to *(i)* highlight the limitations of a traditional sampling strategy, called Uniform Density Sampling (UDS), *(ii)* explain how to circumvent those limitations using a Non-uniform Density Sampling (NDS) strategy, *e.g.*, Variable Density Sampling (VDS) or Multilevel Density Sampling (MDS) introduced in the seminal works [7, 61], respectively, and *(iii)* explain how those sampling strategies control the stability and robustness of the signal recovery. We introduce the notions of global, local, and multilevel coherence and study their role on the efficiency of UDS, VDS, and MDS strategies, respectively. Furthermore, we explain the notions of uniform and non-uniform recovery guarantee that distinguishes VDS from MDS.

In VDS scheme, we observe that the reconstruction quality of a signal critically depends on the accuracy of the bound on the weighted noise power corrupting the compressed measurements. The (random) weights are corresponding to the sampling probability mass function. As a consequence, common noise power estimators, *e.g.*, a χ^2-bound for Gaussian noise, are not useful. Our novelty in this chapter is to derive a new estimator for weighted noise power that only depends on *(i)* the unweighted noise power and *(ii)* the maximum magnitude of the unweighted noise. This bound is of practical interests where the statistical parameters, *i.e.*, mean and variance, of the unweighted noise can be estimated during the calibration. Moreover, our bound is instantiated to the case of an additive Gaussian noise. The role of this contribution is essential in the analysis of the computational interferometry systems in Chapter 4 and Chapter 5.

In Chapter 3, we instantiate the theory of CS introduced in Chapter 2 to the Hadamard-Haar systems, where the sensing basis is set to the 1-D (or 2-D) Hadamard basis and the sparsity basis is set to the 1-D (respectively, 2-D) discrete Haar wavelet basis. This problem is of vested interest in a wide range of imaging applications, *e.g.*, optical multiplexing or single-pixel camera. An example of these applications is a hyperspectral imaging modality based on the single-pixel FTI, *i.e.*, covered in Chapter 5.

Hadamard-Haar system is one of the examples where UDS strategy fails to provide a satisfactory signal recovery. In this chapter, we resort to the VDS and MDS sampling strategies supported in Chapter 2 to develop successful compressive Hadamard-Haar systems for the acquisition of 1-D and 2-D signals and in turn for the stable and robust signal recovery. Our novelty in this chapter are the followings:

- We provide uniform and non-uniform recovery guarantees associated with VDS and MDS, respectively, for compressive Hadamard-Haar systems that are stable with respect to the non-sparse signals and robust to the measurement noise.
- Our results cover the recovery of both 1-D and 2-D signals. In the latter case, we treat two constructions of the 2-D discrete Haar wavelet basis, *i.e.*, using either the tensor product or the multi-resolution analysis of the 1-D Haar basis. We prove that either construction results in a different efficient sampling strategy.
- By computing the exact values of the local and multilevel coherence for Hadamard-Haar systems, we provide tight sample-complexity bounds, *i.e.*, sufficient number of measurements for stable and robust signal recovery, relatively to the VDS and MDS frameworks.

In Chapter 4, we introduce the concept of the Michelson interferometry and discuss about how it provokes the acquisition of HS volumes with the FTI. We explain the principles of the FTI, its continuous and discrete acquisition models. We discuss about the advantages and limitations of the conventional FTI systems in high-resolution application, *e.g.*, biomedical spectroscopy. We also address different distortion sources in the FTI acquisition.

The main objective of this chapter is to establish CS schemes for the conventional (or Nyquist rate) FTI, *i.e.*, reducing the total number of measurements compared to the conventional FTI and still allowing reliable and robust estimation of the HS volumes. We show that the proposed compressive FTI models enable HS imaging when total light exposure is a critical parameter, such as in biological microscopy where special biological dyes inserted in a specimen suffer from over-exposure. In a more biologically friendly scenario, we study this context when the light exposure is considered as a constrained resource to be distributed equally over all measurements.

More precisely, this chapter is driven by the following questions: *(i)* For a fixed spectral resolution, to which extent can we reduce the amount of light exposure on the observed specimen and still allow stable and robust reconstruction of the HS data? *(ii)* For a constrained-exposure budget and spectral resolution, what is the best spatiotemporal illumination allocation? These questions are answered through these two main novelties in this chapter.

- *Coded and structured illumination:* We propose two novel compressive FTI frameworks, based on coded and structured illumination techniques. The first system, referred to as Coded Illumination-FTI (CI-FTI), globally codes the light source, *i.e.*, whether the full biological specimen is globally highlighted or not per time slot, while in the second system, referred to as Structured Illumination FTI (SI-FTI), we allow the illumination of each spatial location of the specimen to be independently coded over time. The reconstruction of HS volumes resorts to the theory of CS introduced in Chapter 2. In particular, by using a VDS scheme we derive uniform recovery guarantee ensuring stable and robust HS volume reconstruction for both CI-FTI and SI-FTI models. In the sequel of this chapter, we provide an improved coding strategy for CI-FTI in the context of fluorescence spectroscopy by using an MDS scheme covered in Chapter 2.

- *Biologically-friendly constrained light exposure coding:* In fluorescence spectroscopy, the tolerance of the fluorescent dyes for light exposure can be fixed by the biologists, *e.g.*, according to the specification of fluorescent dyes. In this case, we consider that the total light exposure is a fixed resource that must be consumed exactly over the compressive FTI measurements while ensuring good reconstruction performance. We propose illumination strategies, for both CI-FTI and SI-FTI, that are faithful to this context. Those strategies adapt the intensity of the light source according to the sample-complexity of CI- or SI-FTI. Interestingly, in this scenario, we observe that full (Nyquist rate) sampling scheme is outperformed by the compressive scheme.

In Chapter 5, we introduce an alteration of the FTI system constrained to operate with single-pixel sensor, *i.e.*, called Single Pixel FTI (SP-FTI). We explain operational principles of SP-FTI, its continuous and discretized acquisition model, and how that sensing model is connected to Hadamard measurements. Since this modality is new in the literature, its profound analysis as well as an efficient design of its spatiospectral illumination coding is lacking.

Our aim in this chapter is to establish a CS scheme for SP-FTI that significantly reduces the number of measurements compared to the conventional (or Nyquist rate) SP-FTI. We explain how this goal is connected to the analysis of the Hadamard-Haar systems studied in Chapter 3, how those analysis can be combined with the VDS of interferometry domain exploited for the FTI system in Chapter 4, and how that combination allows for a stable and robust compressive SP-FTI system. Similar to our contribution in Chapter 4, we show both theoretically and numerically that in the context of biological spectroscopy, the proposed compressive SP-FTI successfully reduces the amount of light exposure received by an observed specimen.

In Chapter 6, we summarize the main contributions of this thesis and provide some lines of research for future works.

To ease the understanding of the structure of this thesis, the dependence between the chapters is depicted in Fig. 1.3. As evident, Chapter 2 is the theoretical underpinning of this thesis. Notice that we do not intend to cover the whole realm of CS in this thesis. Fig. 1.4 illustrates

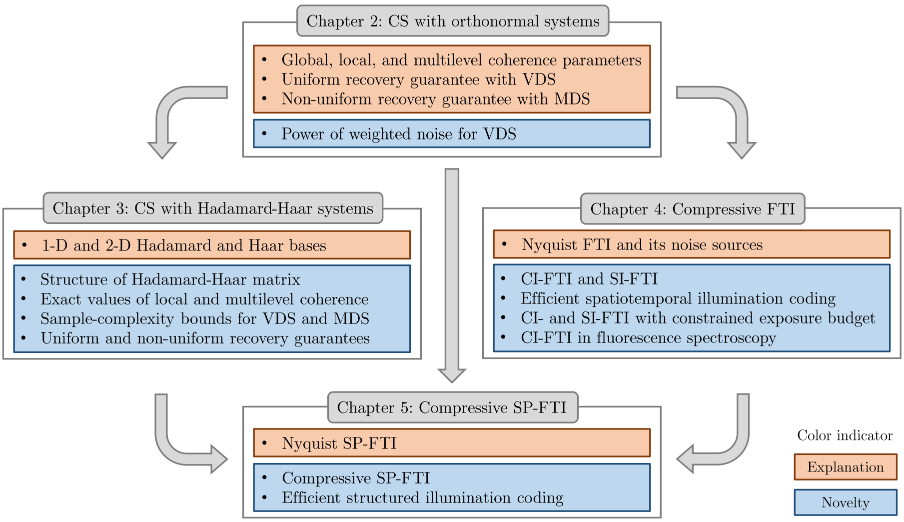

Figure 1.3: Contribution graph of the thesis. Arrows indicate the prerequisite relations among different chapters

the six main aspects of CS theory and highlights the elements exploited in thesis.

We believe that this thesis will be useful for researchers in the domains of signal and image processing, inverse problems, compressive sensing, and computational imaging as well as HS imaging system-developers. As this thesis provides first attempts to propose compressive interferometry systems for HS imaging, further studies must be done to verify the feasibility of such systems in actual optical setup.

Publications: This thesis embeds these personal publications [58, 78–80, 82–86]. However, for the sake of consistency, this former contribution [87], is not included in this thesis. In [87], we studied a reconstruction problem of low-complexity signals when they are observed through quantized compressive observations. Among such low-complexity signals, we considered 1-D sparse vectors, low-rank matrices, or compressible signals that are well approximated by one of these two models. In this context, we proved the estimation efficiency of a recovery problem, called consistent basis pursuit, enforcing the consistency between the

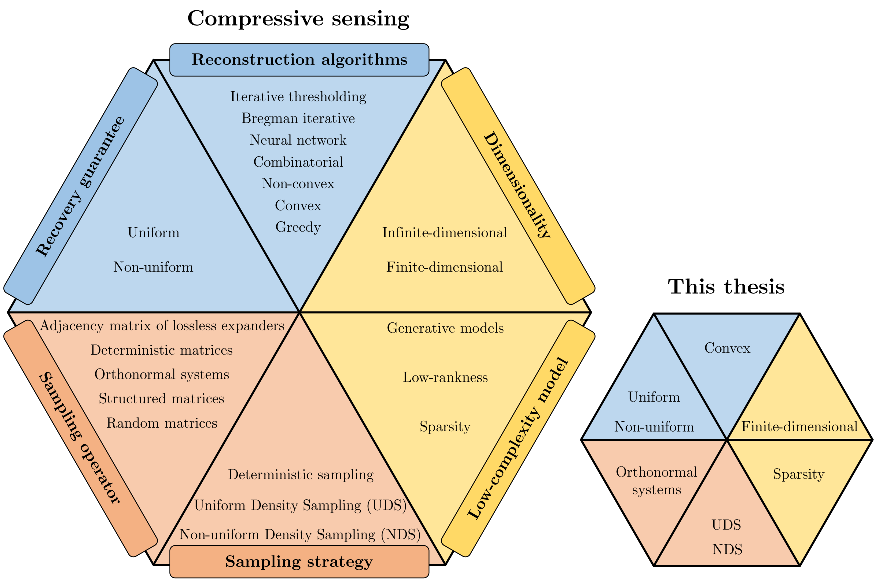

Figure 1.4: Illustration of six aspects of CS theory (left) and the elements used in this thesis form each of those aspects (right).

observations and the re-observed estimate, while promoting its low-complexity nature. We analyzed the dependence of the reconstruction error on the number of the quantized measurements.

1.2 List of Publications

Journal papers

1. **A. Moshtaghpour**, J. Bioucas Dias, and L. Jacques, "Close encounters of the binary kind: Signal reconstruction guarantees for compressive Hadamard sampling with Haar wavelet basis," *arXiv preprint arXive:1907.09795*, submitted to IEEE Transactions on Information Theory, 2019.

2. **A. Moshtaghpour**, L. Jacques, V. Cambareri, P. Antoine, and M. Roblin, "A variable density sampling scheme for compressive Fourier transform interferometry," *SIAM Journal on Imaging Sciences*, vol. 12, no. 2, pp. 671-715, 2019.

3. **A. Moshtaghpour**, L. Jacques, V. Cambareri, K. Degraux, and C. De Vleeschouwer, "Consistent basis pursuit for signal and matrix estimates in quantized compressed sensing," *IEEE Signal Processing Letters*, vol. 23, no. 1, pp. 25-29, 2016.

Conference papers

1. **A. Moshtaghpour**, J. Bioucas Dias, and L. Jacques, "Compressive single-pixel Fourier transform imaging using structured illumination," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2019, pp. 7810-7814.

2. **A. Moshtaghpour**, J. Bioucas Dias, and L. Jacques, "Single pixel hyperspectral imaging using Fourier transform interferometry," in *the International Biomedical and Astronomical Signal Processing (BASP) Frontiers workshop*, 2019, pp. 78.

3. L. Jacques and **A. Moshtaghpour**, "Structured illumination and variable density sampling for compressive Fourier transform interferometry," in *the International Biomedical and Astronomical Signal Processing (BASP) Frontiers workshop*, 2019, pp. 54.

4. **A. Moshtaghpour**, J. Bioucas Dias, and L. Jacques, "Compressive hyperspectral imaging: Fourier transform interferometry meets single pixel camera," in *the 4th International Traveling Workshop on Interactions between Sparse models and Technology (iTWIST)*, 2018.

5. **A. Moshtaghpour**, J. Bioucas Dias, and L. Jacques, "Compressive hyperspectral imaging: Fourier transform interferometry meets single pixel camera," in *the 4th International Traveling Workshop on Interactions between Sparse models and Technology (iTWIST)*, 2018.

6. **A. Moshtaghpour** and L. Jacques, "Multilevel illumination coding for Fourier transform interferometry in fluorescence spectroscopy," in *the 25th IEEE International Conference on Image Processing (ICIP)*, 2018, pp. 1433-1437.

7. **A. Moshtaghpour**, V. Cambareri, L. Jacques, P. Antoine, and M. Roblin, "Compressive hyperspectral imaging using coded Fourier transform interferometry," in *Signal Processing with Adaptive Sparse Structured Representations workshop (SPARS)*, 2017.

8. V. Cambareri, **A. Moshtaghpour**, and L. Jacques, "A greedy blind calibration method for compressed sensing with unknown sensor gains," in *IEEE International Symposium on Information Theory (ISIT)*, 2017, pp. 1132-1136.

9. **A. Moshtaghpour**, V. Cambareri, K. Degraux, A. C. Gonzalez Gonzalez, M. Roblin, L. Jacques, and P. Antoine, "Coded-illumination Fourier transform interferometry," in *the Golden Jubilee Meeting of the Royal Belgian Society for Microscopy (RBSM)*, 2016, pp. 65-66.

10. **A. Moshtaghpour**, K. Degraux, V. Cambareri, A. Gonzalez, M. Roblin, L. Jacques, and P. Antoine, "Compressive hyperspectral imaging with Fourier transform interferometry," in *the 3rd International Traveling Workshop on Interactions between Sparse models and Technology (iTWIST)*, 2016, pp. 27-29.

CHAPTER 2

SAMPLING STRATEGIES FOR COMPRESSIVE SENSING WITH ORTHONORMAL SYSTEMS

This Chapter provides the main tools from CS theory and its recent extension to Non-uniform Density Sampling (NDS) of an orthogonal sensing basis (*e.g.*, Fourier and Hadamard). As depicted in Fig. 1.3, the content of this chapter constitutes the backbone of the thesis. We here emphasize on those aspects of CS illustrated in Fig. 1.4, *i.e.*, sparsity of finite-dimensional signals in orthonormal bases as low-complexity model, uniform and non-uniform density sampling of orthonormal sensing bases, and uniform and non-uniform recovery guarantees associated with a convex reconstruction algorithm. After a short introduction to the theory of CS, we explain the principles and limitations of a traditional sampling strategy, *i.e.*, Uniform Density Sampling (UDS). We then present two sampling strategies, *i.e.*, the Variable Density Sampling (VDS) scheme of Krahmer and Ward [61] and the Multilevel Density Sampling (MDS) approach of Adcock *et al.* [7], allowing us to mitigate the limitations of UDS. Finally, as our novelty in this chapter, we address an aspect of VDS seemingly missing in the literature.

The content of this chapter has been published in [80, 84].

2.1 An Overview of Compressive Sensing

The theory of CS, introduced by Donoho [32] and Candès and Tao [24], is now a versatile sampling paradigm in many real-world applications, *e.g.*, Magnetic Resonance Imaging (MRI) [72], fluorescence microscopy [99, 106], and single-pixel imaging [34]. Mathematically, CS considers the problem of recovering a signal $\boldsymbol{x} \in \mathbb{C}^N$ from M noisy measurements

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{n} \in \mathbb{C}^M. \tag{2.1}$$

In (2.1), the matrix $\boldsymbol{A} \in \mathbb{C}^{M\times N}$ approximates the physical sensing process of $\boldsymbol{x}$, and $\boldsymbol{n}$ denotes an additive noise vector. A typical goal in CS is to minimize the number of measurements M while guaranteeing the quality of the signal recovery. This is indeed a critical aspect of the applications of CS. For example, the number of measurements in computed tomography application can be translated into the X-ray dose, which has to be minimized.

2.2 Uniform Density Sampling: Principles and Limitations

When restricted to signal sensing with random orthonormal basis ensembles [24, 32] (*e.g.*, random Fourier or Hadamard ensembles), CS theory targets the recovery of a *low-complexity* signal $\boldsymbol{x} \in \mathbb{C}^N$ from a vector of noisy measurements $\boldsymbol{y} = \boldsymbol{P}_\Omega \boldsymbol{\Phi}^* \boldsymbol{x} + \boldsymbol{n}$, where $\boldsymbol{\Phi} \in \mathbb{C}^{N\times N}$ is an orthonormal sensing system, $\Omega \subset [\![N]\!]$ is a subset of indices chosen at random with $|\Omega| = M \ll N$, and $\boldsymbol{n}$ accounts for some additive observation noise. The low-complexity nature of $\boldsymbol{x}$ generally amounts to assuming it K-sparse, *i.e.*, with $\|\boldsymbol{x}\|_0 := |\mathrm{supp}(\boldsymbol{x})| \leq K$, and we write $\boldsymbol{x} \in \Sigma_K$, or at least well approximated by a sparse signal, *i.e.*, compressible.

In order to estimate $\boldsymbol{x}$ from $\boldsymbol{y}$, the Restricted Isometry Property (RIP) [23] defined below is a sufficient condition on the matrix $\boldsymbol{P}_\Omega \boldsymbol{\Phi}^*$

for most classes of algorithms, *e.g.*, based on convex optimization [23], greedy [130], and thresholding [16] strategies.

Definition 2.1. *Given $K \leq N$, the Restricted Isometry Constant (RIC) δ_K associated with a matrix $\boldsymbol{A} \in \mathbb{C}^{M\times N}$ is the smallest number δ for which*

$$(1-\delta)\|\boldsymbol{x}\|^2 \leq \|\boldsymbol{A}\boldsymbol{x}\|^2 \leq (1+\delta)\|\boldsymbol{x}\|^2 \tag{2.2}$$

holds for all K-sparse vectors $\boldsymbol{x} \in \mathbb{C}^N$. Alternatively, if (2.2) *holds for $\delta = \delta_K$, we say that $\boldsymbol{A}$ satisfies the RIP of order K and constant δ_K.*

When the RIP holds and the observation noise is bounded, *i.e.*, $\|\boldsymbol{n}\| \leq \varepsilon$ for some $\varepsilon > 0$, the Basis Pursuit DeNoise (BPDN) program expressed as

$$\hat{\boldsymbol{x}} \;=\; \operatorname*{arg\,min}_{\boldsymbol{u}\in\mathbb{C}^N} \|\boldsymbol{u}\|_1 \ \text{s.t.}\ \|\boldsymbol{y}-\boldsymbol{A}\boldsymbol{u}\| \leq \varepsilon, \tag{2.3}$$

provides an accurate estimate of the original signal [24].

Proposition 2.2 (Thm. 2.1 [19])**.** *Assume that the RIC $\delta = \delta_{2K}$ of $\boldsymbol{A} \in \mathbb{C}^{M\times N}$ satisfies*[a] *$\delta_{2K} < \frac{1}{\sqrt{2}}$. Then, for all $\boldsymbol{x} \in \mathbb{C}^N$ observed through the noisy CS model $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{n}$ with $\|\boldsymbol{n}\| \leq \varepsilon$, the solution $\hat{\boldsymbol{x}}$ of* (2.3) *satisfies*

$$\|\boldsymbol{x}-\hat{\boldsymbol{x}}\| \leq c_1 \tfrac{2\sigma_K(\boldsymbol{x})_1}{\sqrt{K}} + c_2\varepsilon, \quad c_1 = 2\tfrac{\delta+\sqrt{\delta(1/\sqrt{2}-\delta)}}{1-\sqrt{2}\delta}, c_2 = 2\tfrac{\sqrt{2(1+\delta)}}{1-\sqrt{2}\delta},$$

where $\sigma_K(\boldsymbol{x})_1 := \|\boldsymbol{x}-\mathcal{H}_K(\boldsymbol{x})\|_1$ is the best K-term approximation error (in the ℓ_1 sense), and $\mathcal{H}_K$ is the hard thresholding operator that maps all but the K largest-magnitude entries of the argument to zero. In particular, the reconstruction is exact, i.e., *$\hat{\boldsymbol{x}} = \boldsymbol{x}$, if $\boldsymbol{x}$ is K-sparse and $\varepsilon = 0$. Note that for $\delta = 1/3$, $c_1 \leq 2.58$ and $c_2 \leq 6.18$.*

[a] There exist alternative versions of this condition on the RIC, *e.g.*, in [61] and [41, Theorem 5].

More generally, if the signal $\boldsymbol{x}$ has a sparse or compressible representation in a general orthonormal basis $\boldsymbol{\Psi} \in \mathbb{C}^{N\times N}$ (*e.g.*, in a wavelet basis), *i.e.*, $\boldsymbol{x} = \boldsymbol{\Psi}\boldsymbol{s}$ where the vector of coefficients $\boldsymbol{s} \in \mathbb{C}^N$ is K-sparse or compressible, one must ensure that $\boldsymbol{P}_\Omega \boldsymbol{\Phi}^* \boldsymbol{\Psi}$ respects the RIP.

Interestingly, according to [95], choosing each of the M elements of Ω uniformly at random in $[\![N]\!]$, *i.e.*, constitutes the UDS scheme, gives that, with probability exceeding $1-\epsilon$, the RIC of the matrix $\boldsymbol{A} = \sqrt{\frac{N}{M}} \boldsymbol{P}_\Omega \boldsymbol{\Phi}^* \boldsymbol{\Psi}$ satisfies $\delta_K \leq \delta$, if

$$M \gtrsim \delta^{-2} N \mu^2(\boldsymbol{\Phi}^* \boldsymbol{\Psi}) K \log(\epsilon^{-1}), \tag{2.4}$$

where $\mu(\boldsymbol{\Phi}^* \boldsymbol{\Psi})$ is the *mutual coherence*[1] of $\boldsymbol{\Phi}^* \boldsymbol{\Psi}$, *i.e.*,

$$\mu(\boldsymbol{\Phi}^* \boldsymbol{\Psi}) := \max_{l,j} |(\boldsymbol{\Phi}^* \boldsymbol{\Psi})_{l,j}| \in [1/\sqrt{N}, 1]. \tag{2.5}$$

However, the impact of this result is unfortunately limited in the case where, for instance, $\boldsymbol{\Phi}$ and $\boldsymbol{\Psi}$ are the discrete Fourier (or Hadamard) and wavelet bases, respectively. This is in fact an important special combination appearing in, *e.g.*, Magnetic Resonance Imaging [9], radio-interferometry [125] and in the compressive FTI and SP-FTI schemes considered in this thesis (see § 4.4, § 4.5, and § 5.3, respectively). The coherence in this case is $\mu(\boldsymbol{\Phi}^* \boldsymbol{\Psi}) = 1$ and (2.4) states that all the samples are required for reconstruction ($M \gtrsim N$), even when the signal is highly sparse in a wavelet basis. This drawback is often called the "coherence barrier" in the literature [8].

Fortunately, this limitation is actually induced by the way Ω is built, *i.e.*, according to a UDS of the columns of $\boldsymbol{\Phi}$. Adopting an NDS scheme allows us to break this "coherence barrier" [7, 61, 93, 99, 124]. As explained in § 3.1.1, there exist fundamentally-different theoretical results that suggest an NDS in the case of coherent bases, with the same idea of breaking the coherence barrier (*e.g.*, [7, 17, 61, 93]). In this thesis we restrict our analysis to the VDS scheme of Krahmer and Ward [61] and the MDS scheme of Adcock *et al.* [7]. Note that the term "VDS" is often used in the literature for other non-uniform density sampling strategies,

[1]In some references, *e.g.*, [99], it is defined as $\mu(\boldsymbol{\Phi}, \boldsymbol{\Psi}) := \max_{l,j} |\langle \boldsymbol{\phi}_l, \boldsymbol{\psi}_j \rangle|$.

while we here use VDS and MDS terms to distinguish the frameworks of [61] and [7, 8].

An important aspect of CS theory is the difference between *uniform* and *non-uniform*[2] (or fixed signal) recovery guarantees [42, Chapter 9]. The former is tightly connected to the RIP defined in Def. 2.1. Essentially, a uniform recovery guarantee claims that a single draw of the sampling matrix is, with high probability, sufficient for the recovery of *all* sparse signals. A non-uniform recovery guarantee asserts that a single draw of the sampling matrix is, with high probability, sufficient for recovery of a *fixed* sparse signal.

2.3 Variable Density Sampling

Krahmer and Ward showed in [61] that the sample-complexity bound in (2.4) can be modified by assigning higher sampling probability to the columns of the sensing basis $\mathbf{\Phi}$ (or equivalently, to the rows of $\mathbf{\Phi}^*$) that are highly coherent with the columns of the sparsity basis $\mathbf{\Psi}$. This idea has generalized the concept of mutual (or global) coherence (2.5) to *local coherence*, *i.e.*, the quantity

$$\mu_l^{\mathrm{loc}}(\mathbf{\Phi}^*\mathbf{\Psi}) := \max_{1\le j\le N} |(\mathbf{\Phi}^*\mathbf{\Psi})_{l,j}| \in [1/\sqrt{N}, 1],\ \forall l \in [\![N]\!], \tag{2.6}$$

with $\max_l \mu_l^{\mathrm{loc}}(\mathbf{\Phi}^*\mathbf{\Psi}) = \mu(\mathbf{\Phi}^*\mathbf{\Psi})$. We denote the local coherence vector $\boldsymbol{\mu}^{\mathrm{loc}} := [\mu_1^{\mathrm{loc}}, \cdots, \mu_N^{\mathrm{loc}}]^\top$ formed by all the local coherence values. Krahmer and Ward [61] proved that this quantity determines a sufficient condition on the construction of the subsampling set Ω in order to obtain a RIP matrix with $M \ll N$, even when the sensing and sparsity bases are too coherent, *i.e.*, $\mu(\mathbf{\Phi}^*\mathbf{\Psi}) \approx 1$. Throughout this thesis, the term "VDS" refers to the sampling strategy proposed by Krahmer and Ward [61].

[2] A word of caution. Uniform and non-uniform density sampling strategies are related to the way the rows of a matrix are subsampled and should not be confused with uniform and non-uniform recovery guarantees.

Proposition 2.3 (RIP for VDS [61, Thm. 5.2]). *Let $\mathbf{\Phi} \in \mathbb{C}^{N\times N}$ and $\mathbf{\Psi} \in \mathbb{C}^{N\times N}$ be orthonormal sensing and sparsity bases, respectively, with local coherence values $\mu_l^{\rm loc}(\mathbf{\Phi}^*\mathbf{\Psi}) \leq \kappa_l$ for some values $\kappa_l \in \mathbb{R}_+$. Let us define $\boldsymbol{\kappa} := [\kappa_1, \cdots, \kappa_N]^\top$. Suppose $K \gtrsim \log(N)$,*

$$M \gtrsim \delta^{-2}\|\boldsymbol{\kappa}\|^2 K \log(\epsilon^{-1}),$$

and choose M (possibly not distinct) indices $l \in \Omega \subset [\![N]\!]$ independent and identically distributed (iid) with respect to the probability distribution p on $[\![N]\!]$ given by

$$p(l) := \tfrac{\kappa_l^2}{\|\boldsymbol{\kappa}\|^2}.$$

Consider the diagonal matrix $\boldsymbol{D} = \operatorname{diag}(\boldsymbol{d}) \in \mathbb{R}^{M\times M}$ with $d_j = 1/\sqrt{p(\Omega_j)}$, $j \in [\![M]\!]$. Then with probability exceeding $1-\epsilon$, the RIC δ_K of the preconditioned matrix $\frac{1}{\sqrt{M}}\boldsymbol{D}\boldsymbol{P}_\Omega\mathbf{\Phi}^*\mathbf{\Psi}$ satisfies $\delta_K \leq \delta$.

We will see later that this VDS offers new means for compressive FTI and SP-FTI. Note that, in practice, although $\boldsymbol{P}_\Omega\mathbf{\Phi}^*$ models the analog or optical sensing procedure of a growing number of CS applications (*e.g.*, in compressive MRI or radio-interferometry), the conditioning by $\boldsymbol{D}$ proposed in Prop. 2.3 is rather achieved by post-processing the acquired data [61]. Therefore, our estimation of the HS data will rather follow this straightforward adaptation of Prop. 2.2, also turned to an analysis-based sparsity framework [35] that better suits the rest of our developments.

Theorem 2.4 (VDS and uniform recovery guarantee for CS, adapted from [61] and [19]). *Let $\mathbf{\Phi} \in \mathbb{C}^{N\times N}$ and $\mathbf{\Psi} \in \mathbb{C}^{N\times N}$ be orthonormal sensing and sparsity bases, respectively, with $\mu_l^{\rm loc}(\mathbf{\Phi}^*\mathbf{\Psi}) \leq \kappa_l$ for some values $\kappa_l \in \mathbb{R}_+$. Let us define $\boldsymbol{\kappa} := [\kappa_1, \cdots, \kappa_N]^\top$. Fix $\epsilon \in (0,1]$ and $\delta < 1/\sqrt{2}$ and suppose $K \gtrsim \log(N)$,*

$$M \gtrsim \delta^{-2}\|\boldsymbol{\kappa}\|^2 K \log(\epsilon^{-1}), \tag{2.7}$$

and choose M (possibly not distinct) indices $l \in \Omega \subset [\![N]\!]$ i.i.d. with respect to the probability distribution η on $[\![N]\!]$ given by

$$p(l) := \frac{\kappa_l^2}{\|\boldsymbol{\kappa}\|^2}. \tag{2.8}$$

Consider the diagonal matrix $\boldsymbol{D} = \mathrm{diag}(\boldsymbol{d}) \in \mathbb{R}^{M\times M}$ with $d_j = 1/\sqrt{\eta(\Omega_j)}$, $j \in [\![M]\!]$. With probability exceeding $1-\epsilon$, for all $\boldsymbol{x} \in \mathbb{C}^N$ observed through the noisy CS model $\boldsymbol{y} = \boldsymbol{P}_\Omega \boldsymbol{\Phi}^* \boldsymbol{x} + \boldsymbol{n}$ with $\|\boldsymbol{D}\boldsymbol{n}\| \le \varepsilon\sqrt{M}$, the solution $\hat{\boldsymbol{x}}$ of the program

$$\hat{\boldsymbol{x}} \;=\; \underset{\boldsymbol{u}\in\mathbb{C}^N}{\arg\min}\, \|\boldsymbol{\Psi}^*\boldsymbol{u}\|_1 \;\text{ s. t. }\; \frac{1}{\sqrt{M}}\|\boldsymbol{D}(\boldsymbol{y} - \boldsymbol{P}_\Omega\boldsymbol{\Phi}^*\boldsymbol{u})\| \le \varepsilon, \tag{2.9}$$

satisfies

$$\|\boldsymbol{x} - \hat{\boldsymbol{x}}\| \le c_1 \tfrac{\sigma_K(\boldsymbol{\Psi}^*\boldsymbol{x})_1}{\sqrt{K}} + c_2\varepsilon, \quad c_1 = 2\tfrac{\delta+\sqrt{\delta(1/\sqrt{2}-\delta)}}{1-\sqrt{2}\delta}, c_2 = \tfrac{2\sqrt{2(1+\delta)}}{1-\sqrt{2}\delta}, \tag{2.10}$$

where $\sigma_K(\boldsymbol{u})_1 := \|\boldsymbol{u} - \mathcal{H}_K(\boldsymbol{u})\|_1$ is the best K-term approximation error (in the ℓ_1 sense), and $\mathcal{H}_K$ is the hard thresholding operator that maps all but the K largest-magnitude entries of the argument to zero. Note that for $\delta = 1/3$, $c_1 < 2.58,\ c_2 < 6.18$. In particular, the reconstruction is exact, i.e.*, $\hat{\boldsymbol{x}} = \boldsymbol{x}$, if $\boldsymbol{x}$ is K-sparse and $\varepsilon = 0$.*

Proof. See § 2.6.1 □

This recovery guarantee is uniform in the sense that a single construction of the measurement matrix $\boldsymbol{P}_\Omega\boldsymbol{\Phi}^*$ with respect to the sample-complexity bound in (2.7) and sampling probability mass function (pmf) in (2.8) is sufficient to ensure (with high probability) the recovery of all sparse vectors. This result is of interest in those applications of CS where sparsity (or compressibility) of the target signal is the only possible prior knowledge. In the next section we describe a method, which takes into account local sparsity of the target signal.

In § 4.4.2, § 4.5.2, and § 5.3.2, we will leverage Thm. 2.4 in order to characterize the sample-complexity bounds and the stability of our compressive FTI and SP-FTI strategies.

2.4 Multilevel Density Sampling

Adcock and co-authors [7] have also advocated that the global notion of coherence (2.5) and sparsity must be replaced by proper local versions in order to obtain a better subsampling strategy.

To fix the ideas, we first introduce the CS setup proposed in [7]. For a fixed $r \in \mathbb{N}$ we decompose the signal (or sparsity) domain $[\![N]\!]$ into r disjoint *sparsity levels* $\mathcal{S} := \{\mathcal{S}_1, \ldots, \mathcal{S}_r\}$ such that $\bigcup_{l=1}^r \mathcal{S}_l = [\![N]\!]$. Given a vector of sparsity parameters $\boldsymbol{k} = [k_1, \ldots, k_r]^\top \in \mathbb{N}^r$, a vector $\boldsymbol{s} \in \mathbb{C}^N$ is called $(\mathcal{S}, \boldsymbol{k})$-sparse-in-level, and we write $\boldsymbol{s} \in \Sigma_{\mathcal{S},\boldsymbol{k}}$, if $|\operatorname{supp} \boldsymbol{P}_{\mathcal{S}_l} \boldsymbol{s}| \leq k_l$ for all $l \in [\![r]\!]$. For an arbitrary vector $\boldsymbol{s}$, its $(\mathcal{S}, \boldsymbol{k})$-approximation error is denoted by

$$\sigma_{\mathcal{S},\boldsymbol{k}}(\boldsymbol{s}) := \min\{\|\boldsymbol{s} - \boldsymbol{z}\|_1 : \boldsymbol{z} \in \Sigma_{\mathcal{S},\boldsymbol{k}}\} = \sum_l \sigma_{k_l}(\boldsymbol{P}_{\mathcal{S}_l} \boldsymbol{s})_1.$$

We quickly observe that the sparsity-in-level model reduces to the global sparsity model by setting $r = 1$ and $\mathcal{S} = [\![N]\!]$.

Similarly, we decompose the sampling domain $[\![N]\!]$ into r disjoint *sampling levels* defined as $\mathcal{W} := \{\mathcal{W}_1, \ldots, \mathcal{W}_r\}$ with $\bigcup_{l=1}^r \mathcal{W}_l = [\![N]\!]$. Given $\boldsymbol{m} = [m_1, \ldots, m_r]^\top \in \mathbb{N}^r$, the set $\Omega_{\mathcal{W},\boldsymbol{m}} := \bigcup_{t=1}^r \Omega_t$ provides an MDS scheme, or $(\mathcal{W}, \boldsymbol{m})$-MDS, if, for each $1 \leq t \leq r$, $\Omega_t \subseteq \mathcal{W}_t$, $|\Omega_t| = m_t \leq |\mathcal{W}_t|$, and if the entries of Ω_t are chosen uniformly at random (without replacement) in $\mathcal{W}_t$.

We further need to define two quantities controlling the sample-complexity bound in MDS scheme (see below). Given an orthonormal matrix $\boldsymbol{U} \in \mathbb{C}^{N \times N}$ and local sparsity values $\boldsymbol{k}$, the t^{th} *relative sparsity* is defined as

$$K_t^{\mathcal{W},\mathcal{S}}(\boldsymbol{U},\boldsymbol{k}) = \max_{\boldsymbol{z}\in\Sigma_{\mathcal{S},\boldsymbol{k}}:\,\|\boldsymbol{z}\|_\infty\le 1} \|\boldsymbol{P}_{\mathcal{W}_t}\boldsymbol{U}\boldsymbol{z}\|^2. \tag{2.11}$$

In the cases where the computation of the exact relative sparsity values given in (2.11) is not feasible, one can instead upper bound it as stated in the next lemma, which is adapted from [9, Eq. 13].

Lemma 2.5. $\sqrt{K_t^{\mathcal{W},\mathcal{S}}(\boldsymbol{U},\boldsymbol{k})} \le \sum_{l=1}^{|\mathcal{S}|} \|\boldsymbol{P}_{\mathcal{W}_t}\boldsymbol{U}\boldsymbol{P}_{\mathcal{S}_l}^\top\|_{2,2}\sqrt{k_l}$.

Moreover, the $(t,l)^{\text{th}}$ *multilevel coherence* of $\boldsymbol{U}$ with respect to the sampling and sparsity levels $\mathcal{W}$ and $\mathcal{S}$, respectively, is defined as

$$\mu_{t,l}^{\mathcal{W},\mathcal{S}}(\boldsymbol{U}) := \mu(\boldsymbol{P}_{\mathcal{W}_t}\boldsymbol{U})\,\mu(\boldsymbol{P}_{\mathcal{W}_t}\boldsymbol{U}\boldsymbol{P}_{\mathcal{S}_l}^\top). \tag{2.12}$$

Within this context, the following guarantee can be reformulated from [7, Thm. 4.4].

Theorem 2.6 (MDS and non-uniform recovery guarantee for CS, adapted from [7])**.** *Let* $\boldsymbol{\Phi} \in \mathbb{C}^{N\times N}$ *and* $\boldsymbol{\Psi} \in \mathbb{C}^{N\times N}$ *be orthonormal sensing and sparsity bases, respectively. Fix sampling and sparsity levels* $\mathcal{W}$ *and* $\mathcal{S}$*, respectively. Let* $\Omega = \Omega_{\mathcal{W},\boldsymbol{m}}$ *be a* $(\mathcal{W},\boldsymbol{m})$*-MDS and* $(\mathcal{S},\boldsymbol{k})$ *be any pair such that the following holds: for* $0 < \epsilon \le \exp(-1)$*,* $K = \|\boldsymbol{k}\|_1$*,* $\hat{m}_t$ *is such that for all* $l \in [\![|\mathcal{S}|]\!]$*,*

$$1 \gtrsim \sum_{t=1}^{r} \left(\left(\frac{|\mathcal{W}_t|}{\hat{m}_t} - 1\right) \mu_{t,l}^{\mathcal{W},\mathcal{S}}(\boldsymbol{\Phi}^*\boldsymbol{\Psi})\, K_t^{\mathcal{W},\mathcal{S}}(\boldsymbol{\Phi}^*\boldsymbol{\Psi},\boldsymbol{k})\right), \tag{2.13}$$

and m_t *such that for all* $t \in [\![|\mathcal{W}|]\!]$*,*

$$m_t \gtrsim \hat{m}_t \log(K\epsilon^{-1})\log(N), \tag{2.14}$$

$$m_t \gtrsim |\mathcal{W}_t| \Big(\sum_{l=1}^{r} \mu_{t,l}^{\mathcal{W},\mathcal{S}}(\boldsymbol{\Phi}^*\boldsymbol{\Psi})\, k_l\Big) \log(K\epsilon^{-1})\log(N). \tag{2.15}$$

Given the noisy CS measurements $\boldsymbol{y} = \boldsymbol{P}_\Omega \boldsymbol{\Phi}^* \boldsymbol{x} + \boldsymbol{n}$ *with* $\|\boldsymbol{n}\| \le \varepsilon$*, suppose that* $\hat{\boldsymbol{x}} \in \mathbb{C}^N$ *is a minimizer of*

$$\hat{\boldsymbol{x}} = \arg\min_{\boldsymbol{u} \in \mathbb{C}^N} \|\boldsymbol{\Psi}^* \boldsymbol{u}\|_1 \ \text{ s.t. } \ \|\boldsymbol{y} - \boldsymbol{P}_\Omega \boldsymbol{\Phi}^* \boldsymbol{u}\| \le \varepsilon. \tag{2.16}$$

Then, with probability exceeding $1 - \epsilon$*, we have*

$$\|\boldsymbol{x} - \hat{\boldsymbol{x}}\| \le c_1\, \sigma_{\mathcal{S},\boldsymbol{k}}(\boldsymbol{\Psi}^* \boldsymbol{x}) + c_2 (1 + C\sqrt{K})\, \varepsilon \sqrt{q}, \tag{2.17}$$

where $q := \max_t \frac{|\mathcal{W}_t|}{m_t}$ *for some constant* $0 < c_1 \le 22, 0 < c_2 \le 11$*, and where* $0 \le C \le 3\sqrt{6} + \frac{4\sqrt{6}\sqrt{\log(6N\epsilon^{-1})}}{\log(N)}$.

We remark the main differences between this theorem and Thm. 2.4. First, Thm. 2.6 provides a non-uniform recovery guarantee: the measurement matrix $\boldsymbol{P}_\Omega \boldsymbol{\Phi}^*$ satisfying the conditions of the proposition needs (in theory) to be redrawn when a new vector is to be recovered. Second, the MDS scheme requires a prior information about the local sparsity of the target signal; while the VDS scheme in Thm. 2.4 does not need such information. Third, the parameter ε in optimization program (2.16) is a bound on the observation noise power, while in the VDS scheme (2.9) it is a bound on the weighted noise power. However, we show in the next section that one can determine a bound on $\|\boldsymbol{D}\boldsymbol{n}\|$, which holds with controllable probability, and that depends on the $\|\boldsymbol{n}\|$, $\|\boldsymbol{n}\|_\infty$ (or on estimations bounding these quantities with high probability) and a parameter fixed by the pmf defining the VDS scheme.

2.5 Noise Level Estimation in VDS

We now study an important aspect of the VDS scheme that seems not covered in the literature: estimating the noise level ε in (2.9) by integrating the influence of the weighting (random) matrix $\boldsymbol{D}$ introduced in Prop. 2.3. This estimation is indeed critical to achieve robust reconstruction of HS volumes from the two compressive FTI schemes studied in

§ 4.4 and § 4.5 and from the compressive SP-FTI covered in § 5.3. The content of this section is voluntarily written in general notations, as it could also be of general interest for any VDS-based compressive sensing application corrupted by some additive measurement noise.

The next theorem bounds with controlled probability the (weighted) ℓ_2-norm of any vector $\boldsymbol{n} \in \mathbb{C}^M$, *e.g.*, the fixed realization of a noise vector corrupting the sensing model $\boldsymbol{y} = \boldsymbol{P}_\Omega \boldsymbol{\Phi}^* \boldsymbol{x} + \boldsymbol{n}$, when this norm is weighted by a random diagonal matrix $\boldsymbol{D}$ associated with Ω. The bound only depends on the unweighted ℓ_2- and ℓ_∞-norms of $\boldsymbol{n}$, and on a quantity $\rho > 1$ fixed by the density defining the VDS.

Theorem 2.7. *Given two integers $M < N$, let us consider a discrete random variable (r.v.) $\beta \in [\![N]\!]$ associated with the pmf $\eta(l) := \mathbb{P}[\beta = l]$ for $l \in [\![N]\!]$, and assume there exists a $\rho \in [1, \infty)$ such that*

$$\tfrac{1}{N} \sup_{q\geq 1} \tfrac{1}{q} (\mathbb{E}_\beta\, \eta(\beta)^{-q})^{1/q} \leq \rho. \tag{2.18}$$

We define a random index set $\Omega = \{\omega_j\}_{j=1}^M \subset [\![N]\!]$ made of M (possibly non-distinct) indices $\omega_j \sim_{\rm iid} \beta$, and a random diagonal matrix $\boldsymbol{D} \in \mathbb{R}^{M\times M}$ such that $D_{jj} = 1/\sqrt{\eta(\omega_j)}$. Given $s > 0$ and $\boldsymbol{n} \in \mathbb{C}^M$, we have

$$\tfrac{1}{M}\, \|\boldsymbol{D}\boldsymbol{n}\|^2 \leq \tfrac{N}{M} \|\boldsymbol{n}\|^2 + 4e \max\big(\tfrac{s}{M}, \tfrac{\sqrt{s}}{\sqrt{M}}\big) \rho\, \|\boldsymbol{n}\|_\infty^2 N, \tag{2.19}$$

*with probability exceeding $1 - e^{-s/2}$ (*e.g.*, for $s = 6$, this probability exceeds 0.95).*

For the proof see § 2.6.2. The fact that $\rho \geq 1$ is a simple consequence of $\mathbb{E}(\eta(\beta))^{-1} = \sum_l 1 = N$.

Remark 2.8. *Observe that, up to the normalization in $1/N$, the parameter ρ in* (2.18) *is actually a bound on the sub-exponential norm $\|1/\eta(\beta)\|_{\psi_1} := \sup_{q\geq 1} \frac{1}{q} (\mathbb{E}_\beta\, \eta(\beta)^{-q})^{1/q}$ of the r.v. $1/\eta(\beta)$ [120, 121]. From this definition, and since $\|X\|_{\psi_1} \leq L$ for any bounded r.v. X with $|X| \leq L$ and $L >$*

0, we can directly identify that in the UDS case, $\eta(l) = 1/N$ *implies* $\|1/\eta(\beta)\|_{\psi_1} \leq N$, i.e., *we can set* $\rho = 1$.

More generally, for any VDS scheme, e.g., see Prop. 4.10, where $\eta(l) = c_\eta \min(1, |l-l_0|^{-\alpha})$ *for some exponent* $\alpha > 0$*, an offset parameter* $l_0 \in [\![N]\!]$ *centering* η *on* l_0*, and* $c_\eta > 0$*, we find*

$$\|1/\eta(\beta)\|_{\psi_1} = c_\eta^{-1} \| \max\big(1, |\beta - l_0|^\alpha\big) \|_{\psi_1} \leq c_\eta^{-1}(N - l_0)^\alpha,$$

which provides $\rho = c_\eta^{-1}|N - l_0|^\alpha/N$*. In Chapters 4 and 5, the parameter* l_0 *denotes the index of the OPD origin. In particular, for* $\alpha = 1$*,* ρ *mainly depends on the pmf normalization constant* c_η, i.e., $c_\eta^{-1} \simeq \log N$, *as computed in Prop. 4.10.*

In conclusion, for $\alpha = 0$ *(UDS) and for* $\alpha = 1$*, we can expect that* ρ *is either constant or that it grows slowly (logarithmically) when* N *increases, hence ensuring a good control over the bound* (2.19).

The previous theorem is expected to provide useful bounds if the leading terms in the right-hand side of (2.19) is $\frac{N}{M}\|\boldsymbol{n}\|^2$, *i.e.*, if $\rho\|\boldsymbol{n}\|_\infty^2 N \lesssim \frac{N}{M}\|\boldsymbol{n}\|^2$. This happens for ρ slowly growing when N increases, *e.g.*, $\rho = O(\log N)$, (see Remark 2.8) and for any vector $\boldsymbol{n}$ that is not too sparse, *i.e.*, such that $\|\boldsymbol{n}\|_\infty^2 \lesssim \|\boldsymbol{n}\|^2/M$. As shown in the next corollary, a Gaussian random noise respects this requirement with high probability[3].

Corollary 2.9. *In the context of Thm. 2.7, let us consider the Gaussian random vector* $\boldsymbol{n} \in \mathbb{R}^M$ *with* $n_k \sim_{\rm iid} \mathcal{N}(0, \sigma^2)$*,* $k \in [\![M]\!]$*. Given some* $s > 0$*, in the UDS case,* i.e., $\eta(l) = 1/N$ *and* $\boldsymbol{D}^2 = N\boldsymbol{I}_N$*, we have*

$$\frac{1}{\sqrt{M}}\,\|\boldsymbol{D}\boldsymbol{n}\| \;\leq\; \varepsilon^0_{\sigma,s}(N, M) := \sigma\sqrt{N}\,(1 + \frac{\sqrt{s}}{\sqrt{2M}} + \frac{s}{M})^{1/2}, \quad (2.20)$$

[3] While Cor. 2.9 can be easily extended to the complex field, we restrict it anyway to a real Gaussian noise, *i.e.*, the noise of interest in this thesis.

with probability exceeding $1 - \exp(-\frac{s}{2})$*. More generally, in a VDS context with arbitrary pmf* η*,*

$$\frac{1}{\sqrt{M}}\,\|\boldsymbol{D}\boldsymbol{n}\| \;\leq\; \varepsilon_{\sigma,s}(N, M, \rho),$$

where

$$\varepsilon_{\sigma,s}(N, M, \rho) := \sigma\sqrt{N}\,\big[(1 + \tfrac{\sqrt{s}}{\sqrt{2M}} + \tfrac{s}{M}) + 4e(2\log M + s)\max\big(\tfrac{s}{M}, \tfrac{\sqrt{s}}{\sqrt{M}}\big)\rho\big]^{\frac{1}{2}}, \tag{2.21}$$

with probability exceeding $1 - 3\exp(-\frac{s}{2})$ *(e.g., with probability greater than* 0.95 *for* $s = 8.2$*).*

We postpone the proof of this corollary to § 2.6.3.

Remark 2.10. *By comparing* (2.21) *to* (2.20)*, we observe that the variability of* $\boldsymbol{D}$ *induces a bias behaving like* $O(\rho\log M\max(\frac{1}{M}, \frac{\sqrt{1}}{\sqrt{M}}))$ *in* $\varepsilon_{\sigma,s}(N, M, \rho)$ *compared to the simpler bound* $\varepsilon^0_{\sigma,s}$ *reached by the UDS. Therefore, if* ρ *is slowly growing when* N *increases (see Remark 2.8), and for a large* M*, this bias is small and the two noise levels, for the VDS and UDS schemes, are thus comparable.*

The previous remark will be used in conjunction with Cor. 2.9 in § 4.7.1, § 4.5.2, and § 5.3.2 in order to bound the noise in our compressive FTI and SP-FTI under a Gaussian noise assumption.

2.6 Proofs

We now present the proofs of the main results presented in this chapter.

2.6.1 Proof of Thm. 2.4

The proof involves three steps: *(i)* in Prop. 2.2 let $\boldsymbol{A} = \boldsymbol{P}_\Omega\boldsymbol{\Phi}^*\boldsymbol{\Psi}$; *(ii)* precondition the matrix $\boldsymbol{A}$ according to Prop. 2.3 so that it satisfies RIP and yields $M^{-1/2}\boldsymbol{D}\boldsymbol{y} = M^{-1/2}\boldsymbol{D}\boldsymbol{\Phi}^*\boldsymbol{\Psi}\boldsymbol{x} + M^{-1/2}\boldsymbol{D}\boldsymbol{n}$; *(iii)* apply a change of variable $\boldsymbol{u} \to \boldsymbol{\Psi}^* \boldsymbol{u}$ in (2.3) using the fact that $\boldsymbol{\Psi}$ is an orthonormal basis.

2.6.2 Proof of Thm. 2.7

In the context defined by Thm. 2.7, we define M independent, positive r.v.s $X_j \sim |n_j|^2 \, \eta(\beta)^{-1}$ (for $1 \leq j \leq M$), such that $\mathbb{E} X_j = |n_j|^2 \, \mathbb{E}\eta(\beta)^{-1} = |n_j|^2 N$, and $\mathbb{E} \sum_j X_j = N\|\boldsymbol{n}\|^2$. Therefore, the left-hand side of (2.19) reads

$$\tfrac{1}{M}\|\boldsymbol{D}\boldsymbol{n}\|^2 = \tfrac{1}{M} \textstyle\sum_{j=1}^M X_j = \tfrac{N}{M}\|\boldsymbol{n}\|^2 + \tfrac{1}{M} \textstyle\sum_{j=1}^M (X_j - \mathbb{E}X_j). \qquad (2.22)$$

Moreover, for all $j \in [\![M]\!]$ and any integer $p \geq 1$, we have $\mathbb{E}X_j^p \leq \|\boldsymbol{n}\|_\infty^{2p} \, \mathbb{E}_\beta \eta(\beta)^{-p}$, so that, for $|s| < 1/(4e\rho N \|\boldsymbol{n}\|_\infty^2)$ and using $p! \geq e^{1-p} p^p$ (Stirling bound [97]),

$$\begin{aligned}
\mathbb{E}e^{s(X_j - \mathbb{E}X_j)} &= \textstyle\sum_{p=0}^{+\infty} \frac{1}{p!} s^p \, \mathbb{E}(X_j - \mathbb{E}X_j)^p \\
&= 1 + \textstyle\sum_{p=2}^{+\infty} \frac{1}{p!} s^p \, \mathbb{E}(X_j - \mathbb{E}X_j)^p \\
&\leq 1 + \textstyle\sum_{p=2}^{+\infty} \frac{1}{p!} 2^p |s|^p \, \mathbb{E}X_j^p \\
&\leq 1 + \textstyle\sum_{p=2}^{+\infty} 2^p |s|^p \, \|\boldsymbol{n}\|_\infty^{2p} \, \frac{1}{p!} \mathbb{E}_\beta \eta(\beta)^{-p} \\
&\leq 1 + \textstyle\sum_{p=2}^{+\infty} 2^p |s|^p \, \|\boldsymbol{n}\|_\infty^{2p} \, e^{p-1} \big(\tfrac{1}{p} (\mathbb{E}_\beta \eta(\beta)^{-p})^{1/p}\big)^p \\
&\leq 1 + \textstyle\sum_{p=2}^{+\infty} 2^p |s|^p \, \|\boldsymbol{n}\|_\infty^{2p} \, e^{p-1} \rho^p N^p \\
&= 1 + 4es^2 \, \|\boldsymbol{n}\|_\infty^4 \, \rho^2 N^2 \textstyle\sum_{p=0}^{+\infty} 2^p e^p |s|^p \, \|\boldsymbol{n}\|_\infty^{2p} \, \rho^p N^p \\
&\leq 1 + 8es^2 \, \|\boldsymbol{n}\|_\infty^4 \, \rho^2 N^2 \;\leq\; \exp(16e^2 s^2 \, \|\boldsymbol{n}\|_\infty^4 \, \rho^2 N^2/2),
\end{aligned}$$

where the second line uses $\mathbb{E}_X|X - \mathbb{E}X|^p = \mathbb{E}_X|\mathbb{E}_{X'}(X - X')|^p \leq \mathbb{E}_X\mathbb{E}_{X'}(|X| + |X'|)^p \leq 2^{p-1}\mathbb{E}_X\mathbb{E}_{X'}(|X|^p + |X'|^p) \leq 2^p \mathbb{E}|X|^p$, for two iid r.v.s X and X'.

This shows that the r.v.s $\{X_j - \mathbb{E}X_j : 1 \leq j \leq M\}$ are positive sub-exponential r.v.s with parameter $\lambda = 4e\|\boldsymbol{n}\|_\infty^2 \rho N$ [97], since $\mathbb{E}e^{s(X_j - \mathbb{E}X_j)} \leq \exp(s^2\lambda^2/2)$ for $|s| < 1/\lambda$. Therefore, Bernstein inequality [97, Thm.

1.13] shows that their sum concentrates around their mean, *i.e.*,

$$\mathbb{P}[\tfrac{1}{M}\textstyle\sum_{j=1}^{M}(X_j - \mathbb{E}X_j) > t\lambda] \le \exp(-\tfrac{M}{2}\min(t^2,t)).$$

With the change of variable $\min(t^2,t) = s$, *i.e.*, $t = \max(\sqrt{s}, s)$, this gives $\mathbb{P}[\frac{1}{M}\sum_{j=1}^{M}(X_j - \mathbb{E}X_j) > \max(s,\sqrt{s})\lambda] \le \exp(-\frac{M}{2}s)$, or equivalently, with $s \leftarrow s/M$,

$$\mathbb{P}[\tfrac{1}{M}\textstyle\sum_{j=1}^{M}(X_j - \mathbb{E}X_j) > \max\big(\tfrac{s}{M}, \tfrac{\sqrt{s}}{\sqrt{M}}\big)\lambda\,] \le \exp(-\tfrac{s}{2}).$$

Therefore, with probability exceeding $1-\exp(-\frac{s}{2})$, (2.22) provides

$$\tfrac{1}{M}\,\|\boldsymbol{D}\boldsymbol{n}\|^2 \le \tfrac{N}{M}\|\boldsymbol{n}\|^2 + 4e\max\big(\tfrac{s}{M}, \tfrac{\sqrt{s}}{\sqrt{M}}\big)\|\boldsymbol{n}\|_\infty^2\rho N,$$

which gives the result.

2.6.3 Proof of Cor. 2.9

By union bound, we first know that $\mathbb{P}[\exists i \in [\![M]\!] : |n_i| \ge t\sigma] \le M\exp(-t^2/2)$. Therefore, $\|\boldsymbol{n}\|_\infty \le \sigma\sqrt{2\log M + s}$ with probability exceeding $1-\exp(-s/2)$. Moreover, since n_i^2 is a χ^2 distribution, we have $\|\boldsymbol{n}\|^2 \le \sigma^2(M+\sqrt{M}\sqrt{s/2}+s)$ with probability exceeding $1-\exp(-s/2)$ (see, *e.g.*, [64, Lem. 1]).

Therefore, by union bound over the failure of these two events and over the one covered by Thm. 2.7 in (2.19), we have

$$\begin{aligned}\tfrac{1}{M}\,\|\boldsymbol{D}\boldsymbol{n}\|^2 &\le \tfrac{N}{M}\|\boldsymbol{n}\|^2 + 4e\max\big(\tfrac{s}{M}, \tfrac{\sqrt{s}}{\sqrt{M}}\big)\|\boldsymbol{n}\|_\infty^2\rho\, N\\ &\le \sigma^2\big[\tfrac{N}{M}(M + \tfrac{\sqrt{Ms}}{\sqrt{2}} + s) + 4e(2\log M + s)\max\big(\tfrac{s}{M}, \tfrac{\sqrt{s}}{\sqrt{M}}\big)\rho\, N\big]\\ &= \sigma^2 N\,\big[(1 + \tfrac{1}{\sqrt{2}}\tfrac{\sqrt{s}}{\sqrt{M}} + \tfrac{s}{M}) + 4e(2\log M + s)\max\big(\tfrac{s}{M}, \tfrac{\sqrt{s}}{\sqrt{M}}\big)\rho\big],\end{aligned}$$

with probability exceeding $1 - 3\exp(-\frac{s}{2})$, which provides the result. Note that in the UDS case, we just need to consider the χ^2-bound on $\|\boldsymbol{n}\|^2$ above since $\boldsymbol{D}^2 = N\boldsymbol{I}_N$ is deterministic.

Chapter 3

Compressive Hadamard Sensing with Haar Sparsity Basis

We have seen in the previous chapter that an NDS scheme (VDS or MDS) can bridge the gap between the theory and application of CS by adjusting the sampling strategy to the values of the local and multilevel coherence between the sensing and sparsity bases. The efficacy of the VDS and MDS schemes depend on the accurate computation of the coherence values. One can indeed numerically compute those values, but it does not provide theoretical insights. Closed-form local and multilevel coherence values, on the other hand, yields meaningful sample-complexity bound which in turn provides explicit signal recovery guarantees.

This content of this chapter has been published in [84].

3.1 Definition of the Problem

In this chapter, we tackle an important problem in the applications of CS theory: recovering a signal from subsampled Hadamard measurements using the Haar wavelet sparsity basis. In particular, in a wide range of imaging modalities, *e.g.*, optical multiplexing or single-pixel camera [34], the sensing process can be modeled as taking measurements from the Hadamard transform. Moreover, considering the Haar wavelet basis paves the way to study other wavelet bases in combination with the Hadamard matrix. In this context, the main question becomes: how to design an efficient sampling strategy for subsampling the Hadamard measurements.

As discussed in § 2.2, traditional CS relying on orthonormal sensing systems [21] suggests selecting the rows of the sensing matrix (in our case, the Hadamard matrix) uniformly at random, *i.e.*, according to a UDS. This approach fails when the target signal is sparse or compressible in a sparsity basis that is too coherent with the sensing basis (see § 2.2). One example of this failure is the Hadamard-Haar system, where the sensing basis (Hadamard) is maximally coherent with the sparsity basis (Haar wavelet).

Several empirical [12, 81, 82] and theoretical evidences [7, 8, 61] suggest using an NDS strategy [7, 61, 93], which densifies the subsampling of the lower Hadamard frequencies, to obtain superior signal reconstruction quality. In a general context, Krahmer and Ward [61] and Adcock *et al.* [7] arguably began to replace the notion of global coherence with its local versions, *i.e.*, *local coherence* and *multilevel coherence* parameters, respectively. The idea in these works is to discriminate the elements of the sensing basis (*e.g.*, Hadamard) in favor of those that are highly coherent with all the elements of the sparsity basis (*e.g.*, Haar).

Although there are other versions of non-uniform sampling strategies, *e.g.*, [15, 17, 66, 93], we here focus on the VDS scheme of Krahmer and Ward [61] (for uniform recovery guarantee) and the MDS scheme of Adcock *et al.* [7] (for non-uniform recovery guarantee). These allow us to derive suitable sampling strategies for Hadamard-Haar systems.

3.1.1 Related Works

Non-uniform density sampling (theory and application): The idea of non-uniform density sampling dates back to the emergence of CS. Donoho in [32] proposed a two-level sampling approach for recovering wavelet coefficients, where the coarse wavelet scale coefficients are fully sampled and the remaining coefficients are subsampled with UDS strategy. This idea was later extended by Tsaig and Donoho in [116] to a multiscale setup. Puy *et al.* [93] advocated a convex optimization procedure for minimizing the coherence between the sensing and sparsity bases. A non-uniform density sampling approach is proposed by Wang and Arce in [124] based on the statistical models of natural images. Bigot *et al.* [15] introduced the notion of block sampling for CS, based on acquiring the blocks of measurements instead of isolated measurements; see also a similar study of Polak *et al.* [91]. Boyer *et al.* incorporated the idea of block sampling with structured sparsity in [17], whose stable and robust recovery guarantee was later proved by Adcock *et al.* in [6]. A RIP-based recovery guarantee is presented by Krahmer and Ward [61] based on the notion of *random bounded orthonormal systems* introduced in [42, 95]. The sampling strategy in [61] is controlled by the local coherence between the sensing and sparsity bases. Adcock *et al.* [7] provided a novel MDS scheme based on the local sparsity and multilevel coherence between the sensing and sparsity bases. A generalization of the RIP for MDS strategy of [7] in finite dimensions (and infinite dimensions) has been analyzed by Li and Adcock in [66] (respectively, Adcock *et al.* [5]).

Most of the works above also tackled the problem of signal recovery from subsampled Fourier measurements using wavelet sparsity basis (Fourier-Wavelet system), *e.g.*, [7, 9, 17, 61, 93]. In this context, the applications of non-uniform density sampling have shown promising results in MRI [71, 72, 99] and interferometric HS imaging [78–80, 83, 86].

Imaging applications of the Hadamard transform: The Hadamard matrix has become an emerging element in many computational imaging applications relying on optical multiplexing or single-pixel imaging, such as Hadamard spectroscopy [28, 99, 106], lensless camera [56], 3-D video imaging [131], laser-based failure-analysis [109], compressive holography [27], single pixel Fourier transform interferometry [59, 81, 82], digital holography [76], intracranial electroencephalogram acquisition [12], single-pixel camera [132], and micro-optoelectromechanical systems [29].

Most of the works above have already been designed based on different NDS schemes, that can be categorized, based on the sampling design, in four groups, *i.e.*, the rows of the Hadamard matrix are selected with respect to *(i)* UDS [28], *(ii)* MDS [12, 81, 99], *(iii)* low-pass sampling where the first M rows are selected [76, 131], and *(iv)* half-half sampling where the first $M/2$ rows are always selected and the other $M/2$ rows are selected uniformly at random among the rest of the rows [106]. Although all these works have obtained high quality signal recovery, they do not provide an explicit recovery guarantee.

Moreover, other contributions, *e.g.*, on video CS [123], 3-D imaging [108], remote sensing [73], terahertz imaging [25], and single-pixel camera [34, 111], which utilize random binary patterns (*e.g.*, Bernoulli matrices) can potentially adopt Hadamard sensing with no hardware burden.

Hadamard-wavelet systems (theory): The recovery of 1-D signals that are sparse in an orthonormal wavelet basis (from subsampled Hadamard measurements) has been studied in [10] in the context of MDS. The problem of signal reconstruction from the Hadamard (or binary) measurements has recently received attention in other contexts than CS, *e.g.*, in generalized sampling methods where the goal is to recover an infinite-dimensional signal from a full set of measurements (without subsampling) via a linear reconstruction. In this context, there exist several works where the sampling space is assumed to be the domain of Hadamard transform and the reconstruction takes place in the span of some wavelet basis (see, *e.g.*, [5, 53] and [20] for a survey). In this work, however, we address the recovery of both 1-D and 2-D finite-dimensional signals using the Haar wavelet sparsity basis. To the best of our knowledge, four other papers have addressed the relationship between the 1-D Hadamard and 1-D Haar wavelet bases [36, 40, 94, 115]. In the next section, as well as Table 3.1, we compare our contribution

		Krahmer and Ward [61]	Adcock *et al.* [7]	Adcock *et al.* [9]	Li and Adcock [66]	Antun [10]	Adcock *et al.* [5]	Hansen and Thesing [52]	Thesing and Hansen [114]	Hansen and Terhaar [53]	Thesing and Hansen [113]	This chapter: Thm. 3.8	This chapter: Thm. 3.10
Sensing basis	any orthonormal	✓	✓		✓		✓						
	Fourier	✓	✓	✓	✓								
	Hadamard					✓	✓	✓	✓	✓	✓	✓	✓
Sparsity basis	any orthonormal	✓	✓		✓		✓						
	Haar wavelet	✓		✓	✓	✓	✓		✓			✓	✓
	Daubechies wavelet					✓					✓		
	orthogonal wavelet		✓				✓	✓		✓			
Signal dimensions	1-D (vector)	✓	✓	✓	✓	✓	✓				✓	✓	✓
	2-D (matrix)	✓										✓	✓
	d-D (tensor)							✓	✓	✓			
Sampling strategy	VDS	✓						—	—	—		✓	
	MDS		✓	✓	✓	✓	✓	—	—	—	✓		✓
Signal type	finite-dimensional	✓	✓	✓	✓	✓						✓	✓
	infinite-dimensional						✓	✓	✓	✓	✓		
Recovery guarantee	uniform	✓			✓		✓	—	—	—		✓	
	non-uniform		✓	✓		✓		—	—	—	✓		✓
Context	compressive sensing	✓	✓	✓	✓	✓	✓				✓	✓	✓
	generalized sampling							✓	✓	✓			

Table 3.1: Comparison between the state-of-the-art works and our contribution. In this table, we consider those contributions in the field of CS that are based on the local coherence (developed by Krahmer and Ward in [61]) and multilevel coherence (developed by Adcock *et al.* in [7]) parameters.

with the state-of-the-art works. This chapter provides explicit CS strategies for Hadamard-Haar systems; an association that was seemingly not covered by the related literature (see, *e.g.*, Table 3.1). Main results are presented in Thm. 3.8 (for the uniform recovery guarantee) and in Thm. 3.10 (for the non-uniform guarantee). Note that the result of this chapter will be applied to the CS of HS data with single-pixel imaging (when the light illumination is spatially coded with Hadamard system) in Chapter 5. During the finalization of this thesis, we became aware of this recent survey of Calderbank *et al.* [20] that pursue similar objectives. Compared to our result in Thm. 3.10, Thm. 5.8 in [20], which is stated from [113], covers the problem of infinite-dimensional signal recovery using Daubechies wavelets. Moreover, Fig. 3 in [20] advocates the same structure as in Fig. 3.6-bottom-right for infinite-dimensional

2-D Hadamard-Haar system. However, the mathematical expression in Prop. 3.6-(ii) for modeling those structures is original.

In this chapter, we first deliver a short introduction to the Hadamard and Haar wavelet bases in § 3.2. In § 3.3, we present our main results in Thm. 3.8 and Thm. 3.10. The technical proofs are postponed to § 3.5. Finally, we conduct a series of numerical tests in § 3.4, which confirms the efficiency of our analysis.

3.2 Definition of the Haar and Hadamard Bases

1-D Discrete Haar Wavelet (DHW) basis: Fix $N = 2^r$ for some $r \in \mathbb{N}$. The DHW basis of $\mathbb{R}^N$ consists of N functions

$$\{\psi_j^{1\mathrm{d}}\}_{j=1}^N := \{\bar{h}\} \cup \{h_{s,p}^{(1)} : 0 \le s \le r-1, 0 \le p \le 2^s - 1\},$$

where, for $\tau \in [\![N-1]\!]_0$, $\bar{h}(\tau) := 2^{-r/2}$ is the constant (scaling) function and $h_{s,p}^{(1)}(\tau) := 2^{\frac{s-r}{2}} h(2^{s-r}\tau - p)$ is the wavelet function at scale (or resolution) s and position p, with $h(\tau)$ equals 1, -1, and 0 over $[0, 1/2)$, $[1/2, 1)$, and $\mathbb{R}\backslash[0, 1)$, respectively (see [74, Page 2], [105, Page 6], or [9]), *i.e.*,

$$h_{s,p}^{(1)}(\tau) = \begin{cases} 2^{\frac{s-r}{2}}, & \text{for } p2^{r-s} \le \tau < (p+\frac{1}{2})2^{r-s}, \\ -2^{\frac{s-r}{2}}, & \text{for } (p+\frac{1}{2})2^{r-s} \le \tau < (p+1)2^{r-s}, \\ 0, & \text{otherwise.} \end{cases} \tag{3.1}$$

In a matrix form, DHW basis in $\mathbb{R}^{N\times N}$ can be constructed [36, 40] from the recursive relation

$$\boldsymbol{\Psi}_{\mathrm{dhw}} := \boldsymbol{W}_r^{(1)} := \frac{1}{\sqrt{2}} \left[\boldsymbol{W}_{r-1}^{(1)} \otimes \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \boldsymbol{I}_{2^{r-1}} \otimes \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right], \text{ with } \boldsymbol{W}_0^{(1)} := [1], \tag{3.2}$$

which collects in its columns all the functions of $\{\psi_j^{1\mathrm{d}}\}_{j=1}^N$ (see Lemma 3.1).

In order to extend the DHW basis to the 2-D Haar wavelet basis, we need to define $N-1$ window functions

$$h_{s,p}^{(0)}(\tau) = \begin{cases} 2^{\frac{s-r}{2}}, & \text{for } p2^{r-s} \le \tau < (p+1)2^{r-s}, \\ 0, & \text{otherwise}, \end{cases} \tag{3.3}$$

for the resolution $0 \le s \le r-1$ and position $0 \le p \le 2^s - 1$ parameters. Similar to the construction of the DHW basis in (3.2), we define the matrix

$$\boldsymbol{W}_r^{(0)} := \frac{1}{\sqrt{2}} \left[\boldsymbol{W}_{r-1}^{(0)} \otimes \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \boldsymbol{I}_{2^{r-1}} \otimes \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right], \quad \text{with } \boldsymbol{W}_0^{(0)} := [1], \tag{3.4}$$

which collects in its first column the function $\bar{h}$ and in the other columns all the functions $\{h_{s,p}^{(0)}\}$ (see Lemma 3.1).

Associated with the DHW basis, the *1-D dyadic levels* $\mathcal{T}^{1\mathrm{d}} := \{\mathcal{T}_l^{1\mathrm{d}}\}_{l=0}^r$ gather coefficient indices with identical wavelet levels; they are defined as

$$\mathcal{T}_l^{1\mathrm{d}} := [\![2^l]\!] \backslash [\![2^{l-1}]\!], \quad \text{for } l \in [\![r]\!], \quad \text{and } \mathcal{T}_0^{1\mathrm{d}} := \{1\}, \tag{3.5}$$

with cardinality $|\mathcal{T}_l^{1\mathrm{d}}| = 2^{l-1}$, for $l \in [\![r]\!]$. We also define the left-complement of dyadic levels as $\mathcal{T}_{<l}^{1\mathrm{d}} := \bigcup_{j=0}^{l-1} \mathcal{T}_j^{1\mathrm{d}} = [\![2^{l-1}]\!]$ for $l \in [\![r]\!]$ and $\mathcal{T}_{<0}^{1\mathrm{d}} := \emptyset$. These levels are important to isolate the indices of the columns (components) of $\boldsymbol{\Psi}_{\mathrm{dhw}} = \boldsymbol{W}_r^{(1)}$ (resp. $(\boldsymbol{\Psi}_{\mathrm{dhw}})^\top \boldsymbol{x}$) associated with a given scale, as well as those of $\boldsymbol{W}_r^{(0)}$

Lemma 3.1. *For $l \in [\![r]\!]_0$ and $a \in \{0,1\}$, the matrix $\boldsymbol{W}_r^{(a)} \boldsymbol{P}_{\mathcal{T}_l^{1\mathrm{d}}}^\top \in \mathbb{R}^{2^r \times |\mathcal{T}_l^{1\mathrm{d}}|}$ collects in its columns all the functions $\{h_{l-1,p}^{(a)}\}_{p=0}^{2^{l-1}-1}$, if $l \in [\![r]\!]$, and if $l = 0$, it collects $\bar{h}$ in its single column.*

Proof. See § 3.5.1. □

There exist two natural ways to construct a 2-D wavelet basis from a 1-D basis, *i.e.*, by tensor product of two 1-D bases, and by following a multi-resolution analysis (see [74, § 7.7], [14], or [90]), which amounts to multiplying all possible pairs of wavelet and scaling functions sharing the same resolution. We describe below those two approaches for 2-D Haar wavelet construction.

2-D Anisotropic Discrete Haar Wavelet (ADHW) basis: For the first approach, the tensor product of two DHW bases leads to an anisotropic 2-D DHW basis. For $N = 2^r$ and some $r \in \mathbb{N}$, we consider the scaling and wavelet functions $\bar{h}$ and $h_{s,p}^{(1)}$ defined above, and we build the ADHW basis of $\mathbb{R}^{N^2}$ as

$$\{\psi_j^{\text{aniso}}\}_{j=1}^{N^2} := \{\psi_{j_1}^{\text{1d}}\psi_{j_2}^{\text{1d}} : j \overset{N}{\rightleftharpoons} (j_1, j_2)\},$$

which provides N^2 possible functions. This basis is of interest for image compression [30], sparsity basis for MRI images [17], and sparsity basis for monochromatic images in fluorescence spectroscopy [80]. In particular, Neumann and von Sachs [89] showed that if a multi-dimensional signal has different degrees of smoothness in different directions, the tensor wavelet construction is a better choice for signal estimation.

In a matrix form, the ADHW basis in $\mathbb{R}^{N^2 \times N^2}$ can be constructed [74, 90] as

$$\boldsymbol{\Psi}_{\text{adhw}} := \boldsymbol{\Psi}_{\text{dhw}} \otimes \boldsymbol{\Psi}_{\text{dhw}},$$

where $\boldsymbol{\Psi}_{\text{adhw}}$ collects in its columns all the functions of $\{\psi_j^{\text{aniso}}\}_{j=1}^{N^2}$. Associated with the ADHW basis, we define the *2-D anisotropic wavelet levels* $\mathcal{T}^{\text{aniso}} := \{\mathcal{T}_l^{\text{aniso}}\}_{l=1}^{r^2}$ where $\mathcal{T}_l^{\text{aniso}} := \overline{\mathcal{T}_{l_1}^{\text{1d}} \times \mathcal{T}_{l_2}^{\text{1d}}}$, for $l \in [\![(r+1)^2]\!]$ and $l_1, l_2 \in [\![r]\!]_0$, with the relation $l \overset{r+1}{\rightleftharpoons} (l_1+1, l_2+1)$ and hence $|\mathcal{T}_l^{\text{aniso}}| = |\mathcal{T}_{l_1}^{\text{1d}}| \cdot |\mathcal{T}_{l_2}^{\text{1d}}|$. These levels thus gather the indices of wavelet coefficients associated with the constant resolution (see the illustration on Fig. 3.1-right for $N = 8$).

Remark 3.2. *According to the construction of $\boldsymbol{\Psi}_{\text{adhw}}$ and $\mathcal{T}^{\text{aniso}}$, one can use Lemma 3.1 and Lemma 3.13 to show that, for $l \overset{r+1}{\rightleftharpoons} (l_1+1, l_2+1)$,*

$$\boldsymbol{\Psi}_{\text{adhw}} \boldsymbol{P}_{\mathcal{T}_l^{\text{aniso}}}^{\top} = \left(\boldsymbol{\Psi}_{\text{dhw}} \boldsymbol{P}_{\mathcal{T}_{l_2}^{\text{1d}}}^{\top}\right) \otimes \left(\boldsymbol{\Psi}_{\text{dhw}} \boldsymbol{P}_{\mathcal{T}_{l_1}^{\text{1d}}}^{\top}\right).$$

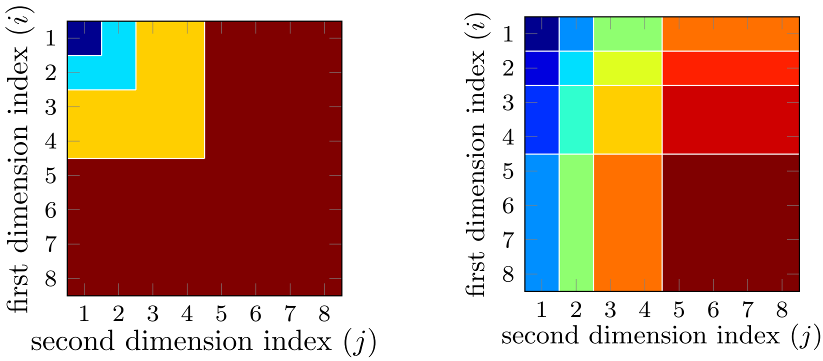

Figure 3.1: An example of the 2-D isotropic wavelet levels $\mathcal{T}_l^{\mathrm{iso}}$ for $l \in [\![r]\!]_0$ (left) versus 2-D anisotropic wavelet levels $\mathcal{T}_l^{\mathrm{aniso}}$ for $l \in [\![(r+1)^2]\!]$ (right) with $N = 8$ (or $r = 3$). Each area represents the subset of pairs of indices (i, j) that belong to a level l.

2-D Isotropic Discrete Haar Wavelet (IDHW) basis: The second type of the 2-D DHW basis is built from a multi-resolution analysis [74]. Fix $N = 2^r$ for some $r \in \mathbb{N}$. Let $\bar{h}$, $h^{(1)}$, and $h^{(0)}$ be the scaling, wavelet, and window functions defined above. Following [74] the IDHW basis $\{\psi_j^{\mathrm{iso}}\}_{j=1}^{N^2}$ of $\mathbb{R}^{N^2}$ consists of the functions

$$\{\phi^{(00)}\} \cup \{\phi^{(ab)}_{s,(p_1,p_2)} : 0 \le s \le r-1,\ 0 \le p_1, p_2 \le 2^s - 1,\\ (a,b) \in \{0,1\}^2 \backslash \{0,0\}\},$$

such that

$$\begin{aligned}\phi^{(00)}(\tau_1,\tau_2) &= \bar{h}(\tau_1)\bar{h}(\tau_2),\\ \phi^{(ab)}_{s,(p_1,p_2)}(\tau_1,\tau_2) &= h^{(a)}_{s,p_1}(\tau_1) h^{(b)}_{s,p_2}(\tau_2),\end{aligned}$$

where $0 \le s \le r-1$ and $0 \le p_1, p_2 \le 2^s - 1$ are the resolution and position indices, respectively, *i.e.*, there are N^2 possible functions.

To construct the orthonormal matrix $\mathbf{\Psi}_{\mathrm{idhw}} \in \mathbb{R}^{N^2 \times N^2}$ associated with the 2-D IDHW basis, we leverage the 1-D partitions $\mathcal{T}_l^{\mathrm{1d}}$ defined above so that the column ordering of $\boldsymbol{\Psi}_{\rm idhw}$ will ease any further column selection[1] (*e.g.*, in § 3.3).

We first define, for $l \in [\![r]\!],\ (a,b) \in \{0,1\}^2\backslash\{(0,0)\}$, the submatrices

$$\boldsymbol{\Psi}_l^{(ab)} := \Big(\boldsymbol{W}^{(a)}\boldsymbol{P}_{\mathcal{T}_l^{\rm 1d}}^\top\Big) \otimes \Big(\boldsymbol{W}^{(b)}\boldsymbol{P}_{\mathcal{T}_l^{\rm 1d}}^\top\Big) \in \mathbb{R}^{N^2\times|\mathcal{T}_l^{\rm 1d}|^2},$$

and $\boldsymbol{\Psi}_0^{(00)} := \mathbf{1}_{N^2}$. For each level l, these submatrices clearly contain all the functions $\{\phi_{l-1,(p_1,p_2)}^{(ab)} : p_1, p_2 \in [\![2^l-1]\!]_0$. Moreover, since $\mathcal{T}_{<l}^{\rm 1d} = [\![2^{l-1}]\!]$ and $\mathcal{T}_l^{\rm 1d} = [\![2^l]\!]/[\![2^{l-1}]\!]$, the following disjoint sets

$$\begin{aligned}\mathcal{T}_0^{(00)} &:= \overline{\mathcal{T}_0^{\rm 1d}\times\mathcal{T}_0^{\rm 1d}},\ \mathcal{T}_l^{(11)} := \overline{\mathcal{T}_l^{\rm 1d}\times\mathcal{T}_l^{\rm 1d}},\\ \mathcal{T}_l^{(10)} &:= \overline{\mathcal{T}_{<l}^{\rm 1d}\times\mathcal{T}_l^{\rm 1d}},\ \mathcal{T}_l^{(01)} := \overline{\mathcal{T}_l^{\rm 1d}\times\mathcal{T}_{<l}^{\rm 1d}},\end{aligned}$$

are such that $|\mathcal{T}_0^{(00)}| = 1$, $|\mathcal{T}_l^{(11)}| = |\mathcal{T}_l^{(01)}| = |\mathcal{T}_l^{(10)}| = |\mathcal{T}_l^{\rm 1d}|^2 = 2^{2(l-1)}$, and

$$\mathcal{T}_0^{(00)} \cup \bigcup_{l\in[\![r]\!]}\Big(\mathcal{T}_l^{(01)}\cup\mathcal{T}_l^{(11)}\cup\mathcal{T}_l^{(10)}\Big) = [\![N^2]\!].$$

Therefore, as illustrated in Fig. 3.1-left, we can order the columns of $\boldsymbol{\Psi}_{\rm idhw}$ such that, for the 2-D isotropic wavelet levels $\mathcal{T}^{\rm iso} := \{\mathcal{T}_l^{\rm iso}\}_{l=0}^r$ defined by

$$\mathcal{T}_0^{\rm iso} := \mathcal{T}_0^{(00)},\ \text{and}\ \mathcal{T}_l^{\rm iso} := \mathcal{T}_l^{(01)} \,\underline{\cup}\, \mathcal{T}_l^{(11)} \,\underline{\cup}\, \mathcal{T}_l^{(10)},\ \in [\![r]\!]_0, \tag{3.6}$$

we have $\boldsymbol{\Psi}_{\rm idhw}\boldsymbol{P}_{\mathcal{T}_0^{\rm iso}}^\top = \boldsymbol{\Psi}_0^{(00)}$, $\boldsymbol{\Psi}_{\rm idhw}\boldsymbol{P}_{\mathcal{T}_l^{(ab)}}^\top = \boldsymbol{\Psi}_l^{(ab)}$, and

$$\boldsymbol{\Psi}_{\rm idhw}\boldsymbol{P}_{\mathcal{T}_l^{\rm iso}}^\top = [\boldsymbol{\Psi}_l^{(01)}, \boldsymbol{\Psi}_l^{(11)}, \boldsymbol{\Psi}_l^{(10)}]$$

for $l \in [\![r]\!]$ and $(a,b) \in \{0,1\}^2\backslash\{(0,0)\}$.

[1]Note that the ordering of the functions in the definition of the Haar bases $\psi^{\rm 1d}$, $\psi^{\rm aniso}$, and $\psi^{\rm iso}$ are arbitrary. In this thesis, however, we are concerned only by the column ordering of the matrix version of those bases, *i.e.*, $\boldsymbol{\Psi}_{\rm dhw}$, $\boldsymbol{\Psi}_{\rm adhw}$, and $\boldsymbol{\Psi}_{\rm idhw}$, as specified in the text.

(Paley-ordered) Hadamard matrix: We now present an important family of orthogonal matrices introduced by J. Hadamard [51], *i.e.*, the Hadamard matrix, that has appeared in various fields, *e.g.*, coding theory [92], harmonic analysis [60], and optics [59]. There exist mainly three constructions of the Hadamard matrix, each with specific row ordering, called ordinary (or Sylvester)-, sequency-, and Paley-ordered Hadamard matrix [110, 133]. In this thesis, we focus only on the Paley-ordered Hadamard matrix. But all our results are clearly extendable to the other two constructions after proper reordering (see [10, § 4] for the row ordering).

Given $r \in \mathbb{N}$, the $2^r \times 2^r$ Hadamard matrix [36, 55] $\boldsymbol{\Phi}_{\mathrm{had}} := \boldsymbol{H}_r \in \{\pm 2^{-r/2}\}^{2^r \times 2^r}$ is defined by

$$\boldsymbol{H}_r := \frac{1}{\sqrt{2}} \left[\boldsymbol{H}_{r-1} \otimes \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \boldsymbol{H}_{r-1} \otimes \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right], \boldsymbol{H}_0 := [1]. \tag{3.7}$$

Note that this recurrence relation bears some resemblance with the one of the Haar wavelet basis in (3.2). Moreover, from (3.7), we can easily show that $\boldsymbol{\Phi}_{\mathrm{had}}$ is symmetric, *i.e.*, $\boldsymbol{\Phi}_{\mathrm{had}}^\top = \boldsymbol{\Phi}_{\mathrm{had}}$. The Hadamard transformation of a signal $\boldsymbol{x} \in \mathbb{C}^N$ with $N = 2^r$ reads $\boldsymbol{z} = \boldsymbol{\Phi}_{\mathrm{had}}^\top \boldsymbol{x}$. For 2-D signals, the Hadamard basis is defined by $\boldsymbol{\Phi}_{\mathrm{2had}} := \boldsymbol{\Phi}_{\mathrm{had}} \otimes \boldsymbol{\Phi}_{\mathrm{had}} \in \mathbb{R}^{N^2 \times N^2}$ so that the Hadamard transformation of a matrix $\boldsymbol{X} \in \mathbb{C}^{N \times N}$ is $\boldsymbol{Z} = \boldsymbol{\Phi}_{\mathrm{had}}^\top \boldsymbol{X} \boldsymbol{\Phi}_{\mathrm{had}}$, or equivalently $\mathrm{vec}(\boldsymbol{Z}) = \boldsymbol{\Phi}_{\mathrm{2had}}^\top \mathrm{vec}(\boldsymbol{X})$.

Remark 3.3. *Following the definition of $\boldsymbol{\Phi}_{\mathrm{2had}}$, $\mathcal{T}^{\mathrm{aniso}}$, and $\mathcal{T}^{\mathrm{iso}}$ and using Lemma. 3.13 we directly deduce that, for $l_1, l_2 \in [\![r]\!]_0$ and $l \in [\![(r+1)^2]\!]$,*

$$\boldsymbol{P}_{\mathcal{T}_l^{\mathrm{aniso}}} \boldsymbol{\Phi}_{\mathrm{2had}}^\top = \left(\boldsymbol{P}_{\mathcal{T}_{l_2}^{\mathrm{1d}}} \boldsymbol{\Phi}_{\mathrm{had}}^\top \right) \otimes \left(\boldsymbol{P}_{\mathcal{T}_{l_1}^{\mathrm{1d}}} \boldsymbol{\Phi}_{\mathrm{had}}^\top \right), l \xrightleftharpoons{r+1} (l_1 + 1, l_2 + 1),$$

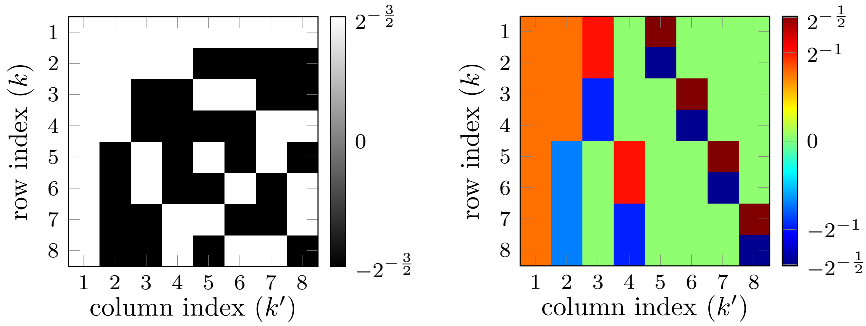

Figure 3.2: An example of the 1-D Hadamard, $(\mathbf{\Phi}_{\text{had}})_{k,k'}$, (left) and DHW, $(\mathbf{\Psi}_{\text{dhw}})_{k,k'}$, (right) matrices with $N = 8$ (or $r = 3$).

and for $l \in [\![r]\!]_0$,

$$\boldsymbol{P}_{\mathcal{T}_l^{\text{iso}}} \mathbf{\Phi}_{2\text{had}}^{\top} = \begin{bmatrix} \boldsymbol{P}_{\overline{\mathcal{T}_l^{1\text{d}} \times \mathcal{T}_{<l}^{1\text{d}}}} (\mathbf{\Phi}_{\text{had}} \otimes \mathbf{\Phi}_{\text{had}}) \\ \boldsymbol{P}_{\overline{\mathcal{T}_l^{1\text{d}} \times \mathcal{T}_l^{1\text{d}}}} (\mathbf{\Phi}_{\text{had}} \otimes \mathbf{\Phi}_{\text{had}}) \\ \boldsymbol{P}_{\overline{\mathcal{T}_{<l}^{1\text{d}} \times \mathcal{T}_l^{1\text{d}}}} (\mathbf{\Phi}_{\text{had}} \otimes \mathbf{\Phi}_{\text{had}}) \end{bmatrix}$$
$$= \begin{bmatrix} (\boldsymbol{P}_{\mathcal{T}_{<l}^{1\text{d}}} \mathbf{\Phi}_{\text{had}}) \otimes (\boldsymbol{P}_{\mathcal{T}_t^{1\text{d}}} \mathbf{\Phi}_{\text{had}}) \\ (\boldsymbol{P}_{\mathcal{T}_l^{1\text{d}}} \mathbf{\Phi}_{\text{had}}) \otimes (\boldsymbol{P}_{\mathcal{T}_l^{1\text{d}}} \mathbf{\Phi}_{\text{had}}) \\ (\boldsymbol{P}_{\mathcal{T}_l^{1\text{d}}} \mathbf{\Phi}_{\text{had}}) \otimes (\boldsymbol{P}_{\mathcal{T}_{<l}^{1\text{d}}} \mathbf{\Phi}_{\text{had}}) \end{bmatrix}.$$

An example of the 1-D Hadamard and DHW matrices, for $N = 8$, is illustrated in Fig. 3.2.

3.3 Uniform and Non-uniform Recovery Guarantees

Equipped with the definitions above, we are now ready to develop our main results. To do so, we need to calculate the local coherence (2.6), multilevel coherence (2.12), and relative sparsity (2.11) for the Hadamard-Haar systems in one and two dimensions. Note that the proofs of this section are all postponed to § 3.5.

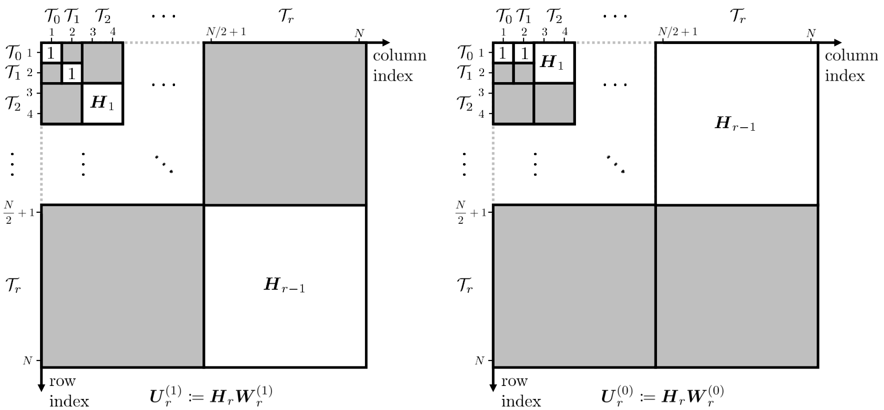

Figure 3.3: Block structure of the matrices $\boldsymbol{H}_r \boldsymbol{W}_r^{(1)}$ (left) and $\boldsymbol{H}_r \boldsymbol{W}_r^{(0)}$ (right) where $\mathcal{T}_l = \mathcal{T}_l^{1\mathrm{d}}$ for $l \in [\![r]\!]_0$. Gray color represents zero value.

We start with the following crucial proposition; it captures a particular recursive block structure of the Hadamard-Haar matrix obtained by multiplying the 1-D Hadamard and Haar matrices.

Proposition 3.4. *Given the integer $r \geq 0$ and defining the Hadamard-Haar matrix $\boldsymbol{U}_r^{(a)} := \boldsymbol{H}_r^\top \boldsymbol{W}_r^{(a)}$ for $a \in \{0,1\}$, we observe that $\boldsymbol{U}_0^{(1)} = \boldsymbol{U}_0^{(0)} = [1]$, and for $r \geq 1$,*

$$\boldsymbol{U}_r^{(1)} = \begin{bmatrix} \boldsymbol{U}_{r-1}^{(1)} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{H}_{r-1} \end{bmatrix}, \qquad \boldsymbol{U}_r^{(0)} = \begin{bmatrix} \boldsymbol{U}_{r-1}^{(0)} & \boldsymbol{H}_{r-1} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}.$$

In particular, the matrix $\boldsymbol{U}_r^{(1)}$ is clearly symmetric, and $\boldsymbol{U}_r^{(1)}$ and $\boldsymbol{U}_r^{(0)}$ contain the structure illustrated in Fig. 3.3.

Proof. See § 3.5.2. □

Remark 3.5. *In the context of Prop. 3.4, from the definition of $\mathcal{T}_l = \mathcal{T}_l^{1\mathrm{d}}$ in* (3.5) *and block structure of $\boldsymbol{U}_r^{(0)}$ and $\boldsymbol{U}_r^{(1)}$ unfolded in Fig. 3.3, we easily deduce the following relations:*

for $t, l \in [\![r]\!]_0$*, we have*

$$\boldsymbol{P}_{\mathcal{T}_t}\boldsymbol{U}_r^{(1)}\boldsymbol{P}_{\mathcal{T}_l}^\top = \begin{cases} \boldsymbol{H}_{(t-1)_+}, & \text{if } t = l, \\ \boldsymbol{0}, & \text{otherwise,} \end{cases} \tag{3.9a}$$

$$\boldsymbol{P}_{\mathcal{T}_{<t}}\boldsymbol{U}_r^{(1)}\boldsymbol{P}_{\mathcal{T}_l}^\top = \begin{cases} \boldsymbol{U}_r^{(1)}\boldsymbol{P}_{\mathcal{T}_l}^\top, & \text{if } t > l, \\ \boldsymbol{0}, & \text{otherwise,} \end{cases} \tag{3.9b}$$

$$\boldsymbol{P}_{\mathcal{T}_t}\boldsymbol{U}_r^{(0)}\boldsymbol{P}_{\mathcal{T}_t}^\top = \begin{cases} 1, & \text{if } t = 0, \\ \boldsymbol{0}, & \text{otherwise,} \end{cases} \tag{3.9c}$$

$$\boldsymbol{P}_{\mathcal{T}_{<t}}\boldsymbol{U}_r^{(0)}\boldsymbol{P}_{\mathcal{T}_t}^\top = \boldsymbol{H}_{t-1}, \text{ for } t > 0, \tag{3.9d}$$

where $(u)_+ \coloneqq \max(u, 0)$.

Noting that $|(\boldsymbol{H}_r)_{k,k'}| = 2^{-r/2}$ for $r \geq 0$ and $k, k' \in [\![2^r]\!]$, Fig. 3.4-top confirms the result in Prop. 3.4 for $N = 8$.

We now focus on the 2-D Hadamard-Haar systems to extract a similar structure.

Proposition 3.6. *Given an integer* $r \geq 0$*, we observe that*
(i) for $t \overset{r+1}{\rightleftharpoons} (t_1 + 1, t_2 + 1)$, $l \overset{r+1}{\rightleftharpoons} (l_1 + 1, l_2 + 1)$, $t_1, t_2, l_1, l_2 \in [\![r]\!]_0$,

$$\boldsymbol{P}_{\mathcal{T}_t^{\mathrm{aniso}}}\boldsymbol{\Phi}_{\mathrm{2had}}^\top\boldsymbol{\Psi}_{\mathrm{adhw}}\boldsymbol{P}_{\mathcal{T}_l^{\mathrm{aniso}}}^\top = \begin{cases} \boldsymbol{H}_{(t_2-1)_+} \otimes \boldsymbol{H}_{(t_1-1)_+}, & \text{if } t_1 = l_1,\ t_2 = l_2, \\ \boldsymbol{0}, & \text{otherwise,} \end{cases} \tag{3.10a}$$

(ii) $\boldsymbol{P}_{\mathcal{T}_0^{\mathrm{iso}}}\boldsymbol{\Phi}_{\mathrm{2had}}^\top\boldsymbol{\Psi}_{\mathrm{idhw}}\boldsymbol{P}_{\mathcal{T}_0^{\mathrm{iso}}}^\top = 1$*, and for* $t, l \in [\![r]\!]_0$ *with* $(t, l) \neq (0, 0)$,

$$\boldsymbol{P}_{\mathcal{T}_t^{\mathrm{iso}}}\boldsymbol{\Phi}_{\mathrm{2had}}^\top\boldsymbol{\Psi}_{\mathrm{idhw}}\boldsymbol{P}_{\mathcal{T}_l^{\mathrm{iso}}}^\top = \begin{cases} \boldsymbol{I}_3 \otimes \big(\boldsymbol{H}_{t-1} \otimes \boldsymbol{H}_{t-1}\big), & \text{if } t = l, \\ \boldsymbol{0}, & \text{otherwise.} \end{cases} \tag{3.10b}$$

Proof. See § 3.5.3. □

In Fig. 3.4-bottom, we depict the structure of the 2-D Hadamard-Haar matrices obtained by multiplying the 2-D Hadamard and Haar matrices.

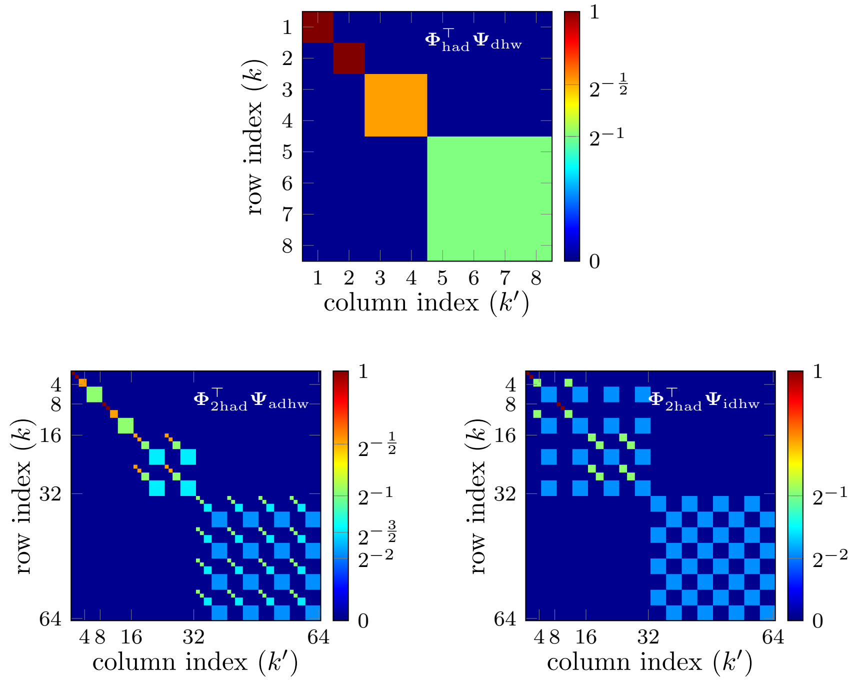

Figure 3.4: Structure of the matrix $|(\Phi_{\mathrm{had}}^\top \Psi_{\mathrm{dhw}})_{k,k'}|$ (top), $|(\Phi_{\mathrm{2had}}^\top \Psi_{\mathrm{adhw}})_{k,k'}|$ (bottom-left), and $|(\Phi_{\mathrm{2had}}^\top \Psi_{\mathrm{idhw}})_{k,k'}|$ (bottom-right) for $N = 8$. We observe that the figure in the middle is the Kronecker product of the matrix on the left with itself. This is actually the consequence of the construction of the 2-D Hadamard matrix and ADHW basis using the Kronecker product.

Prop. 3.6 provides a meaningful expression for those structures. We emphasize that the key aspects in the proof of this proposition is the design of the 2-D isotropic and anisotropic levels, as well as the specific column ordering of the IDHW matrix explained in § 3.2.

The scaling relations in Prop. 3.4 and Prop. 3.6 allow us to determine the local and multilevel coherence of the Hadamard-Haar systems; a result that is at the heart of the proofs of Thm. 3.8 and Thm. 3.10.

Proposition 3.7 (Local coherence of Hadamard-Haar systems)**.** *Given integers $r \geq 1$ and $N = 2^r$, the following equalities hold:*

(i) for the 1-D Hadamard-Haar system: for $l \in [\![N]\!]$,

$$\begin{cases} \mu_l^{\mathrm{loc}}(\mathbf{\Phi}_{\mathrm{had}}^\top \mathbf{\Psi}_{\mathrm{dhw}}) = \min\left(1, 2^{-\frac{\lfloor \log_2(l-1) \rfloor}{2}}\right), \\ \|\boldsymbol{\mu}^{\mathrm{loc}}(\mathbf{\Phi}_{\mathrm{had}}^\top \mathbf{\Psi}_{\mathrm{dhw}})\|^2 = \log_2(N) + 1, \end{cases} \tag{3.11a}$$

(ii) for the 2-D isotropic Hadamard-Haar system: for $l \overset{N}{\rightleftharpoons} (l_1, l_2)$,

$$\begin{cases} \mu_l^{\mathrm{loc}}(\mathbf{\Phi}_{\mathrm{2had}}^\top \mathbf{\Psi}_{\mathrm{idhw}}) = \min\left(1, 2^{-\lfloor \log_2(\max(l_1, l_2)-1) \rfloor}\right), \\ \|\boldsymbol{\mu}^{\mathrm{loc}}(\mathbf{\Phi}_{\mathrm{2had}}^\top \mathbf{\Psi}_{\mathrm{idhw}})\|^2 = 3\log_2(N) + 1, \end{cases} \tag{3.11b}$$

(iii) for the 2-D anisotropic Hadamard-Haar system: for $l \overset{N}{\rightleftharpoons} (l_1, l_2)$,

$$\begin{cases} \mu_l^{\mathrm{loc}}(\mathbf{\Phi}_{\mathrm{2had}}^\top \mathbf{\Psi}_{\mathrm{adhw}}) = \min\left(1, 2^{-\frac{\lfloor \log_2(l_1-1) \rfloor}{2}}\right) \cdot \min\left(1, 2^{-\frac{\lfloor \log_2(l_2-1) \rfloor}{2}}\right), \\ \|\boldsymbol{\mu}^{\mathrm{loc}}(\mathbf{\Phi}_{\mathrm{2had}}^\top \mathbf{\Psi}_{\mathrm{adhw}})\|^2 = \left(\log_2(N) + 1\right)^2. \end{cases} \tag{3.11c}$$

Proof. See § 3.5.4. □

The exact values of the local coherence are illustrated in Fig. 3.5 for $N = 8$. We thus observe that those values are well-controlled in Prop. 3.7, while the global coherence of the Hadamard-Haar systems is equal to one. The local coherence values (3.11b) will be used in § 5.3.2 to design an efficient structured light illumination in an imaging problem. Since the values of the local coherence are closed-form in all the three cases considered in Prop. 3.7, following the argument of Thm. 2.4, we can set the upper bounds κ_l to μ_l^{loc} to characterize the associated systems in the following theorem.

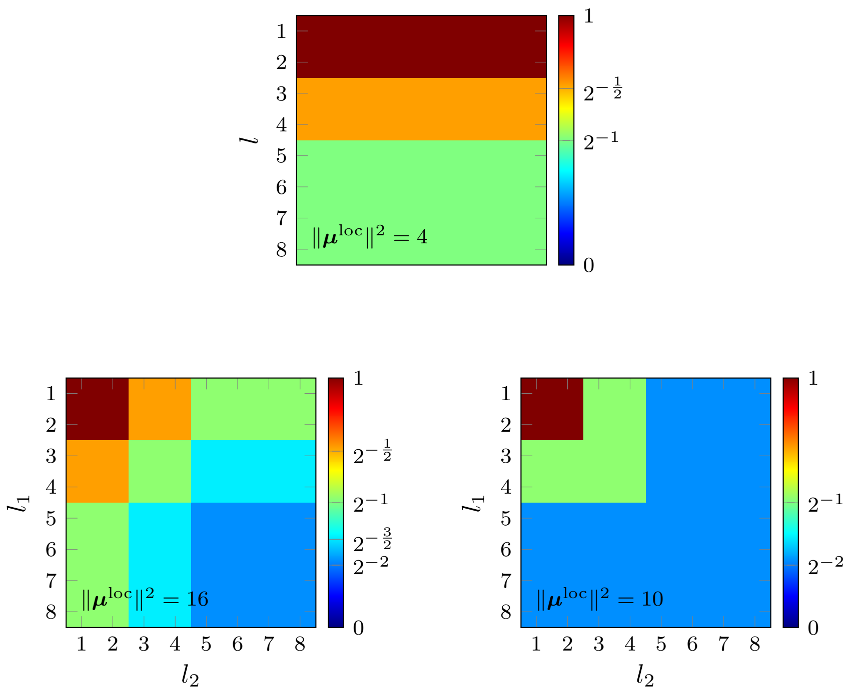

Figure 3.5: The exact local coherence values for $\mu_l^{\mathrm{loc}}(\boldsymbol{\Phi}_{\mathrm{had}}^\top \boldsymbol{\Psi}_{\mathrm{dhw}})$ (top), $\mu_l^{\mathrm{loc}}(\boldsymbol{\Phi}_{\mathrm{2had}}^\top \boldsymbol{\Psi}_{\mathrm{adhw}})$ (bottom-left), and $\mu_l^{\mathrm{loc}}(\boldsymbol{\Phi}_{\mathrm{2had}}^\top \boldsymbol{\Psi}_{\mathrm{idhw}})$ (bottom-right) for $N = 8$, with l_1 and l_2 defined in Prop. 3.7. The values shown here are equal to the estimated values in Prop. 3.7. The block structure of these figures, as represented by the constant color areas fits the definition of the wavelet levels, *i.e.*, with 1-D dyadic (top), 2-D anisotropic (bottom-left), and 2-D isotropic (bottom-right) levels.

Theorem 3.8 (Uniform guarantee for Hadamard-Haar systems)**.** *Fix $N = 2^r$ for some integer $r \in \mathbb{N}$. We provide below, for three Hadamard-Haar systems $(\boldsymbol{\Phi}, \boldsymbol{\Psi})$ in one and two dimensions, the sample-complexity bound and sampling pmf ensuring* (2.7) *and* (2.8) *in Thm. 2.4:*
(i) for the 1-D Hadamard-Haar system:
$\boldsymbol{\Phi} = \boldsymbol{\Phi}_{\mathrm{had}} \in \mathbb{R}^{N \times N}$, $\boldsymbol{\Psi} = \boldsymbol{\Psi}_{\mathrm{dhw}} \in \mathbb{R}^{N \times N}$*, and*

$$\begin{cases} M \gtrsim K \log(N) \log(\epsilon^{-1}), \\ \eta(l) = \frac{\min\left(1, 2^{-\lfloor \log_2(l-1) \rfloor}\right)}{\log_2(N)+1}, \; l \in [\![N]\!], \end{cases} \tag{3.12a}$$

(ii) for the 2-D isotropic Hadamard-Haar system:
$\mathbf{\Phi} = \mathbf{\Phi}_{\text{2had}} \in \mathbb{R}^{N^2 \times N^2}$, $\mathbf{\Psi} = \mathbf{\Psi}_{\text{idhw}} \in \mathbb{R}^{N^2 \times N^2}$, *and*

$$\begin{cases} M \gtrsim K \log(N) \log(\epsilon^{-1}), \\ \eta(l) = \frac{\min\left(1, 2^{-2\lfloor \log_2(\max(l_1, l_2)-1)\rfloor}\right)}{3\log_2(N)+1}, \ l \overset{N}{\rightleftharpoons} (l_1, l_2), \end{cases} \tag{3.12b}$$

(iii) for the 2-D anisotropic Hadamard-Haar system:
$\mathbf{\Phi} = \mathbf{\Phi}_{\text{2had}} \in \mathbb{R}^{N^2 \times N^2}$, $\mathbf{\Psi} = \mathbf{\Psi}_{\text{adhw}} \in \mathbb{R}^{N^2 \times N^2}$, *and*

$$\begin{cases} M \gtrsim K \log^2(N) \log(\epsilon^{-1}), \\ \eta(l) = \frac{\min\left(1, 2^{-\lfloor \log_2(l_1-1)\rfloor}\right) \cdot \min\left(1, 2^{-\lfloor \log_2(l_2-1)\rfloor}\right)}{(\log_2(N)+1)^2}, \ l \overset{N}{\rightleftharpoons} (l_1, l_2). \end{cases} \tag{3.12c}$$

According to this theorem, the optimal sampling pmf $\eta(l)$ is a non-increasing function of l. Since $\eta(l) \propto (\mu_l^{\text{loc}})^2$, (up to a normalization factor $\|\boldsymbol{\mu}^{\text{loc}}\|^2$) the values in Fig. 3.5 indicate the decay behavior of the sampling pmf. In all the Hadamard-Haar systems, the total number of measurements M is on the order of global sparsity K. However, following the computation of the local coherence values in Prop. 3.7, it could be noticed that the use of the UDS strategy gives $M \gtrsim NK \log(\epsilon^{-1}) \log^{\alpha}(N)$ for some $\alpha \in \{1, 2\}$. Moreover, the required number of measurements in (3.12c) is larger than the one in (3.12b) by a $\log(N)$ factor: for those signals that have the same sparsity in IDHW and ADHW bases, *i.e.*, $\sigma_K(\mathbf{\Psi}_{\text{idhw}}^\top \boldsymbol{x})_1 \approx \sigma_K(\mathbf{\Psi}_{\text{adhw}}^\top \boldsymbol{x})_1$, by considering IDHW basis as the sparsity basis we would require smaller number of measurements for signal recovery.

We now turn our attention to the non-uniform guarantee. Following the sample-complexity bounds (2.15) and (2.13), for a fixed signal and fixed sensing and sparsity bases, the efficiency of the MDS scheme relies on *(i)* a suitable partitioning of the sampling and sparsity domains and *(ii)* the ability to estimate the accurate multilevel coherence and relative sparsity values.

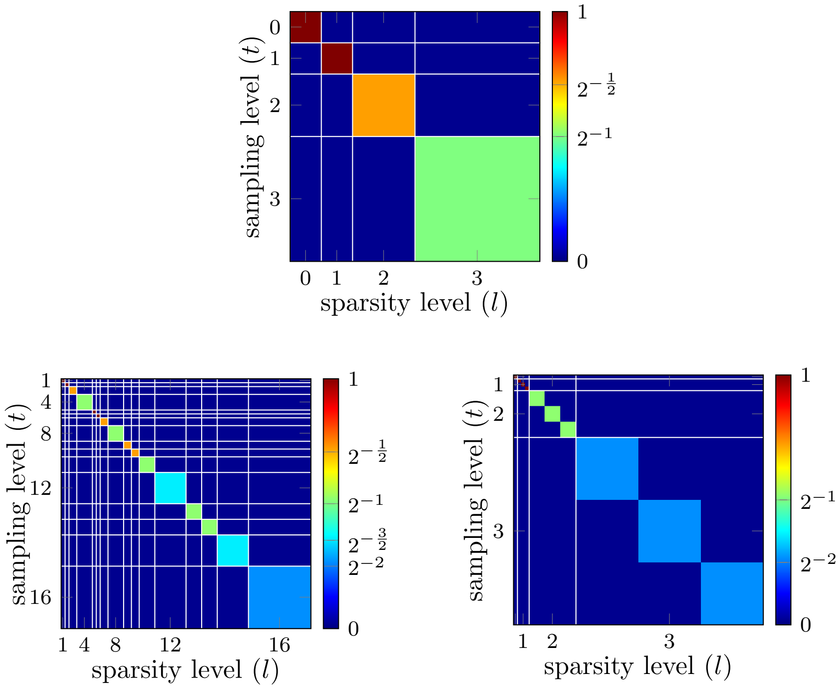

Figure 3.6: Rearrangement of the rows and columns of the matrices shown in Fig. 3.4 with respect to the sampling and sparsity levels. Each white rectangle centered at (t, l) corresponds to the $(t, l)^{\text{th}}$ block, *i.e.*, $\boldsymbol{P}_{\mathcal{T}_t^{\text{1d}}} \boldsymbol{\Phi}_{\text{had}}^\top \boldsymbol{\Psi}_{\text{dhw}} \boldsymbol{P}_{\mathcal{T}_l^{\text{1d}}}^\top$ (top), $\boldsymbol{P}_{\mathcal{T}_t^{\text{aniso}}} \boldsymbol{\Phi}_{\text{2had}}^\top \boldsymbol{\Psi}_{\text{adhw}} \boldsymbol{P}_{\mathcal{T}_l^{\text{aniso}}}^\top$ (bottom-left), and $\boldsymbol{P}_{\mathcal{T}_t^{\text{iso}}} \boldsymbol{\Phi}_{\text{2had}}^\top \boldsymbol{\Psi}_{\text{idhw}} \boldsymbol{P}_{\mathcal{T}_l^{\text{iso}}}^\top$ (bottom-right).

One way to design the sampling and sparsity levels is to leverage the structure of the Hadamard-Haar systems observed in Prop. 3.4 and Prop. 3.6. To visualize those structure, one can properly permute the columns and the rows of the matrices in Fig. 3.4 according to specific wavelet levels, *e.g.*, the 1-D dyadic, 2-D isotropic, or 2-D anisotropic, and obtain the matrices in Fig. 3.6. Each white rectangle in Fig. 3.6 centered at the index (t, l) corresponds to a single partition. Note that the horizontal and vertical axis in Fig. 3.6 denotes the sparsity and sampling level index, respectively; while the axis in Fig. 3.4 represent the column and row indices. The observed structures in Fig. 3.6, specially the ones related to the ADHW and IDHW bases, confirms the statements of Prop. 3.4

and Prop. 3.6. With these structures in mind, we can now compute the following values for multilevel coherence and relative sparsity in different Hadamard-Haar systems.

Proposition 3.9 (Multilevel coherence and relative sparsity values of the Hadamard-Haar systems)**.** *Fix integers r and $N = 2^r$. We consider the levels $\mathcal{T}^{\mathrm{1d}}$, $\mathcal{T}^{\mathrm{iso}}$, and $\mathcal{T}^{\mathrm{aniso}}$ defined above and, for each of them, a vector $\boldsymbol{k}$ whose size equals the number of levels. Then, the following holds:*
(i) for the 1-D Hadamard-Haar system:
for $t, l \in [\![r]\!]_0$,

$$\begin{cases} \mu_{t,l}^{\mathcal{T}^{\mathrm{1d}},\mathcal{T}^{\mathrm{1d}}}(\boldsymbol{\Phi}_{\mathrm{had}}^{\top}\boldsymbol{\Psi}_{\mathrm{dhw}}) = 2^{-(t-1)_+} \cdot \delta_{t,l}, \\ K_t^{\mathcal{T}^{\mathrm{1d}},\mathcal{T}^{\mathrm{1d}}}(\boldsymbol{\Phi}_{\mathrm{had}}^{\top}\boldsymbol{\Psi}_{\mathrm{dhw}}, \boldsymbol{k}) \leq k_t, \end{cases} \tag{3.13a}$$

(ii) for the 2-D isotropic Hadamard-Haar system:
for $t, l \in [\![r]\!]_0$,

$$\begin{cases} \mu_{t,l}^{\mathcal{T}^{\mathrm{iso}},\mathcal{T}^{\mathrm{iso}}}(\boldsymbol{\Phi}_{\mathrm{2had}}^{\top}\boldsymbol{\Psi}_{\mathrm{idhw}}) = 2^{-2(t-1)_+} \cdot \delta_{t,l}, \\ K_t^{\mathcal{T}^{\mathrm{iso}},\mathcal{T}^{\mathrm{iso}}}(\boldsymbol{\Phi}_{\mathrm{2had}}^{\top}\boldsymbol{\Psi}_{\mathrm{idhw}}, \boldsymbol{k}) \leq k_t, \end{cases} \tag{3.13b}$$

(iii) for the 2-D anisotropic Hadamard-Haar system:
for $t \xleftrightarrow{r+1} (t_1+1, t_2+1)$, $l \xleftrightarrow{r+1} (l_1+1, l_2+1)$,

$$\begin{cases} \mu_{t,l}^{\mathcal{T}^{\mathrm{aniso}},\mathcal{T}^{\mathrm{aniso}}}(\boldsymbol{\Phi}_{\mathrm{2had}}^{\top}\boldsymbol{\Psi}_{\mathrm{adhw}}) = 2^{-(t_1-1)_+} \cdot 2^{-(t_2-1)_+} \cdot \delta_{t_1,l_1} \cdot \delta_{t_2,l_2}, \\ K_t^{\mathcal{T}^{\mathrm{aniso}},\mathcal{T}^{\mathrm{aniso}}}(\boldsymbol{\Phi}_{\mathrm{2had}}^{\top}\boldsymbol{\Psi}_{\mathrm{adhw}}, \boldsymbol{k}) \leq k_t, \end{cases} \tag{3.13c}$$

where the relative sparsity $K_t^{\mathcal{W},\mathcal{S}}$ and the multilevel coherence $\mu_{t,l}^{\mathcal{W},\mathcal{S}}$ are defined in (2.11) *and* (2.12)*, respectively, and where $\delta_{t,l}$ is the Kronecker function,* i.e., *$\delta_{k,l} = 1$ if $t = l$ (and zero, otherwise).*

Proof. See § 3.5.5. □

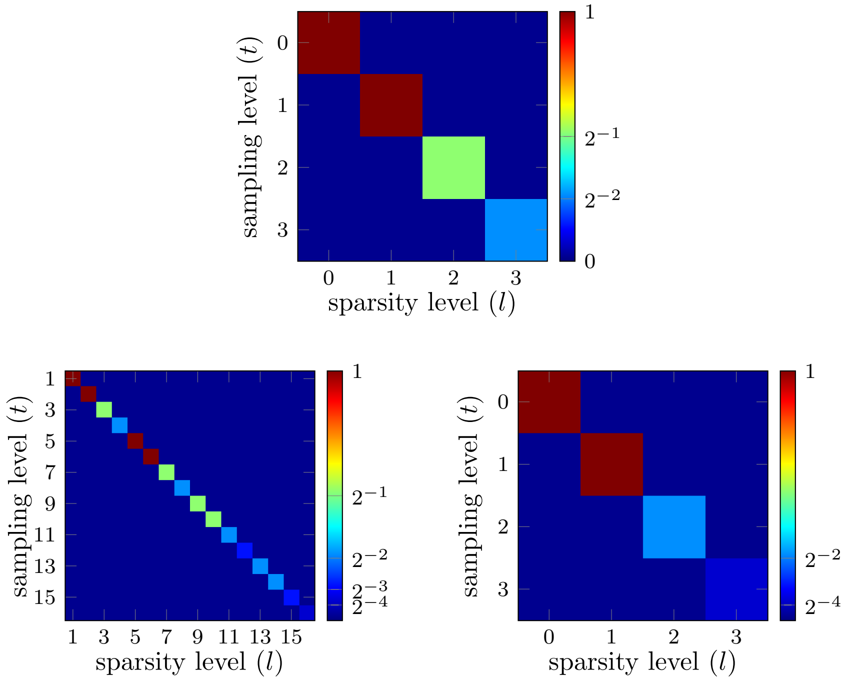

Figure 3.7: The exact multilevel coherence values for Hadamard-Haar systems with $N = 8$: (top) $\mu_{t,l}^{\mathcal{T}^{\rm 1d},\mathcal{T}^{\rm 1d}}(\mathbf{\Phi}_{\rm had}^{\top}\mathbf{\Psi}_{\rm dhw})$, (bottom-left) $\mu_{t,l}^{\mathcal{T}^{\rm iso},\mathcal{T}^{\rm iso}}(\mathbf{\Phi}_{\rm 2had}^{\top}\mathbf{\Psi}_{\rm adhw})$, and (bottom-right) $\mu_{t,l}^{\mathcal{T}^{\rm iso},\mathcal{T}^{\rm iso}}(\mathbf{\Phi}_{\rm 2had}^{\top}\mathbf{\Psi}_{\rm idhw})$. The multilevel coherence values in this figure confirm our estimations in Prop. 3.9.

According to Prop. 3.9, the multilevel coherence of the Hadamard-Haar systems is an exponentially-decreasing function of the level index (see also Fig. 3.7 for an illustration of the multilevel coherence for $N = 8$). Moreover, as an advantage of our sampling and sparsity levels design, the multilevel coherence of the Hadamard-Haar systems at level (t, l) vanishes when $t \neq l$ and thus, the sample-complexity bounds (2.15) and (2.13) become

$$m_t \gtrsim |\mathcal{W}_t|\, \mu_{t,t}^{\mathcal{W},\mathcal{S}}(\mathbf{\Phi}^{\top}\mathbf{\Psi})\, k_t\, \log(K\epsilon^{-1})\, \log(N),$$

$$1 \gtrsim \left(\frac{|\mathcal{W}_l|}{\hat{m}_l} - 1\right) \mu_{l,l}^{\mathcal{W},\mathcal{S}}(\mathbf{\Phi}^{\top}\mathbf{\Psi})\, K_l^{\mathcal{W},\mathcal{S}}(\mathbf{\Phi}^{\top}\mathbf{\Psi}, \boldsymbol{k}).$$

If we ignore the second sample-complexity bound, the first bound relates the number of measurements m_t at level t to the sparsity value k_t at the same level t (and not to the sparsity values at the other levels). This is exactly as one expects when the matrix $\mathbf{\Phi}^\top \mathbf{\Psi}$ is block-diagonal (see [7, § 4.2.1] for more insights) and an application of the sample-complexity bound (2.4) on every block gives the sufficient conditions on the number of measurements.

We are now ready to combine the proposition above with Thm. 2.6 and present the following non-uniform recovery guarantees of Hadamard-Haar systems.

Theorem 3.10 (Non-uniform guarantee for the Hadamard-Haar systems)**.** *Given $N = 2^r$ for some integer $r \in \mathbb{N}$, if we fix*

$$m_t \gtrsim k_t \log(K\epsilon^{-1}) \log(N) \tag{3.14}$$

with either:
(i) for the 1-D Hadamard-Haar system:
$t \in [\![r]\!]_0, \mathbf{\Phi} = \mathbf{\Phi}_{\rm had} \in \mathbb{R}^{N\times N}$,
$\mathbf{\Psi} = \mathbf{\Psi}_{\rm dhw} \in \mathbb{R}^{N\times N}, \mathcal{W} = \mathcal{S} = \mathcal{T}^{\rm 1d}$;
(ii) for the 2-D isotropic Hadamard-Haar system:
$t \in [\![r]\!]_0, \mathbf{\Phi} = \mathbf{\Phi}_{\rm 2had} \in \mathbb{R}^{N^2\times N^2}$,
$\mathbf{\Psi} = \mathbf{\Psi}_{\rm idhw} \in \mathbb{R}^{N^2\times N^2}, \mathcal{W} = \mathcal{S} = \mathcal{T}^{\rm iso}$;
(iii) for the 2-D anisotropic Hadamard-Haar system:
$t \in [\![(r+1)^2]\!], \mathbf{\Phi} = \mathbf{\Phi}_{\rm 2had} \in \mathbb{R}^{N^2\times N^2}$,
$\mathbf{\Psi} = \mathbf{\Psi}_{\rm adhw} \in \mathbb{R}^{N^2\times N^2}, \mathcal{W} = \mathcal{S} = \mathcal{T}^{\rm aniso}$;
then (2.15) *and* (2.13) *in Thm. 2.6 are satisfied.*

Proof. See § 3.5.6. □

It is worth mentioning that Thm. 3.10 provides the tightest sample-complexity bounds, since the multilevel coherence values that lead to these estimates are accurately computed in Prop. 3.9. We observe in Thm. 3.10 that the local number of measurements m_t for the covered Hadamard-Haar systems is on the order of the corresponding local sparsity k_t. A similar observation has recently been made for the infinite-dimensional Hadamard-Haar system in [5, Thm. 4.13]. Unlike the observation in (3.14), for an arbitrary orthonormal wavelet basis the local number of measurements m_t scales as a linear combination of the local sparsities (see in [113] or [20, Thm. 5.8]), which is due to the fact that the Hadamard-wavelet system is not exactly block-diagonal.

Remark 3.11. *One can question how to set the local number of measurements m_t given the local sparsity values k_l and the total number of measurements M. We provide an approach for the 1-D signal recovery problem that is easily extendable to the 2-D cases. The idea here is based on the fact that the local number of measurements in* (3.14) *can be written as $m_t = Ck_t$ for $t \in [\![r]\!]_0$ with $C > 0$ independent of t and k_t. Therefore, the total number of measurements is $M = \sum_t m_t = CK$ where K is the total sparsity value. Therefore, up to a rounding error, the local number of measurements reads*

$$m_t = \frac{M}{K} k_t,\ t \in [\![r]\!]_0.$$

Remark 3.12. *Recall that the recovery guarantee in Thm. 3.10 is non-uniform. The authors in [66] developed a uniform recovery in the context of MDS scheme where the sample-complexity bound writes*

$$m_t \gtrsim |\mathcal{W}_t| \,\big(\sum_{l=1}^{r} \mu(\boldsymbol{P}_{\mathcal{W}_t} \boldsymbol{\Phi}^* \boldsymbol{\Psi} \boldsymbol{P}_{\mathcal{S}_l}^\top)\, k_l\big)\, \big(r\, \log(M)\, \log(N)\, \log^2(K) + \log(\epsilon^{-1})\big), \tag{3.15}$$

compared to the bounds (2.13), (2.14), *and* (2.15). *Following the proof of Prop. 3.9 in § 3.5.5, we observe that for all the three Hadamard-Haar systems covered in Prop. 3.9, $\mu_{t,l}^{\mathcal{W},\mathcal{S}}(\boldsymbol{\Phi}^* \boldsymbol{\Psi}) = \mu(\boldsymbol{P}_{\mathcal{W}_t} \boldsymbol{\Phi}^* \boldsymbol{\Psi} \boldsymbol{P}_{\mathcal{S}_l}^\top)$. Combining this* *with Prop. 3.9, the sample-complexity bound* (3.15) *reads*

$$m_t \gtrsim k_t \left(\log(M) \log^{2+\alpha}(N) \log^2(K) + \log(\epsilon^{-1})\right), \tag{3.16}$$

where $\alpha = 1$ *for the 2-D anisotropic Hadamard-Haar system, and zero otherwise. As one may expect, the sample-complexity bound* (3.16) *associated with the uniform recovery guarantee requires more number of measurements than the bound* (3.14) *associated with the non-uniform guarantee.*

3.4 Numerical results

In this section we carry out several simulations to verify the obtained theoretical results in Thm. 3.8 and Thm. 3.10. In the first set of simulations we address the problem of 1-D signal recovery from subsampled Hadamard measurements and later we focus on the 2-D signal recovery problem, which is associated with single pixel imaging application of CS.

The general setup of the simulations is as follows. Given a ground truth signal $\boldsymbol{x} \in \mathbb{C}^N$ we follow the sensing model (2.1) for some dimensions and sensing bases to be specified later, where we suppose the noise components $n_l \sim_{\text{i.i.d.}} \mathcal{N}(0, \sigma)$ and σ is fixed with respect to the desired SNR in dB defined as

$$\text{SNR} := 20 \log_{10}(\|\boldsymbol{x}\|/(\sigma\sqrt{N})). \tag{3.17}$$

For all the experiments we report the Signal-to-Reconstruction Ratio (SRE) in dB, *i.e.*,

$$\text{SRE} := 20 \log_{10} \mathbb{E}_{\text{e}} \|\boldsymbol{x}\|/\|\boldsymbol{x} - \hat{\boldsymbol{x}}\|, \tag{3.18}$$

where $\mathbb{E}_{\text{e}}$ is the empirical mean over several trials of the sensing context (as specified in the text). In this section the term "VDS" (or "MDS") implies the sampling strategies defined in Thm. 3.8 (resp. Thm. 3.10). For the MDS scheme we respect the approach described in Remark 3.11. We consider two algorithms for signal reconstruction: *(i)* CS reconstruction, that refers to the ℓ_1 minimization problems (2.9) or (2.16) (depending on the recovery guarantee type) for some sparsity basis to be specified later; and *(ii)* Minimal Energy (ME) reconstruction [22], which corresponds to applying the right pseudo-inverse of $\boldsymbol{P}_\Omega \boldsymbol{\Phi}^*$ to the measurement vector. CS reconstructions (2.9) and (2.16) are performed with the Spectral Projected Gradient for ℓ_1 minimization (SPGL1) [117, 118]. In our experiments, the parameter ε in (2.9) (and (2.16)) is set to the oracle value of $\|\boldsymbol{D}\boldsymbol{n}\|$ (respectively, $\|\boldsymbol{n}\|$). Matrices and operators are implemented using the Spot toolbox [4].

The MDS schemes in Thm. 3.10 require to set the values of the local sparsity parameter k_l. For these simulations, when the signal of interest is not exactly sparse we perform the following procedure that is proposed by Adcock *et al.* in [7, Eq. 2.8] and used in [86]: *(i)* given a parameter $\rho \in (0,1]$ and a signal $\boldsymbol{x} \in \mathbb{C}^N$ we first compute the vector of coefficients $\boldsymbol{s} \in \mathbb{C}^N$ in the sparsity basis $\boldsymbol{\Psi}$, *i.e.*, $\boldsymbol{s} = \boldsymbol{\Psi}^\top \boldsymbol{x}$; *(ii)* the effective global sparsity value K is then computed such that after applying the hard thresholding operator $\mathcal{H}_K$ to s, the ratio of the energy that is preserved by K coefficients equals ρ; mathematically,

$$K = K(\rho) = \min\{n : \|\mathcal{H}_n(\boldsymbol{s})\|/\|\boldsymbol{s}\| \geq \rho\},$$

where we set $\rho = 0.995$ in all the experiments here ; *(iii)* we finally compute the effective local sparsity values by simply localizing the number of non-zero coefficients of the hard thresholded signal $\mathcal{H}_K(\boldsymbol{s})$, *i.e.*, for all l

$$k_l = k_l(\rho) = |\mathrm{supp}(\boldsymbol{P}_{\mathcal{S}_l} \mathcal{H}_{K(\rho)}(\boldsymbol{s}))|.$$

Note that this procedure does not sparsify the signal $\boldsymbol{x}$ in the basis $\boldsymbol{\Psi}$, but it is only used to estimate the parameters k_l.

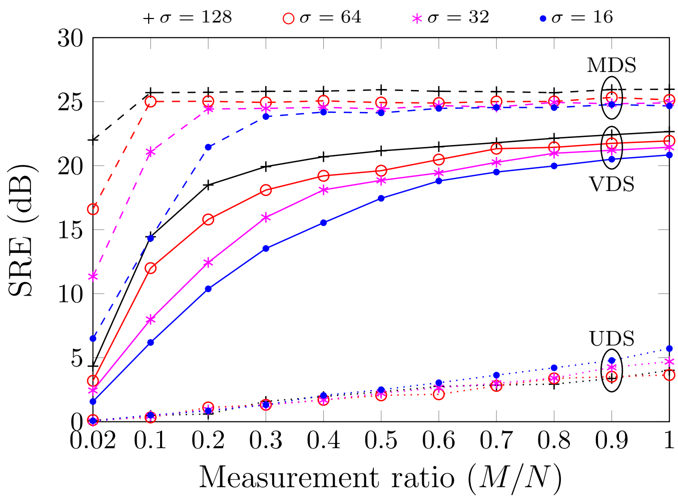

Figure 3.8: The reconstruction performance comparison of the proposed MDS and VDS schemes with the traditional UDS scheme.

3.4.1 Reconstruction Performance on 1-D Signals

We here examine the VDS and MDS schemes defined in Thm. 3.8 and Thm. 3.10 by comparing their SRE values with the one achieved by UDS scheme for different signals. In this part, the sensing and sparsity bases are set to the 1-D Hadamard and DHW bases, respectively, and the signals are recovered via only CS reconstruction. In the first simulation, a Gaussian-shape signal $\boldsymbol{x} \in \mathbb{R}^N$ of size $N = 512$, *i.e.*,

$$x_i = \frac{1}{\sigma\sqrt{2\pi}} \exp\Big(- \frac{(i-i_0)^2}{2\sigma^2}\Big), \ \forall i \in [\![N]\!],$$

is generated as the ground truth. The variables i_0 and σ determine the center and the width of the Gaussian curve. Essentially, by increasing σ the coefficients of the signal in Haar wavelet domain become sparser. The variable $\sigma \in \{16, 32, 64, 128\}$ and the parameter i_0 is generated uniformly at random in the range $[\sigma, N - \sigma]$. We set the variance of the noise to read an SNR of 20 dB. Fig. 3.8 displays the reconstruction quality of the generated signals as a function of the measurement ratio (M/N)

for different values of σ and sampling strategies (UDS, VDS, and MDS). Each point of the curves in Fig. 3.8 is an average of 100 trials (*i.e.*, over random generation of the noise, subsampling set Ω, and parameter i_0).

In the simulations here with the MDS scheme, the effective local sparsities $k_l(\rho)$ are fixed for each value of σ a priori. In particular, given σ we first generate 100 Gaussian-shape signals (different from the ones to be recovered) whose locations i_0 are selected uniformly at random; and then compute their effective local sparsities as prescribed above. Finally, we consider the worst local sparsity values k_l with $l \in [\![r]\!]$ over all 100 trials for designing our MDS scheme. This approach gives a near-optimal MDS strategy, yet it is of practical interest where the true values of the local sparsity are not accessible.

From Fig. 3.8, we can make the following observations: *(i)* by increasing the value of σ the signal becomes sparser in the Haar domain, and thus, all reconstructions yield better SRE values; *(ii)* the UDS scheme yields a poor reconstruction quality; this is aligned with the large value of the global coherence between the Hadamard and Haar bases, which drives the UDS sample-complexity in (2.4); *(iii)* the VDS scheme provides a stable and robust signal recovery (with respect to the change of sparsity and noise level); *(iv)* the SRE of the Hadamard-Haar system is further increased by using the MDS scheme, since it adjusts the sampling strategy to the sparsity structure of the signal; *(v)* although the MDS scheme here is not designed based on the ground truth signal, the dashed lines show significant SRE improvement compared to the VDS strategy.

In Fig. 3.9, we apply similar tests on four other functions, *i.e.*, the "Blocks", "Bumps", "HeaviSine", and "Doppler" signals taken from [33]. These signals display various behaviors, hence allowing us to test our scheme in a broader context. They are generated by evenly sampling the continuous functions specified in [33] over $N = 2048$ samples.

The reconstructed signals from 20% subsampled Hadamard measurements using MDS, VDS, and UDS schemes are displayed in Fig. 3.9.

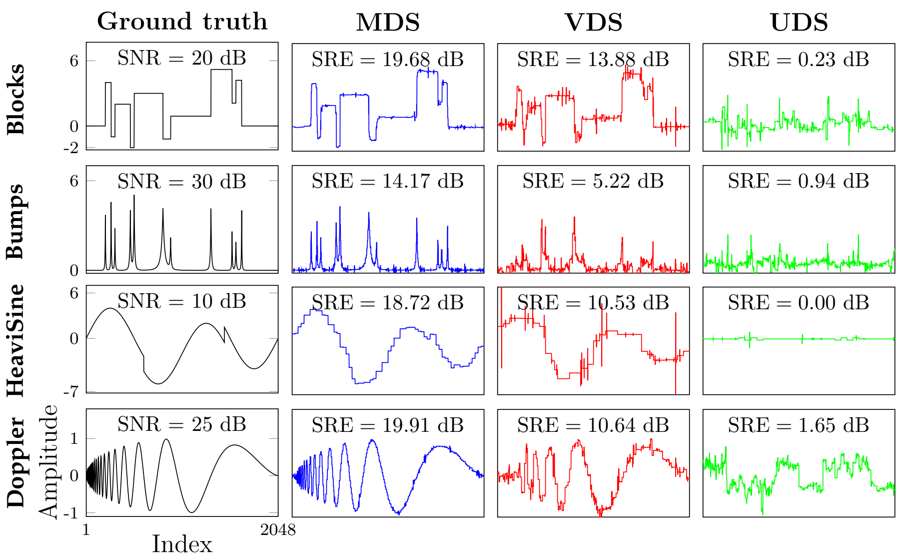

Figure 3.9: Recovering four special 1-D signals from 20% Hadamard measurements.

As can be seen, the UDS strategy does not allow signal recovery. Note that these signals (except the Blocks signal) are not well-compressible in the Haar basis. As a consequence, most reconstructions have blocky artifacts and the VDS scheme does not provide a high quality reconstruction. The MDS scheme, which leverages the local compressibility of the signal, achieves a much higher reconstruction quality in all examples.

3.4.2 Reconstruction Performance on 2-D Signals

We now test the performance of the proposed VDS and MDS schemes in an imaging context. We generate synthetic Shepp-Logan phantom images [103] of size $N \times N$ with $N = 2^r$ and $r \in \{7, \cdots, 11\}$ as the ground truth. The variance of the noise amounts to an SNR of 20 dB. Fig. 3.10 illustrates the SRE values as a function of the measurement ratio (M/N^2) for different resolutions N, sampling strategies (UDS, VDS, and MDS), sparsity bases (IDHW and ADHW), and recovery algorithms (CS and ME). The results are averaged over 10 trials (*i.e.*, over the random

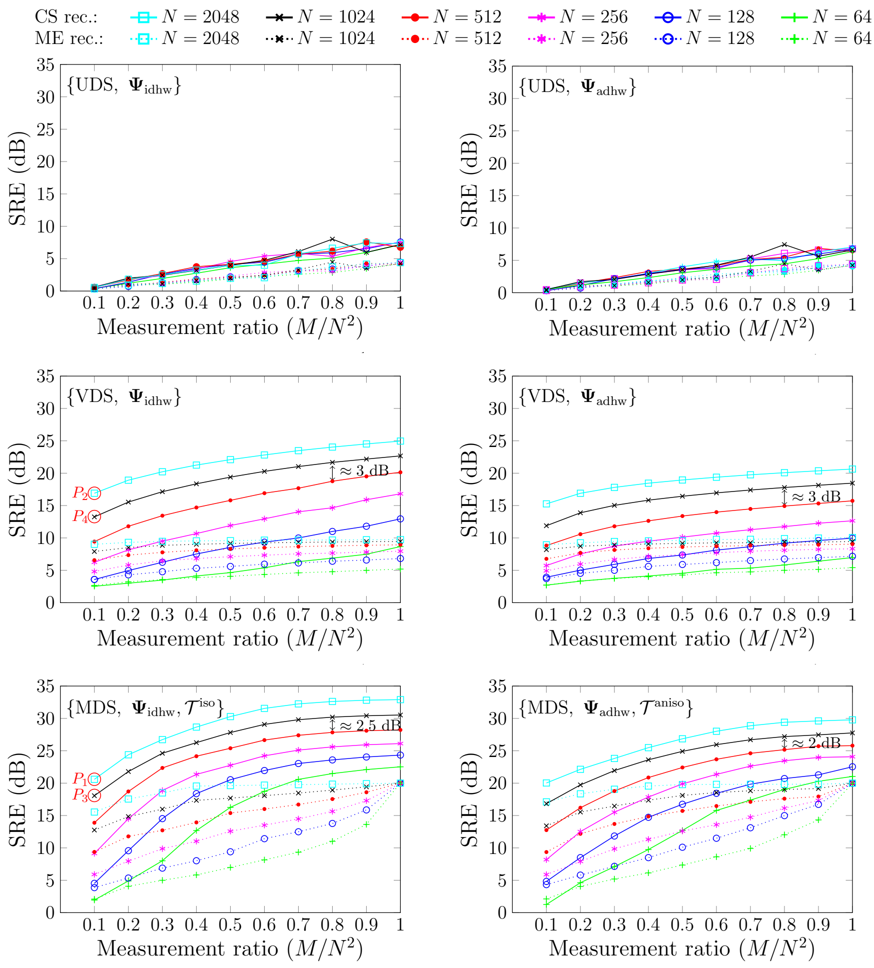

Figure 3.10: The SRE of phantom image recovery from subsampled Hadamard measurements.

generation of both the noise and random selection of the subsampling set Ω according to the sampling strategy). We note that in Fig. 3.10 and in the UDS and VDS cases, since there are repeated indices in the subsampled set Ω, even for $M/N^2 = 1$, we cannot reach the recovery quality of fully-sampled (or Nyquist) Hadamard measurements. On the contrary, since MDS scheme does not allow repeated indices, the recovery quality of the Nyquist Hadamard-Haar system happens when

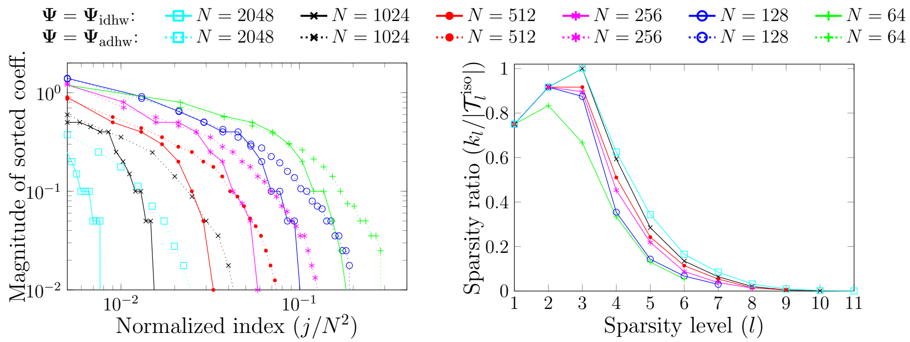

Figure 3.11: Global (left) and local (right) sparsity of the phantom image in 2-D Haar wavelet basis. On the right figure, in order to obtain meaningful curves we assumed $\mathcal{T}_0^{\mathrm{iso}} = \emptyset$ and $\mathcal{T}_1^{\mathrm{iso}} = \{1, 2\}$.

$M/N^2 = 1$. Not surprisingly, ME reconstruction yields SRE = SNR = 20 dB when the Hadamard measurements are fully-sampled. Fig. 3.11 displays the global and local sparsity of the phantom images of different sizes in 2-D Haar wavelet basis. On the left, the sorted coefficients in IDHW and ADHW bases are plotted versus the normalized index axis. Fig. 3.11-right shows an experiment in which we computed the local sparsity ratios for the phantom image of different sizes using IDHW sparsity basis.

From Fig. 3.10 and Fig. 3.11 we can do the following observations. First, similar to the 1-D signal recovery, the UDS scheme performs poorly. Second, the CS reconstruction always outperforms the ME reconstruction, as the latter does not take into account the sparsity prior information. Third, by increasing the resolution of the signal (or the size of the problem) one can obtain a higher SRE value (up to 3 dB), regardless of the CS or ME reconstruction method. Essentially, by going higher in resolution the signal becomes (asymptotically) sparser in the wavelet domain, as represented in Fig. 3.11-left. In this figure, the decay rate of the curves increases as N grows. As already stressed in, *e.g.*, [99], the MDS scheme is thus expected to express its efficacy in high-dimensional applications. Fourth, the IDHW basis yields better SRE

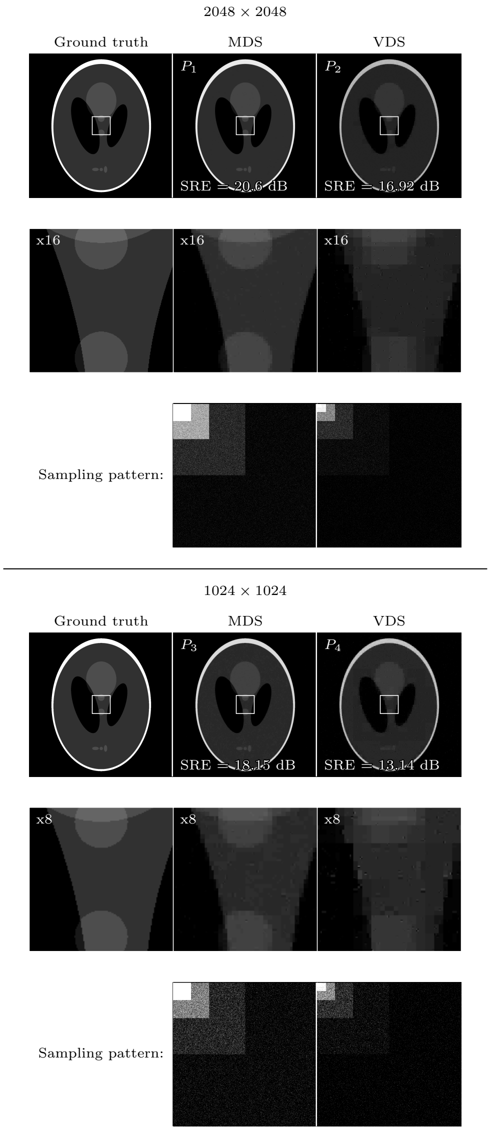

Figure 3.12: An example of the reconstructed images from 10% of the Hadamard measurements. These images correspond to the points P_1, P_2, P_3, and P_4 in Fig. 3.10. Superior quality of the MDS scheme is obvious both visually and quantitatively. We also recall the repetition in the selected indices in VDS scheme which results in less white points in the sampling pattern.

values in comparison to the ADHW basis because the phantom image is more compressible in the IDHW basis. Concretely, by comparing the solid and dotted lines in Fig. 3.11-left, we conclude that the phantom image reaches higher compressibility in the IDHW basis, which further increases the quality of the signal recovery. Fifth, the MDS scheme is resolution dependent: following the sample-complexity bounds in Thm. 3.10, the values in Fig. 3.11-right determine the required number of measurements at each level. Finally, since the MDS scheme leverages the sparsity structure of the signal, it outperforms the VDS scheme in the sense of recovery quality.

An example of the reconstructed images in the simulation above, marked by points P_1, P_2, P_3, and P_4, is depicted in Fig. 3.12. In this figure we notice the effect of the resolution on the MDS strategy and on the image recovery quality.

3.5 Proofs

We now turn our attention to the proofs of the main results. We present first a few auxiliary lemmas used later in this section.

Lemma 3.13. *Let $\boldsymbol{u} \in \mathbb{C}^N$, $\boldsymbol{u}' \in \mathbb{C}^{N'}$, and $\boldsymbol{v} = \boldsymbol{u}' \otimes \boldsymbol{u} \in \mathbb{C}^{\bar{N}}$ with $\bar{N} = NN'$. For two sets $\mathcal{S} \subset [\![N]\!]$ and $\mathcal{S}' \subset [\![N']\!]$, and $\mathcal{S} = \mathcal{S} \times \mathcal{S}'$, we have*

$$\boldsymbol{P}_{\overline{\mathcal{S}}}\boldsymbol{v} = (\boldsymbol{P}_{\mathcal{S}'}\boldsymbol{u}') \otimes (\boldsymbol{P}_{\mathcal{S}}\boldsymbol{u}). \tag{3.19}$$

Proof. Defining $\boldsymbol{e}_i := (\boldsymbol{I}_N)_i$, $\boldsymbol{e}'_j := (\boldsymbol{I}_{N'})_j$, and $\bar{\boldsymbol{e}}_l := (\boldsymbol{I}_{\bar{N}})_l$, we first note that

$$\boldsymbol{u}' \otimes \boldsymbol{u} = \textstyle\sum_{i=1}^{N_1} \sum_{j=1}^{N_2} u_i u'_j \left(\boldsymbol{e}'_j \otimes \boldsymbol{e}_i\right).$$

Therefore,

$$\begin{aligned}
\boldsymbol{P}_{\overline{\mathcal{S}}}\boldsymbol{v} &= \sum_{l\in\overline{\mathcal{S}}} \boldsymbol{v}_l \bar{\boldsymbol{e}}_l = \sum_{(i,j)\in\mathcal{S}} u_i u_j' \left(\boldsymbol{e}_j' \otimes \boldsymbol{e}_i\right) = \sum_{i\in\mathcal{S}} u_i \Big(\sum_{j\in\mathcal{S}'} u_j' \boldsymbol{e}_j' \Big) \otimes \boldsymbol{e}_i \\
&= \sum_{i\in\mathcal{S}} u_i \big((\boldsymbol{P}_{\mathcal{S}'} \boldsymbol{u}_2) \otimes \boldsymbol{e}_i \big) = \big(\boldsymbol{P}_{\mathcal{S}'}\boldsymbol{u}'\big) \otimes \Big(\sum_{i\in\mathcal{S}} u_i \boldsymbol{e}_i \Big) \\
&= \big(\boldsymbol{P}_{\mathcal{S}'}\boldsymbol{u}'\big) \otimes \left(\boldsymbol{P}_{\mathcal{S}}\boldsymbol{u}\right),
\end{aligned}$$

where in the first line we used the fact that $u_l = u_i u_j$ and $\bar{\boldsymbol{e}}_l = \boldsymbol{e}_j' \otimes \boldsymbol{e}_i$ for $l \xrightleftharpoons{N_1,N_2} (i,j)$. □

Lemma 3.14. *For $\boldsymbol{A} \in \mathbb{C}^{M\times N}$ and $\boldsymbol{B} \in \mathbb{C}^{P\times Q}$, we have*

$$\mu(\boldsymbol{A}\otimes\boldsymbol{B}) = \mu(\boldsymbol{A})\cdot\mu(\boldsymbol{B}), \tag{3.20a}$$

$$\mu_l^{\mathrm{loc}}(\boldsymbol{A}\otimes\boldsymbol{B}) = \mu_{l_1}^{\mathrm{loc}}(\boldsymbol{B})\cdot\mu_{l_2}^{\mathrm{loc}}(\boldsymbol{A}), \text{ with } l \xrightleftharpoons{P,M} (l_1,l_2). \tag{3.20b}$$

Proof. From the definition of coherence in (2.5),

$$\mu(\boldsymbol{A}\otimes\boldsymbol{B}) = \max_{i,j} |(\boldsymbol{A}\otimes\boldsymbol{B})_{i,j}| = \max_{i_1,j_1} |a_{i_1,j_1}| \cdot \max_{i_2,j_2} |b_{i_2,j_2}| = \mu(\boldsymbol{A})\cdot\mu(\boldsymbol{B}).$$

For the second relation, following the definition of the local coherence in (2.6), we find

$$\begin{aligned}
\mu_l^{\mathrm{loc}}(\boldsymbol{A}\otimes\boldsymbol{B}) &= \max_{j} \left| \left(\boldsymbol{a}_{l_2}\otimes\boldsymbol{b}_{l_1}\right)_j \right| = \max_{j_1,j_2} |a_{l_2,j_2}\cdot b_{l_1,j_1}| \\
&= \max_{j_2} |a_{l_2,j_2}| \cdot \max_{j_1} |b_{l_1,j_1}|,
\end{aligned}$$

where we used the relation $j \xrightleftharpoons{Q,N} (j_1,j_2)$. □

3.5.1 Proof of Lemma 3.1

Below, to get simpler notation, we write $\mathcal{T}_l$ instead of $\mathcal{T}_l^{\mathrm{1d}}$. We first note from (3.2) and (3.4) that $\boldsymbol{W}^{(a)}\boldsymbol{P}_{\mathcal{T}_0}^{\top} = \mathbf{1}_{2^r}$, since $\boldsymbol{W}_0^{(a)} = [1]$, for

$a \in \{0,1\}$. Since $\mathcal{T}_l \subset \mathcal{T}_{<l+1}$ and $\bar{\boldsymbol{P}}_{\mathcal{T}_{<l+1}} = \boldsymbol{P}_{\mathcal{T}_{<l+1}}^\top \boldsymbol{P}_{\mathcal{T}_{<l+1}}$, we have $\boldsymbol{P}_{\mathcal{T}_l}\bar{\boldsymbol{P}}_{\mathcal{T}_{<l+1}} = \boldsymbol{P}_{\mathcal{T}_l}$ and using Lemma 3.15 proved below we have

$$\boldsymbol{W}_r^{(a)}\boldsymbol{P}_{\mathcal{T}_l}^\top = \boldsymbol{W}_r^{(a)}\bar{\boldsymbol{P}}_{\mathcal{T}_{<l+1}}^\top \boldsymbol{P}_{\mathcal{T}_l}^\top = 2^{\frac{l-r}{2}}\left[\boldsymbol{W}_l^{(a)} \otimes \mathbf{1}_{2^{r-l}}, \mathbf{0}\right]\boldsymbol{P}_{\mathcal{T}_l}^\top. \tag{3.21}$$

Inserting the recursive formulation of $\boldsymbol{W}_r^{(1)}$ and $\boldsymbol{W}_r^{(0)}$ in (3.2) and (3.4), respectively, in (3.21), using $(\boldsymbol{A}\otimes\boldsymbol{B})\otimes\boldsymbol{C} = \boldsymbol{A}\otimes(\boldsymbol{B}\otimes\boldsymbol{C})$ and $[\boldsymbol{A},\boldsymbol{B}]\otimes\boldsymbol{C} = [\boldsymbol{A}\otimes\boldsymbol{C}, \boldsymbol{B}\otimes\boldsymbol{C}]$, and noting that both matrices $\boldsymbol{W}_{l-1}^{(a)}$ and $\boldsymbol{I}_{2^{l-1}}$ have 2^{l-1} columns and the operator $\boldsymbol{P}_{\mathcal{T}_l}^\top$ selects only the columns indexed in $\mathcal{T}_l = \{2^{l-1}+1, \cdots, 2^l\}$, we get

$$\begin{aligned}\boldsymbol{W}_r^{(a)}\boldsymbol{P}_{\mathcal{T}_l}^\top &= 2^{\frac{l-r-1}{2}}\left[\boldsymbol{W}_{l-1}^{(a)} \otimes \mathbf{1}_{2^{r-l+1}}, \boldsymbol{I}_{2^{l-1}} \otimes \begin{bmatrix}\mathbf{1}_{2^{r-l}} \\ (-1)^a\mathbf{1}_{2^{r-l}}\end{bmatrix}, \mathbf{0}\right]\boldsymbol{P}_{\mathcal{T}_l}^\top \\ &= 2^{\frac{l-r-1}{2}}\boldsymbol{I}_{2^{l-1}} \otimes \begin{bmatrix}\mathbf{1}_{2^{r-l}} \\ (-1)^a\mathbf{1}_{2^{r-l}}\end{bmatrix}.\end{aligned} \tag{3.22}$$

By expanding the right-hand side of (3.22), the $(i,j)^{\text{th}}$ component of the matrix $\boldsymbol{W}_r^{(a)}\boldsymbol{P}_{\mathcal{T}_l}^\top$ reads

$$\left(\boldsymbol{W}_r^{(a)}\boldsymbol{P}_{\mathcal{T}_l}^\top\right)_{i,j} = \begin{cases} c^{-1/2}, & \text{if } (j-1)c+1 \le i < (j+\frac{1}{2})c+1, \\ (-1)^a c^{-1/2}, & \text{if } (j+\frac{1}{2})c+1 \le i < (j+1)c+1, \\ 0, & \text{otherwise,}\end{cases} \tag{3.23}$$

where $c = 2^{r-l+1}$. By comparing (3.23) with (3.1) and (3.3), we conclude that $(\boldsymbol{\Psi}_l^{(a)})_{i,j} = \left(\boldsymbol{W}_r\boldsymbol{P}_{\mathcal{T}_l}^\top\right)_{i,j} = h_{l-1,j-1}^{(a)}(i-1)$, which completes the proof.

Lemma 3.15. *For $a \in \{0,1\}$ and $q_l = 2^{r-1}\cdot\sum_{k=0}^{r-l}2^{-k}$,*

$$\boldsymbol{W}_r^{(a)}\bar{\boldsymbol{P}}_{\mathcal{T}_{<l}}^\top = 2^{\frac{l-r-1}{2}}\left[\boldsymbol{W}_{l-1}^{(a)} \otimes \mathbf{1}_{2^{r-l+1}}, \mathbf{0}_{2^r\times q_l}\right].$$

Proof. We prove this lemma by induction over the value of l. From (3.2) or (3.4) and the definition of $\bar{\boldsymbol{P}}_{\mathcal{T}_{<r}}^\top$ one can observe that the base case $\boldsymbol{W}_r^{(a)}\bar{\boldsymbol{P}}_{\mathcal{T}_{<r}}^\top = 2^{\frac{-1}{2}}\left[\boldsymbol{W}_{r-1}^{(a)} \otimes \mathbf{1}_2, \mathbf{0}_{2^r\times 2^{r-1}}\right]$ is true. We now show that if the statement of the lemma holds for $l = j+1$ (induction hypothesis),

then it holds for $l = j$. Since $\mathcal{T}_{<j} \subset \mathcal{T}_{<j+1}$, we have $\bar{\boldsymbol{P}}_{\mathcal{T}_{<j}} \bar{\boldsymbol{P}}_{\mathcal{T}_{<j+1}} = \bar{\boldsymbol{P}}_{\mathcal{T}_{<j}}$, and using the induction hypothesis we can write

$$\boldsymbol{W}_r^{(a)} \bar{\boldsymbol{P}}_{\mathcal{T}_{<j}}^\top = \boldsymbol{W}_r^{(a)} \bar{\boldsymbol{P}}_{\mathcal{T}_{<j+1}}^\top \bar{\boldsymbol{P}}_{\mathcal{T}_{<j}}^\top = 2^{\frac{j-r}{2}} \left[\boldsymbol{W}_j^{(a)} \otimes \mathbf{1}_{2^{r-j}}, \mathbf{0}_{2^r \times q_{j+1}} \right] \bar{\boldsymbol{P}}_{\mathcal{T}_{<j}}^\top. \tag{3.24}$$

By injecting the recursion formula of $\boldsymbol{W}_r^{(0)}$ and $\boldsymbol{W}_r^{(1)}$ from (3.2) and (3.4) (with $r = j$) in (3.24), $(\boldsymbol{A} \otimes \boldsymbol{B}) \otimes \boldsymbol{C} = \boldsymbol{A} \otimes (\boldsymbol{B} \otimes \boldsymbol{C})$ and $[\boldsymbol{A}, \boldsymbol{B}] \otimes \boldsymbol{C} = [\boldsymbol{A} \otimes \boldsymbol{C}, \boldsymbol{B} \otimes \boldsymbol{C}]$, we get

$$\boldsymbol{W}_r^{(a)} \bar{\boldsymbol{P}}_{\mathcal{T}_{<j}}^\top = 2^{\frac{j-r-1}{2}} \left[\boldsymbol{W}_{j-1}^{(a)} \otimes \mathbf{1}_{2^{r-j+1}}, \mathbf{0}_{2^r \times 2^{j-1}}, \mathbf{0}_{2^r \times q_{j+1}} \right], \tag{3.25}$$

since $\mathcal{T}_{<j} = [\![2^{j-1}]\!]$ and $\bar{\boldsymbol{P}}_{\mathcal{T}_{<j}}^\top$ preserves the first 2^{j-1} columns of $\boldsymbol{W}_r^{(a)}$. Noting that $q_{j+1} + 2^{j-1} = q_j$ confirms the statement of the lemma for $n = j$ and thus, completes the proof. $\square$

3.5.2 Proof of Prop. 3.4

From the definitions of the Hadamard and DHW bases in § 3.2, we quickly obtain $\boldsymbol{U}_0^{(1)} = \boldsymbol{H}_0^\top \boldsymbol{W}_0^{(1)} = [1]$ and $\boldsymbol{U}_0^{(0)} = \boldsymbol{H}_0^\top \boldsymbol{W}_0^{(0)} = [1]$. Since $(\boldsymbol{A} \otimes \boldsymbol{B})(\boldsymbol{C} \otimes \boldsymbol{D}) = (\boldsymbol{AC}) \otimes (\boldsymbol{BD})$, we get, for $r \geq 1$,

$$\begin{aligned}
&\boldsymbol{H}_r^\top \boldsymbol{W}_r^{(1)} \\
&\quad = \frac{1}{2} \begin{bmatrix} \left(\boldsymbol{H}_{r-1}^\top \otimes [1 \;\; 1]\right)\left(\boldsymbol{W}_{r-1}^{(1)} \otimes \left[\begin{smallmatrix}1\\1\end{smallmatrix}\right]\right) & \left(\boldsymbol{H}_{r-1}^\top \otimes [1 \;\; 1]\right)\left(\boldsymbol{I}_{2^{r-1}} \otimes \left[\begin{smallmatrix}1\\-1\end{smallmatrix}\right]\right) \\ \left(\boldsymbol{H}_{r-1}^\top \otimes [1 \; -1]\right)\left(\boldsymbol{W}_{r-1}^{(1)} \otimes \left[\begin{smallmatrix}1\\1\end{smallmatrix}\right]\right) & \left(\boldsymbol{H}_{r-1}^\top \otimes [1 \; -1]\right)\left(\boldsymbol{I}_{2^{r-1}} \otimes \left[\begin{smallmatrix}1\\-1\end{smallmatrix}\right]\right) \end{bmatrix} \\
&\quad = \frac{1}{2} \begin{bmatrix} \boldsymbol{H}_{r-1}^\top \boldsymbol{W}_{r-1}^{(1)} \otimes [2] & \boldsymbol{H}_{r-1}^\top \boldsymbol{I}_{2^{r-1}} \otimes [0] \\ \boldsymbol{H}_{r-1}^\top \boldsymbol{W}_{r-1}^{(1)} \otimes [0] & \boldsymbol{H}_{r-1}^\top \boldsymbol{I}_{2^{r-1}} \otimes [2] \end{bmatrix} \\
&\quad = \begin{bmatrix} \boldsymbol{H}_{r-1} \boldsymbol{W}_{r-1}^{(1)} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{H}_{r-1} \end{bmatrix}.
\end{aligned}$$

Similarly, we can write

$$\begin{aligned}
&\boldsymbol{H}_r^\top \boldsymbol{W}_r^{(0)} \\
&\quad = \frac{1}{2}\begin{bmatrix} \left(\boldsymbol{H}_{r-1}^\top \otimes [1 \quad 1]\right)\left(\boldsymbol{W}_{r-1}^{(0)} \otimes \left[\begin{smallmatrix}1\\1\end{smallmatrix}\right]\right) & \left(\boldsymbol{H}_{r-1}^\top \otimes [1 \quad 1]\right)\left(\boldsymbol{I}_{2^{r-1}} \otimes \left[\begin{smallmatrix}1\\1\end{smallmatrix}\right]\right) \\ \left(\boldsymbol{H}_{r-1}^\top \otimes [1 \ -1]\right)\left(\boldsymbol{W}_{r-1}^{(0)} \otimes \left[\begin{smallmatrix}1\\1\end{smallmatrix}\right]\right) & \left(\boldsymbol{H}_{r-1}^\top \otimes [1 \ -1]\right)\left(\boldsymbol{I}_{2^{r-1}} \otimes \left[\begin{smallmatrix}1\\1\end{smallmatrix}\right]\right) \end{bmatrix} \\
&\quad = \frac{1}{2}\begin{bmatrix} \boldsymbol{H}_{r-1}^\top \boldsymbol{W}_{r-1}^{(0)} \otimes [2] & \boldsymbol{H}_{r-1}^\top \boldsymbol{I}_{2^{r-1}} \otimes [2] \\ \boldsymbol{H}_{r-1}^\top \boldsymbol{W}_{r-1}^{(0)} \otimes [0] & \boldsymbol{H}_{r-1}^\top \boldsymbol{I}_{2^{r-1}} \otimes [0] \end{bmatrix}.
\end{aligned}$$

By recursion, and from the definition of the 1-D dyadic levels $\mathcal{T}^{\mathrm{1d}}$ we then get the structure described in Fig. 3.3. Moreover, from Fig. 3.3-left and using the fact that $\boldsymbol{H}_r$ is symmetric, we conclude that $\boldsymbol{U}_r^{(1)}$ is symmetric as well.

3.5.3 Proof of Prop. 3.6

To get simpler notation, we write $\mathcal{T}_l$ for $\mathcal{T}_l^{\mathrm{1d}}$ and define, for $a \in \{0,1\}$, $\boldsymbol{U}^{(a)} := \boldsymbol{H}_r \boldsymbol{W}_r^{(a)}$, $\boldsymbol{U}_{(<t,l)}^{(a)} := \boldsymbol{P}_{\mathcal{T}_{<t}} \boldsymbol{U}^{(a)} \boldsymbol{P}_{\mathcal{T}_l}^\top$, $\boldsymbol{U}_{(t,l)}^{(a)} := \boldsymbol{P}_{\mathcal{T}_t} \boldsymbol{U}^{(a)} \boldsymbol{P}_{\mathcal{T}_l}^\top$.

To prove the first part of the proposition, from Remark 3.2 and Remark 3.3 we can write, for $t \xleftrightarrow{r+1} (t_1+1, t_2+1)$ and $l \xleftrightarrow{r+1} (l_1+1, l_2+1)$,

$$\boldsymbol{P}_{\mathcal{T}_t^{\mathrm{aniso}}} \boldsymbol{\Phi}_{\mathrm{2had}}^\top \boldsymbol{\Psi}_{\mathrm{adhw}} \boldsymbol{P}_{\mathcal{T}_l^{\mathrm{aniso}}}^\top = \boldsymbol{U}_{t_2,l_2}^{(1)} \otimes \boldsymbol{U}_{t_1,l_1}^{(1)} \in \mathbb{R}^{2^{t_1+t_2-2} \times 2^{l_1+l_2-2}}.$$

This matrix, using (3.9a), is $\boldsymbol{H}_{(t_2-1)_+} \otimes \boldsymbol{H}_{(t_1-1)_+}$, if $t_1 = l_1$ and $t_2 = l_2$ (and $\mathbf{0}$ otherwise).

We now prove the second part of the proposition. Recall from the definition of the IDHW basis and the 2-D isotropic wavelet levels in § 3.2 that

$$\boldsymbol{\Psi}_{\mathrm{idhw}} \boldsymbol{P}_{\mathcal{T}_l^{\mathrm{iso}}}^\top = \begin{bmatrix} \left(\left(\boldsymbol{W}_r^{(0)} \boldsymbol{P}_{\mathcal{T}_l}^\top\right) \otimes \left(\boldsymbol{W}_r^{(1)} \boldsymbol{P}_{\mathcal{T}_l}^\top\right)\right)^\top \\ \left(\left(\boldsymbol{W}_r^{(1)} \boldsymbol{P}_{\mathcal{T}_l}^\top\right) \otimes \left(\boldsymbol{W}_r^{(1)} \boldsymbol{P}_{\mathcal{T}_l}^\top\right)\right)^\top \\ \left(\left(\boldsymbol{W}_r^{(1)} \boldsymbol{P}_{\mathcal{T}_l}^\top\right) \otimes \left(\boldsymbol{W}_r^{(0)} \boldsymbol{P}_{\mathcal{T}_l}^\top\right)\right)^\top \end{bmatrix}^\top, \tag{3.26}$$

for $l \in [\![r]\!]$. Define $\boldsymbol{V}^{(t,l)} := \boldsymbol{P}_{\mathcal{T}_t^{\mathrm{iso}}} \boldsymbol{\Phi}_{\mathrm{2had}}^\top \boldsymbol{\Psi}_{\mathrm{idhw}} \boldsymbol{P}_{\mathcal{T}_l^{\mathrm{iso}}}^\top$. For the proof we need to compute $\boldsymbol{V}^{(t,l)}$ for $t, l \in [\![r]\!]_0$. First, we assume that $t, l \in [\![r]\!]$.

From Remark 3.3 and (3.26) we have

$$\boldsymbol{V}^{(t,l)} = \begin{bmatrix} \boldsymbol{U}^{(0)}_{(<t,l)} \otimes \boldsymbol{U}^{(1)}_{(t,l)} & \boldsymbol{U}^{(1)}_{(<t,l)} \otimes \boldsymbol{U}^{(1)}_{(t,l)} & \boldsymbol{U}^{(1)}_{(<t,l)} \otimes \boldsymbol{U}^{(0)}_{(t,l)} \\ \boldsymbol{U}^{(0)}_{(t,l)} \otimes \boldsymbol{U}^{(1)}_{(t,l)} & \boldsymbol{U}^{(0)}_{(t,l)} \otimes \boldsymbol{U}^{(1)}_{(t,l)} & \boldsymbol{U}^{(1)}_{(t,l)} \otimes \boldsymbol{U}^{(0)}_{(t,l)} \\ \boldsymbol{U}^{(0)}_{(t,l)} \otimes \boldsymbol{U}^{(1)}_{(<t,l)} & \boldsymbol{U}^{(0)}_{(t,l)} \otimes \boldsymbol{U}^{(1)}_{(<t,l)} & \boldsymbol{U}^{(1)}_{(t,l)} \otimes \boldsymbol{U}^{(0)}_{(<t,l)} \end{bmatrix}.$$

From Remark 3.5 (with an attention to the conditions on the right-hand side of the relations) we observe that the diagonal blocks in $\boldsymbol{V}^{(t,l)}$ are equal to $\boldsymbol{H}_{t-1} \otimes \boldsymbol{H}_{t-1}$ if $t = l$ and $\boldsymbol{0}$ otherwise. Therefore, if $t = l$,

$$\begin{aligned} \boldsymbol{V}^{(t,l)} &= \begin{bmatrix} \boldsymbol{H}_{t-1} \otimes \boldsymbol{H}_{t-1} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{H}_{t-1} \otimes \boldsymbol{H}_{t-1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{H}_{t-1} \otimes \boldsymbol{H}_{t-1} \end{bmatrix} \\ &= \boldsymbol{I}_3 \otimes \big(\boldsymbol{H}_{t-1} \otimes \boldsymbol{H}_{t-1}\big), \end{aligned}$$

while $\boldsymbol{V}^{(t,l)} = \boldsymbol{0}$ if $t \neq l$. Second, we compute $\boldsymbol{V}^{(t,l)}$ for $t \in [\![r]\!]_0$ and $l = 0$. Since $\boldsymbol{\Psi}_{\mathrm{idhw}} \boldsymbol{P}^\top_{\mathcal{T}_0^{\mathrm{iso}}} = \Big(\boldsymbol{W}_r^{(0)} \boldsymbol{P}^\top_{\mathcal{T}_0}\Big) \otimes \Big(\boldsymbol{W}_r^{(0)} \boldsymbol{P}^\top_{\mathcal{T}_0}\Big)$, $\boldsymbol{U}^{(0)} \boldsymbol{P}^\top_{\mathcal{T}_0} = \boldsymbol{I}_{2^r} \boldsymbol{P}^\top_{\{1\}}$, and from Remark 3.3, we have

$$\boldsymbol{V}^{(t,0)} = \begin{bmatrix} \boldsymbol{U}^{(0)}_{(<t,0)} \otimes \boldsymbol{U}^{(0)}_{(t,0)} \\ \boldsymbol{U}^{(0)}_{(t,0)} \otimes \boldsymbol{U}^{(0)}_{(t,0)} \\ \boldsymbol{U}^{(0)}_{(t,0)} \otimes \boldsymbol{U}^{(0)}_{(<t,0)} \end{bmatrix} = \begin{bmatrix} \boldsymbol{P}_{\overline{\mathcal{T}_t \times \mathcal{T}_{<t}}} \boldsymbol{I}_{2^{2r}} \boldsymbol{P}^\top_{\{1\}} \\ \boldsymbol{P}_{\overline{\mathcal{T}_t \times \mathcal{T}_t}} \boldsymbol{I}_{2^{2r}} \boldsymbol{P}^\top_{\{1\}} \\ \boldsymbol{P}_{\overline{\mathcal{T}_{<t} \times \mathcal{T}_t}} \boldsymbol{I}_{2^{2r}} \boldsymbol{P}^\top_{\{1\}} \end{bmatrix} = \boldsymbol{P}_{\mathcal{T}_t^{\mathrm{iso}}} [1, \boldsymbol{0}]^\top.$$

Finally, we need to compute $\boldsymbol{V}^{(t,l)}$ for $t = 0$ and $l \in [\![r]\!]$, *i.e.*,

$$\boldsymbol{V}^{(0,l)} = \Big[\boldsymbol{U}^{(0)}_{(0,l)} \otimes \boldsymbol{U}^{(1)}_{(0,l)} \quad \boldsymbol{U}^{(1)}_{(0,l)} \otimes \boldsymbol{U}^{(1)}_{(0,l)} \quad \boldsymbol{U}^{(1)}_{(0,l)} \otimes \boldsymbol{U}^{(0)}_{(0,l)}\Big].$$

Using (3.9a) and (3.9c) with $t = 0$ and $l \in [\![r]\!]$ yields $\boldsymbol{V}^{(0,l)} = \boldsymbol{0}$. This completes the proof.

3.5.4 Proof of Prop. 3.7

In this proof we write $\mathcal{T}_l$ for $\mathcal{T}_l^{\mathrm{1d}}$. Recall that $\mu(\boldsymbol{H}_r) = 2^{-r/2}$, and for any $k > 1$, $k \in \mathcal{T}_{\bar{l}(k)}$ with $\bar{l}(k) := \lfloor \log_2(k-1) \rfloor + 1$, since $\mathcal{T}_l = [\![2^l]\!] \backslash [\![2^{l-1}]\!]$, for

$l \geq 1$. We first observe that $\mu_1^{\mathrm{loc}}(\boldsymbol{U}_r^{(1)}) = 1$, since $(\boldsymbol{H}_r)_{1,i} = (\boldsymbol{W}_r^{(1)})_{1,i} = 2^{-r/2}$ for all $i \in [\![2^r]\!]$.

To prove (3.11a), note that, for $k > 1$, since $\bar{\boldsymbol{P}}_\Omega = \boldsymbol{P}_\Omega^\top \boldsymbol{P}_\Omega$, for any subset Ω, and $\boldsymbol{P}_{\{k\}} \bar{\boldsymbol{P}}_{\mathcal{T}_{\bar{l}(k)}} = \boldsymbol{P}_{\{k\}}$, $|(\boldsymbol{H}_r)_{i,j}| = 2^{-r/2}$ for all $i, j \in [\![2^r]\!]$ and using (3.9a),

$$\begin{aligned}\mu_k^{\mathrm{loc}}(\boldsymbol{U}_{2^r}^{(1)}) &= \mu(\boldsymbol{P}_{\{k\}} \bar{\boldsymbol{P}}_{\mathcal{T}_{\bar{l}(k)}} \boldsymbol{U}_{2^r}^{(1)}) = \mu(\boldsymbol{P}_{\{k\}} \boldsymbol{H}_{\bar{l}(k)-1}) \\ &= 2^{-\frac{\bar{l}(k)-1}{2}} = 2^{-\frac{\lfloor \log_2(k-1) \rfloor}{2}}.\end{aligned}$$

In addition,

$$\|\boldsymbol{\mu}^{\mathrm{loc}}(\boldsymbol{U}_r^{(1)})\|_2^2 = 1 + \sum_{k=2}^{N} 2^{-\lfloor \log_2(k-1) \rfloor} = 1 + \sum_{l=0}^{r-1} 2^l \cdot 2^{-l} = \log_2(N) + 1.$$

Next, to prove (3.11b), we first note that $\mu_1^{\mathrm{loc}}(\boldsymbol{\Phi}_{\mathrm{2had}}^\top \boldsymbol{\Psi}_{\mathrm{idhw}}) = 1$, since $(\boldsymbol{\Phi}_{\mathrm{2had}})_{1,i} = (\boldsymbol{\Psi}_{\mathrm{idhw}})_{1,i} = 2^{-r}$ for all $i \in [\![2^{2r}]\!]$. Consider the rule $k \overset{N}{\rightleftharpoons} (k_1, k_2)$. Using (3.10b) and (3.20b), for $1 < k \in \mathcal{T}_t^{\mathrm{iso}}$,

$$\mu_k^{\mathrm{loc}}(\boldsymbol{\Phi}_{\mathrm{2had}}^\top \boldsymbol{\Psi}_{\mathrm{idhw}}) = \mu_{k_1}^{\mathrm{loc}}(\boldsymbol{H}_{(t-1)}) \cdot \mu_{k_2}^{\mathrm{loc}}(\boldsymbol{H}_{(t-1)}) = 2^{-(t-1)}. \tag{3.27}$$

Moreover, we find

$$k \in \mathcal{T}_t^{\mathrm{iso}} \quad \Leftrightarrow \quad \max(k_1, k_2) \in \mathcal{T}_t \quad \Leftrightarrow \quad t - 1 = \lfloor \log_2(\max(k_1, k_2) - 1) \rfloor. \tag{3.28}$$

Combining (3.27) and (3.28) implies the local coherence relation in (3.11b).

Moreover, since $|\mathcal{T}_t^{\mathrm{iso}}| = 3 \cdot 2^{2(t-1)}$ for $t \in [\![r]\!]$, (3.27) provides

$$\begin{aligned}\|\boldsymbol{\mu}^{\mathrm{loc}}(\boldsymbol{\Phi}_{\mathrm{2had}}^\top \boldsymbol{\Psi}_{\mathrm{idhw}})\|_2^2 &= 1 + \sum_{t=1}^{r} \sum_{k \in \mathcal{T}_t^{\mathrm{iso}}} \mu_k^{\mathrm{loc}}(\boldsymbol{\Phi}_{\mathrm{2had}}^\top \boldsymbol{\Psi}_{\mathrm{idhw}})^2 \\ &= 1 + \sum_{t=1}^{r} |\mathcal{T}_t^{\mathrm{iso}}| \cdot 2^{-2(t-1)} = 1 + 3 \cdot r.\end{aligned}$$

Finally, to prove (3.11c), we first observe that $\mu_1^{\mathrm{loc}}(\boldsymbol{\Phi}_{2\mathrm{had}}^\top \boldsymbol{\Psi}_{\mathrm{adhw}}) = 1$, since $(\boldsymbol{\Phi}_{2\mathrm{had}})_{1,i} = (\boldsymbol{\Psi}_{\mathrm{adhw}})_{1,i} = 2^{-r}$ for all $i \in [\![2^{2r}]\!]$. Consider the rules $t \overset{r+1}{\Longleftarrow\!\!\Rightarrow} (t_1+1, t_2+1)$ and $k \overset{N}{\Longleftarrow\!\!\Rightarrow} (k_1, k_2)$. From the construction of the 2-D anisotropic levels we have, for $k > 1$,

$$k \in \mathcal{T}_t^{\mathrm{aniso}} \quad \Leftrightarrow \quad k_1 \in \mathcal{T}_{t_1},\ k_2 \in \mathcal{T}_{t_2} \quad \Leftrightarrow \quad \begin{cases} t_1 - 1 = \lfloor \log_2(k_1 - 1) \rfloor, \\ t_2 - 1 = \lfloor \log_2(k_2 - 1) \rfloor. \end{cases} \tag{3.29}$$

Using (3.10a) and (3.20b), for $1 < k \in \mathcal{T}_t^{\mathrm{aniso}}$, we get

$$\begin{aligned} \mu_k^{\mathrm{loc}}(\boldsymbol{\Phi}_{2\mathrm{had}}^\top \boldsymbol{\Psi}_{\mathrm{adhw}}) &= \mu_k^{\mathrm{loc}}(\boldsymbol{H}_{t_2-1} \otimes \boldsymbol{H}_{t_1-1}) \\ &= \mu_{k_1}^{\mathrm{loc}}(\boldsymbol{H}_{t_1-1}) \cdot \mu_{k_2}^{\mathrm{loc}}(\boldsymbol{H}_{t_2-1}). \end{aligned} \tag{3.30}$$

Combining (3.29) and (3.30) with the relation in 3.11a implies the local coherence value in (3.11c).

In addition, using (3.11a),

$$\begin{aligned} \|\boldsymbol{\mu}^{\mathrm{loc}}(\boldsymbol{\Phi}_{2\mathrm{had}}^\top \boldsymbol{\Psi}_{\mathrm{adhw}})\|_2^2 &= \big(1 + \sum_{k_1=2}^{N} 2^{-\lfloor \log_2(k_1-1) \rfloor}\big) \cdot \big(1 + \sum_{k_2=2}^{N} 2^{-\lfloor \log_2(k_2-1) \rfloor}\big) \\ &= (\log_2(N) + 1)^2. \end{aligned}$$

3.5.5 Proof of Prop. 3.9

Given $\boldsymbol{U}_r^{(1)} = \boldsymbol{H}_r \boldsymbol{W}_r^{(1)}$, and $\mathcal{T}_l = \mathcal{T}_l^{1\mathrm{d}}$ for $l \in [\![r]\!]_0$, we note that $\mu(\boldsymbol{H}_r) = 2^{-r/2}$, $\mu(\boldsymbol{P}_{\mathcal{W}_t}\boldsymbol{A}) = \max_l \mu(\boldsymbol{P}_{\mathcal{W}_t}\boldsymbol{A}\boldsymbol{P}_{\mathcal{S}_l}^\top)$, and for any orthonormal matrix $\boldsymbol{\Phi}$, $\|\boldsymbol{\Phi}\|_{2,2} = \max_{\|\boldsymbol{v}\|_2=1} \|\boldsymbol{\Phi}\boldsymbol{v}\|_2 = \|\boldsymbol{v}\|_2 = 1$.

We first prove (3.13a). From Remark 3.5, note that, for $t, l \in [\![r]\!]$,

$$\mu(\boldsymbol{P}_{\mathcal{T}_t}\boldsymbol{U}_r^{(1)}\boldsymbol{P}_{\mathcal{T}_l}^\top) = \mu(\boldsymbol{H}_{t-1}) \cdot \delta_{t,l} = 2^{-(t-1)/2} \cdot \delta_{t,l}, \tag{3.31}$$

and $\mu(\boldsymbol{P}_{\mathcal{T}_t}\boldsymbol{U}_r^{(1)}\boldsymbol{P}_{\mathcal{T}_0}^\top) = \delta_{t,0}$. Therefore, for $t, l \in [\![r]\!]_0$, $\mu(\boldsymbol{P}_{\mathcal{T}_t}\boldsymbol{U}_r^{(1)}) = 2^{-\frac{(t-1)_+}{2}}$ and

$$\mu_{t,l}^{\mathcal{T},\mathcal{T}}(\boldsymbol{U}^{(1)}) = 2^{-(t-1)_+} \cdot \delta_{t,l}.$$

To compute the relative sparsity, from Lemma 2.5 and Remark 3.5, and since $\|\boldsymbol{H}_r\|_{2,2} = 1$, we find

$$
\begin{aligned}
K_t^{\mathcal{T},\mathcal{T}}(\boldsymbol{U}^{(1)}, \boldsymbol{k})^{1/2} &\le \sum_{l=0}^{r} \|\boldsymbol{P}_{\mathcal{T}_t}\boldsymbol{U}^{(1)}\boldsymbol{P}_{\mathcal{T}_l}^\top\|_{2,2}\sqrt{k_l} \\
&= \|\boldsymbol{H}_{(t-1)_+}\|_{2,2}\sqrt{k_t} = \sqrt{k_t}.
\end{aligned}
$$

To prove (3.13b), note from (3.10b) that, for $l \in [\![r]\!]$,

$$
\begin{aligned}
\mu(\boldsymbol{P}_{\mathcal{T}_t^{\mathrm{iso}}}\boldsymbol{\Phi}_{\mathrm{2had}}^\top\boldsymbol{\Psi}_{\mathrm{idhw}}\boldsymbol{P}_{\mathcal{T}_l^{\mathrm{iso}}}^\top) &= \mu(\boldsymbol{I}_3)\cdot\mu(\boldsymbol{H}_{(t-1)})\cdot\mu(\boldsymbol{H}_{(t-1)})\cdot\delta_{t,l} \\
&= 2^{-(t-1)}\cdot\delta_{t,l},
\end{aligned}
$$

where we used the rule in (3.20a), and $\mu(\boldsymbol{P}_{\mathcal{T}_t^{\mathrm{iso}}}\boldsymbol{\Phi}_{\mathrm{2had}}^\top\boldsymbol{\Psi}_{\mathrm{idhw}}\boldsymbol{P}_{\mathcal{T}_0^{\mathrm{iso}}}^\top) = \delta_{t,0}$. Therefore, for $t,l \in [\![r]\!]_0$ we obtain $\mu(\boldsymbol{P}_{\mathcal{T}_t^{\mathrm{iso}}}\boldsymbol{\Phi}_{\mathrm{2had}}^\top) = 2^{-(t-1)_+}$ and $\mu_{t,l}^{\mathcal{T}^{\mathrm{iso}},\mathcal{T}^{\mathrm{iso}}}(\boldsymbol{\Phi}_{\mathrm{2had}}^\top\boldsymbol{\Psi}_{\mathrm{idhw}}) = 2^{-2(t-1)_+}\cdot\delta_{t,l}$. To compute the relative sparsity, from Lemma 2.5 and Prop. 3.6, we have

$$
\begin{aligned}
&K_t^{\mathcal{T}^{\mathrm{iso}},\mathcal{T}^{\mathrm{iso}}}(\boldsymbol{\Phi}_{\mathrm{2had}}^\top\boldsymbol{\Psi}_{\mathrm{idhw}}, \boldsymbol{k})^{\frac{1}{2}} \\
&\qquad \le k_0^{\frac{1}{2}}\cdot\delta_{t,0} + \sum_{l=1}^{r}\|\boldsymbol{I}_3\otimes\big(\boldsymbol{H}_{(l-1)}\otimes\boldsymbol{H}_{(l-1)}\big)\|_{2,2}\,k_l^{\frac{1}{2}}\cdot\delta_{t,l} \\
&\qquad = k_t^{\frac{1}{2}}.
\end{aligned}
$$

We now prove (3.13c). Consider $t,l \in [\![(r+1)^2]\!]$ such that $t \overset{r+1}{\rightleftharpoons} (t_1+1, t_2+1)$, $l \overset{r+1}{\rightleftharpoons} (l_1+1, l_2+1)$ and $t_1,t_2,l_1,l_2 \in [\![r]\!]_0$. From (3.10a) and using (3.20a) we have

$$
\begin{aligned}
\mu(\boldsymbol{P}_{\mathcal{T}_t^{\mathrm{aniso}}}\boldsymbol{\Phi}_{\mathrm{2had}}^\top\boldsymbol{\Psi}_{\mathrm{adhw}}\boldsymbol{P}_{\mathcal{T}_l^{\mathrm{aniso}}}^\top) &= \mu(\boldsymbol{H}_{(t_1-1)_+})\cdot\mu(\boldsymbol{H}_{(t_2-1)_+})\cdot\delta_{t_1,l_1}\cdot\delta_{t_2,l_2} \\
&= 2^{\frac{-(t_1-1)_+}{2}}\cdot 2^{\frac{-(t_2-1)_+}{2}}\cdot\delta_{t_1,l_1}\cdot\delta_{t_2,l_2}.
\end{aligned}
$$

Therefore, $\mu(\boldsymbol{P}_{\mathcal{T}_t^{\mathrm{aniso}}}\boldsymbol{\Phi}_{\mathrm{2had}}^\top) = 2^{\frac{-(t_1-1)_+}{2}}\cdot 2^{\frac{-(t_2-1)_+}{2}}$ and

$$
\mu_{t,l}^{\mathcal{T}^{\mathrm{aniso}},\mathcal{T}^{\mathrm{aniso}}}(\boldsymbol{\Phi}_{\mathrm{2had}}^\top\boldsymbol{\Psi}_{\mathrm{adhw}}) = 2^{-(t_1-1)_+}\cdot 2^{-(t_2-1)_+}\cdot\delta_{t_1,l_1}\cdot\delta_{t_2,l_2}.
$$

To compute the relative sparsity, from Lemma 2.5 and Prop. 3.6, we have

$$\begin{aligned} K_t^{\mathcal{T}^{\text{aniso}},\mathcal{T}^{\text{aniso}}}(\mathbf{\Phi}_{2\text{had}}^\top \boldsymbol{\Psi}_{\text{adhw}}, \boldsymbol{k})^{\frac{1}{2}} &\leq \sum_{l=1}^{(r+1)^2} \|\boldsymbol{H}_{(l_2-1)_+} \otimes \boldsymbol{H}_{(l_1-1)_+}\|_{2,2}\, k_l^{\frac{1}{2}} \cdot \delta_{t,l} \\ &= k_t^{\frac{1}{2}}. \end{aligned}$$

3.5.6 Proof of Thm. 3.10

Following Thm. 2.6, since in all cases covered by Thm. 3.10 (*i.e.*, 1-D Hadamard-Haar, 2-D isotropic Hadamard-Haar, and 2-D anisotropic Hadamard-Haar) we have $\mathcal{W} = \mathcal{S}$, we need to show that the sample-complexity bound for each case satisfies

$$\begin{aligned} &m_t \gtrsim |\mathcal{S}_t| \cdot \Big(\sum_{l=1}^{|\mathcal{S}|} \mu_{t,l}^{\mathcal{S},\mathcal{S}}(\mathbf{\Phi}^\top \mathbf{\Psi}) \cdot k_l\Big) \cdot \log(K\epsilon^{-1}) \cdot \log(N), \\ &m_t \gtrsim \hat{m}_t \cdot \log(K\epsilon^{-1}) \cdot \log(N), \end{aligned}$$

where $\hat{m}_t$ must satisfy

$$\sum_{t=1}^{|\mathcal{S}|} \frac{|\mathcal{S}_t| \cdot \mu_{t,l}^{\mathcal{S},\mathcal{S}}(\mathbf{\Phi}^\top \mathbf{\Psi}) \cdot K_t^{\mathcal{S},\mathcal{S}}(\mathbf{\Phi}^\top \mathbf{\Psi}, \boldsymbol{k})}{\hat{m}_t} \lesssim 1, \quad \text{for } l \in [\![|\mathcal{S}|]\!].$$

Moreover, since in the three covered cases the multilevel coherence $\mu_{t,l}^{\mathcal{S},\mathcal{S}}(\mathbf{\Phi}^\top \mathbf{\Psi})$ vanishes for $t \neq l$, and $\mu_{t,l}^{\mathcal{S},\mathcal{S}}(\mathbf{\Phi}^\top \mathbf{\Psi}) = |\mathcal{S}_l|^{-1}$ for $t = l$, the proof is further simplified, as the condition on $\hat{m}_t$ holds if $\hat{m}_l \gtrsim K_l^{\mathcal{S},\mathcal{S}}(\mathbf{\Phi}^\top \mathbf{\Psi}, \boldsymbol{k})$. Thus, it suffices to show that in each case

$$m_t \gtrsim \max\Big(K_t^{\mathcal{S},\mathcal{S}}(\mathbf{\Phi}^\top \mathbf{\Psi}, \boldsymbol{k}), k_t\Big) \cdot \log(K\epsilon^{-1}) \cdot \log(N).$$

However, for the three cases, we have $\max\Big(K_t^{\mathcal{S},\mathcal{S}}(\mathbf{\Phi}^\top \mathbf{\Psi}, \boldsymbol{k}), k_t\Big) = k_t$. Therefore, $m_t \gtrsim k_t \cdot \log(K\epsilon^{-1}) \cdot \log(N)$ for all the three cases, which completes the proof.

3.6 Discussion

This chapter has studied the Hadamard-Haar systems in the context of CS theory, *i.e.*, the problem of recovering signals from subsampled Hadamard measurements using Haar wavelet sparsity basis.

Traditional UDS scheme is inapplicable in Hadamard-Haar systems, since the Hadamard and Haar bases are maximally coherent. The new CS principles, *i.e.*, local and multilevel coherence, introduced by Krahmer and Ward [61] and by Adcock *et al.* [9], respectively, inspired us to design sampling strategies that require minimum number of Hadamard measurements and in the same time allow stable and robust signal recovery. By computing the exact values of local and multilevel coherence we achieved the tight sample-complexity bounds for both uniform and non-uniform recovery guarantees. In two-dimensions, we considered two constructions of the 2-D Haar wavelet basis, *i.e.*, using either tensor product of two 1-D Haar bases or the isotropic construction of a multiresolution analysis; and observed that an efficient design of sampling strategy for each system is unique.

Our results have been illustrated by several numerical tests for different types of signals with varying resolution, sparsity, and number of measurements. In particular, we have numerically demonstrated the impact of the resolution in signal recovery.

Our uniform recovery guarantee in Thm. 3.8 is linked to the ℓ_1 minimization problem (2.9). A variant of this problem would be to replace the ℓ_1-norm term with the total variation norm. Following the proof of Thm. 3.1 in [61] we believe that the same sample-complexity bounds and sampling strategies as in Thm. 3.8 provides stable and robust signal recovery (from subsampled Hadamard measurements) via the total variation norm minimization problem. However, we postpone this potential extension to a future study.

Following the uncovered cells in Table. 3.1, a line of study would be to characterize the effect of the other sparsity bases on our local and multilevel coherence analysis, *e.g.*, the 2-D Daubechies wavelets.

Finally, in this respect, it is worth mentioning that the recurrence relations provided by the Kronecker factorization in (3.2) and (3.7) goes beyond the Hadamard and Haar matrices. In fact, the Kronecker product has been used to describe a range of other unitary matrices, *e.g.*, the discrete Fourier transform and the related Sine, Cosine, and Hartley transforms [48, 69, 96]; see also [39] for the factorization of the Daubechies wavelets. An interesting research would be to investigate the combinations of different sensing and sparsity bases and to find other scaling structures.

CHAPTER 4

COMPRESSIVE FOURIER TRANSFORM INTERFEROMETRY

Having discussed about CS in Chapter 2, in this chapter we target an application of CS in capturing HS data using interferometry, *i.e.*, the FTI.

The FTI is an appealing HS imaging modality for many applications demanding high spectral resolution, *e.g.*, in fluorescence microscopy. However, the effective resolution of the FTI is limited by the durability of biological elements when exposed to illuminating light. Over-exposed elements are indeed subject to *photo-bleaching* and become unable to fluoresce. In this context, the acquisition of biological HS volumes based on sampling the Optical Path Difference (OPD) axis at Nyquist rate leads to unpleasant trade-offs between spectral resolution, quality of the HS volume, and light exposure intensity. In this chapter, we propose two variants of the FTI imager, *i.e.*, Coded Illumination FTI (CI-FTI) and Structured Illumination FTI (SI-FTI), based on the theory of CS. These schemes efficiently modulate the light exposure temporally (in CI-FTI) or spatiotemporally (in SI-FTI). Leveraging two VDS strategies (see § 4.4 and § 4.5) and an MDS strategy (see § 4.7), we provide efficient illumination strategies, so that the light exposure imposed on a biological specimen is minimized while the spectral resolution is preserved. Our theoretical analysis focuses on two criteria: *(i)* a trade-off between exposure intensity and the quality of the reconstructed HS volume for a given spectral resolution; *(ii)* maximizing HS volume quality for a fixed spectral resolution and constrained light exposure budget. Our solutions can be adapted to the FTI imager without hardware modifications. The reconstruction of HS volumes from compressive FTI measurements relies on the ℓ_1-norm minimization problem (2.3) promoting a spatiospectral sparsity prior. Numerically, we support the proposed methods on synthetic data and simulated compressive sensing measurements (from actual FTI measurements) under various scenarios. In particular, the biological HS volumes considered in this chapter can be reconstructed with a three-to-ten-fold reduction in the light exposure.

The content of this chapter has been published in [58, 78–80, 83, 86].

4.1 Definition of the Problem

An HS volume refers to a 3-D data cube associated with the stacks of 2-D spatial images along the spectral axis, *i.e.*, one monochromatic (or single-color) image for each wavenumber. Each spatial location records the variation of transmitted or reflected light intensity along a large number of spectral bands from that location in a scene.

Since chemical elements have unique spectral signatures, referred to as their fingerprints, observing the dense spectral content of an HS volume, instead of a RGB image, provides accurate information about the constituents of a scene. This has made HS imaging appealing in a wide range of applications, *e.g.*, agriculture [63], food processing [47], chemical imaging [37], and biology [65, 70, 128].

Due to the extremely large amount of data stored in an HS volume, also referred to as HS cube, the acquisition and processing of these volumes encounters significant challenges. For example, in satellite or airborne applications it is common to transmit volumes with size of several gigabytes to a ground station. Additionally, in biological applications, designers of HS imagers face a myriad of trade-offs related to photon efficiency of the sensors, acquisition time, achievable spatiospectral resolution, and cost.

Compared to other techniques that often reach lower spectral resolutions[1], the FTI is an HS imaging modality that has shown promising results in the acquisition of high spectral resolution biological HS volumes[2], specially in fluorescence spectroscopy. In this application a biological specimen is stained with dyes that fluoresce with a unique spectrum when they are exposed to light. Since each fluorescent dye binds to a certain protein, this technique is frequently used for determining the localization, concentration, and growth of certain target cells (*e.g.*, malignant tumors), whose automatic characterizations are improved at higher spectral resolution.

As described in § 4.2.2, the FTI consists in the MI [77] with one moving mirror controlling the optical path difference of the light. Besides, the raw FTI measurements are intermediate data, *i.e.*, the Fourier coefficients of the actual HS volume with respect to the spectral domain. Shannon-Nyquist sampling theory states that for a fixed wavenumber range, the spectral resolution of an HS volume is proportional to the number of FTI measurements (see § 4.2.2.2). This fact is simultaneously a blessing and a curse in fluorescence spectroscopy, where high spectral resolution is highly desirable for biologists. One may record more FTI measurements, thereby over-exposing fluorescent dyes. As a consequence, the ability of the dyes to fluoresce fades during the experiments. This phenomenon, that is due to the photochemical alteration of the dyes, is called *photo-bleaching* [31]. Therefore, the resolution of an HS volume is limited by the durability of the fluorescent dyes when exposed to the illumination. An alternative is to reduce light intensity in a fashion that is inversely proportional to the increase in the number of FTI measurements, resulting in a low SNR. Considering all these facts, our main motivation in this thesis is to reduce photo-bleaching in fluorescence spectroscopy using the FTI. However, this requires delicate

[1] See, *e.g.*, [102] for a comprehensive classification of different HS acquisition methods.

[2] They contain hundreds of spectral samples in the range of visible (400 nm-700 nm) and or near-infrared (700 nm-1000 nm) wavelengths.

compromises between the goals of achieving high spectral resolution and high SNR while minimizing the light exposure imposed on the specimen.

In this chapter, we propose two compressive FTI imagers, *i.e.*, CI-FTI and SI-FTI, that are near-optimal in the sense of minimizing the light exposure on the biological specimen. In CI-FTI scheme the light source is randomly activated at few time intervals (or slots) during the operation of the FTI, while in SI-FTI, at each time slot, the illumination is activated on a programmed set of spatial locations (*e.g.*, using a Spatial Light Modulator (SLM)). Both models amount to an incomplete set of FTI measurements that must be processed to recover the biological HS volume. While a common, but inaccurate, recovery procedure consists in applying the inverse Fourier transform on the (properly zero-padded) incomplete measurements [22] — with undesirable side-lobes artifacts limiting the usability of the reconstructed HS volume for biomedical purposes — we resort here to the theory of CS explained in Chapter 2 for reconstructing the HS volume. Recall that in CS, the recovery of a signal from a small set of (random) linear measurements is ensured if the number of measurements is higher than the intrinsic dimension of the signal (*e.g.*, its sparsity level [42, 74]). A critical aspect of the adaptation of CI-FTI and SI-FTI to CS theory concerns the way FTI measurements are sub-sampled. As discussed in § 2.2, classical UDS of the signal's frequency domain, is inefficient in the setting of the FTI; the incompatibility between the Fourier sensing and the regularity of the HS volume (*e.g.*, in a wavelet domain) imposes limited sensing compression ratio. However, as most informative FTI measurements are concentrated around low frequencies [79], we here consider the sampling strategy as a degree of freedom in the system model; we leverage the VDS procedure introduced by Krahmer and Ward [61] (see § 2.3) to boost the efficiency of the illumination coding strategy of CI-FTI and SI-FTI. In the sequel, we also apply the MDS scheme of Adcock *et al.* to our CI-FTI system.

Although NDS-based methods are new in the field of HS imaging [99, 106], they have received considerable interest in MRI applications where the recorded data is the 2-D Fourier transform of an image. Our problem in this chapter differs in the sense that the practical FTI devices entail 3-D signals whose 1-D Fourier transform (with respect to spectral domain) is recorded.

4.1.1 Related Works

Hyperspectral imaging is an appealing area of research for the application of CS theory, in terms of acquisition [11, 43, 45, 46, 78, 79, 83, 99, 106, 122, 126] and processing [44, 67, 75, 129]. We here provide a non-exhaustive list of the state-of-the-art HS imaging strategies related to our study in this thesis.

Coded Aperture Snapshot Imager (CASSI): Prior works in [11, 43, 122] have proposed three types of CASSI whose working principles are affiliated to our structured illumination scheme in § 4.5. The imager in [43], termed as dual disperser CASSI, contains two dispersive elements and an aperture in between. The input scene is measured as a 2-D multiplexed spectral projection. The corresponding sensing operator is equivalent to a cyclic S-matrix (*i.e.*, an alteration of the Hadamard matrix) and the solution of the system is estimated from subsampled measurements using the expectation maximization algorithm. A variant of CASSI, called single disperser CASSI, that includes only one dispersive element is presented in [122], wherein a 2-D multiplexed spatiospectral projection of the scene is measured and the HS volume is recovered via the ℓ_1 optimization problem (2.3). In [11], the idea of single-shot CASSI in [43] and [122] is extended to an imager that collects multiple shot measurements.

HS acquisition based on random Gaussian sensing matrix: Besides CASSI systems, [46] and [45] described a theoretical CS scheme for HS data acquisition based on random Gaussian projections. They provided a reconstruction method based on a convex optimization constrained by an ℓ_2-fidelity term with respect to the noisy compressive measurements. The convex objective function in [46] penalizes both the trace-norm and the (spatial) Total Variation (TV) norm of the HS volume; whereas in [45] the TV norm is replaced by an $\ell_{2,1}$ mixed-norm of the data matrix inferring a group-sparse structure in the HS volume.

Single-pixel HS imaging: There exist several works that focus on efficient subsampling strategies for the spatial domain of HS data. Based on the single-pixel imaging technique [34] (see § 3.1.1 for its other variants), the proposed HS microscopes in [106] and [99] target the problem of acquiring high spatial resolution HS volume from compressive measurements using Hadamard sensing model, *e.g.*, by structuring the wide-field illumination [106] as in SI-FTI. The authors in [106] compared the UDS and half-half density samplings in their implementation while [99] employs an MDS [7, 8] for the same microscope used in [106] in addition to taking into account the photon noise as well as the lens point spread function. The authors in [59] have recently advocated a compressive FTI system based on the single-pixel imaging. We have improved the structured illumination coding of the system in [59] using the VDS and MDS schemes; see Chapter 4 for the details.

HS acquisition from subsampled Fourier sensing: To the best of our knowledge, we introduced the first compressive FTI in [79] by subsampling the OPD dimension of the interferometric signals. Our investigations in [78, 79, 83] showed that a VDS, when promoting low-frequencies, captures more informative FTI samples. Recently, a versatile scheme for fluorescence microscopy has appeared in the literature [126]. It applies when the acquisition of the HS volume is done consecutively in the axial domain such that at every axial point the light source is modulated by a waveform. The CS measurements are essentially formed by recording deterministically the primitive part of axial domain, *i.e.*, low-pass frequency sampling. A post-processing approach to increase the spectral resolution of an FTI-like system, *i.e.*, a Sagnac interferometer, coupled with structured light source is demonstrated in [26]. The objective of [26] is to mitigate the physical limitation for achieving a high resolution HS volume; nonetheless this approach significantly increases photo-bleaching.

This chapter provides a clear, fivefold contribution compared to the aforementioned studies. First, we target the problem of compressive HS volume acquisition; second, the physical model of the FTI device imposes the Fourier sensing operator (since it captures interferometric information) that brings about challenges different from ideal Gaussian sensing operator; third, we study the trade-off between the spectral resolution of the HS volume and photo-bleaching deterioration of the fluorescent dyes; forth, we apply a VDS scheme to the field of interferometric HS imaging; and fifth, unlike SP-FTI models [59, 81, 82], the proposed compressive FTI systems are based on 2-D imaging sensors. In the sequel, we present an improved coding scheme for CI-FTI published in [86] by leveraging the sparsity structure of the fluorochrome spectra and combining it with the MDS scheme.

Table 4.1 summarizes the related HS acquisition models and highlights the differences of the problems studied in this thesis. Concerning the state-of-the-art of the NDS schemes we refer the reader to § 3.1.1.

This chapter is organized as follows. We start by explaining the acquisition model of a conventional FTI imager. The problem formulation and theoretical analysis of CI-FTI and SI-FTI frameworks are presented in § 4.4 and § 4.5, respectively. The concept of constrained light exposure budget for both CI-FTI and SI-FTI is revealed in § 4.6. An extension of CI-FTI to a more biologically-related scenario is presented in § 4.7, where the illumination coding is improved by incorporating the structure of the fluorochrome spectral signatures. scheme In § 4.8, we focus on comprehensive synthetic and experimental tests of our schemes demonstrating the efficiency of the proposed methods when the fluorochromes in a biological specimen undergone photo-bleaching. The proofs of the

	Spectrometry		Study type			Prior model			Sampling	
	Interferometric	Direct	Hardware	Numerical	Analytical	3D low-rank	Spectral sparsity	Spatial sparsity	Operator	Strategy
Gehm *et al.* [43]		✓	✓				DW	DW	RSE	UDS
Wagadarikar *et al.* [122]		✓	✓				Dirac	DW	RSE	UDS
Arguello *et al.* [11]		✓	✓				DC	DW	RSE	UDS
Golbabaee *et al.* [46]		✓		✓		✓	Dirac	TV	RGM	—
Golbabaee *et al.* [45]		✓		✓		✓	Dirac	DW	RGM	—
Studer *et al.* [106]		✓	✓				Dirac	DW	RHE	Half-half
Roman *et al.* [99]		✓		✓			Dirac	DW	RHE	MDS
Woringer *et al.* [126]		✓	✓	✓			Dirac	Dirac	LFE	Deter.
CI-FTI: Thm. 4.12	✓			✓	✓		Dirac	DW	RFE	VDS
CI-FTI: Thm. 4.21	✓			✓	✓		Dirac	DW	RFE	MDS
SI-FTI: Thm. 4.16	✓			✓	✓		DW	DW	RFE	VDS

RSE: Random S-matrix Ensembles	RGM: Random Gaussian Mixtures
RHE: Random Hadamard Ensembles	RFE: Random Fourier Ensembles
LFE: Low-pass Fourier Ensembles	TV: Total Variation
DC: Discrete Cosine	DW: Discrete Wavelet
UDS: Uniform Density Sampling	VDS: Variable Density Sampling
MDS: Multilevel Density Sampling	Deter.: Deterministic

Table 4.1: Comparison of the related acquisition modalities in HS imaging. In direct spectrometry the spectral dimension is recorded by either wavelength filtering, grating, or prism. Half-half sampling strategy means half of the M measurements is taken from the lowest frequency components and the other half is taken by randomly sub-sampling the rest of the remaining samples. Note that compressive FTI systems designed based on single-pixel imaging technique, *e.g.*, [59, 81, 82], are not reported in this table.

main results are all postponed to § 4.9.1. Finally, we conclude with some remarks and perspectives in § 4.10.

4.2 Acquisition Model in Conventional FTI

This section describes the principles of the MI before explaining the FTI acquisition, its continuous observation model and its crucial discretization into the forward model that can be inverted numerically.

4.2.1 Principles of the Michelson Interferometer

We here summarize the principle of the MI [77] (see, *e.g.*, [13] for a detailed survey on the MI). The MI is a system that produces interference

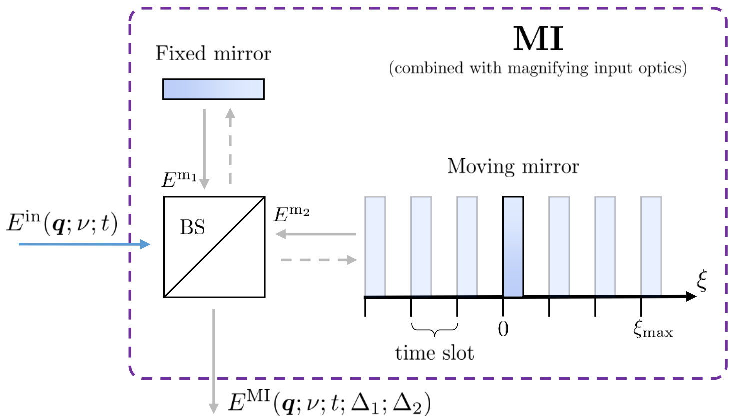

Figure 4.1: Operating principle of the MI.

between two beams of light. In a nutshell, as schematized in Fig. 4.1 a single light beam entering the MI is divided into two beams by a Beam-Splitter (BS), *e.g.*, made of two birefringent prisms. The two beams are then reflected back by a fixed and a moving mirror, respectively, and interfere after being recombined by the BS. In imaging applications the intensity of the resulting beam is next acquired by either an array of sensors (as in the FTI) or a single sensor (as in the SP-FTI).

Let us now develop an idealized optical model of this MI. Given an orthonormal coordinate system $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\} \subset \mathbb{R}^3$, a monochromatic light beam entering the MI and traveling along $\boldsymbol{e}_3$-direction may be formulated at some point $\boldsymbol{q} := [q_1, q_2, q_3]^\top \in \mathbb{R}^3$ and time t as

$$E^{\mathrm{in}}(\boldsymbol{q};\nu;t) = E_0^{\mathrm{in}}(q_1, q_2;\nu;t)e^{i(2\pi q_3 \nu - \omega t)}, \tag{4.1}$$

where $\nu > 0$ is the wavenumber, $\omega = 2\pi\nu \mathrm{c_v}$ is the angular frequency (with $\mathrm{c_v}$ the light velocity in the medium) and $E_0^{\mathrm{in}}(q_1, q_2;\nu;t)$ represents the amplitude of light at $(q_1, q_2;\nu;t)$; see, *e.g.*, [127, Chapter 32] for the principles of electromagnetic wave propagation. The beam (4.1) is first divided into two beams by the BS. After perfect reflection on their respective mirror, and assuming perfect reflection and transmission in the BS, the incident beams at the BS read

$$E^{\mathrm{m}_1}(\boldsymbol{q};\nu;t;\Delta_1) = E_0^{\mathrm{in}}(q_1,q_2;\nu;t)e^{i(2\pi(q_3+\Delta_1)\nu-\omega t)},$$

and

$$E^{\mathrm{m}_2}(\boldsymbol{q};\nu;t;\Delta_2) = E_0^{\mathrm{in}}(q_1,q_2;\nu;t)e^{i(2\pi(q_3+\Delta_2)\nu-\omega t)},$$

where Δ_1 and Δ_2 are the distances that the two beams have traveled inside the MI. Therefore, neglecting the equal additional paths traveled in the BS, the recombined beam is

$$E^{\mathrm{MI}}(\boldsymbol{q};\nu;t;\Delta_1,\Delta_2) = E^{\mathrm{m}_1}(\boldsymbol{q};\nu;t;\Delta_1) + E^{\mathrm{m}_2}(\boldsymbol{q};\nu;t;\Delta_2).$$

Finally, the intensity $I^{\mathrm{MI}}(\xi;\nu;t) = |E^{\mathrm{MI}}(\boldsymbol{q};\nu;t;\Delta_1,\Delta_2)|^2$ of this beam, which will be needed in the next sections, can be computed according to the rule

$$I^{\mathrm{MI}}(\xi;\nu;t) = 2\,|E_0^{\mathrm{in}}(q_1,q_2;\nu;t)|^2(1+\cos(2\pi\nu\xi)), \tag{4.2}$$

where $\xi := \Delta_2 - \Delta_1$ is the OPD parameter.

With this summary in hand, we describe the principles of the FTI and SP-FTI in the following sections.

4.2.2 Fourier Transform Interferometry

In a nutshell, the operating principle of a conventional FTI, called hereafter the *Nyquist FTI*, *i.e.*, collecting as many observations as the number of HS volume voxels so that HS reconstruction can be achieved by a linear (inverse) transform, is based on the MI described above. As shown in Fig. 4.2, a parallel beam of coherent light[3] obtained from a 2-D image of a thin, still, and almost transparent biological specimen magnified by a (confocal) microscope, enters the MI. As explained above, this group

[3]A group of light beams are coherent if they are emitted by coherent sources, *i.e.*, monochromatic (or single-color) sources of the same frequency and with any constant phase relationship (see *e.g.*, [127, Chapter 35.1]).

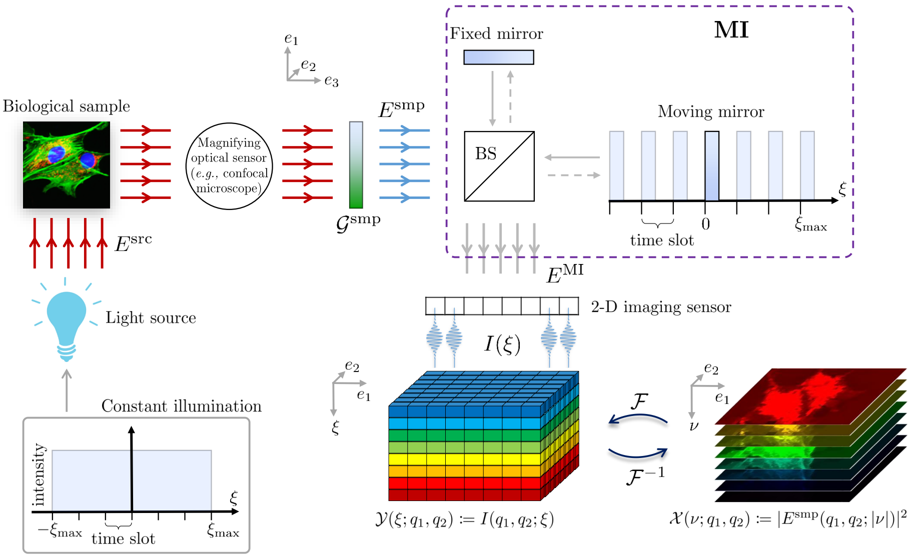

Figure 4.2: Operating principle of FTI.

of beams is divided into two groups (by the BS) that are next reflected back by a fixed and a moving mirror, and interfere after being recombined by the BS. The intensity of the outgoing group of beams from the MI is then acquired by a standard imaging sensor (camera); each image captured by this camera corresponds to a given position of the moving mirror, and each pixel records temporally a specific interference pattern.

4.2.2.1 Continuous Observation Model of the FTI

In order to understand how an HS volume of the observed biological specimen can be deduced from the recorded interferences, let us develop an idealized optical model of this FTI.

Assumption 4.1. *We simplify our description by omitting the magnifying optics that are required to observe the spatial details of the specimen. Such an optical setup* (e.g., *a confocal microscope) can be easily inserted in our* *scheme by an introduction of a scaling factor between the coordinate systems of the biological specimen plane and the 2-D imaging sensor plane.*

Similar to the description of the MI, a coherent wide-band plane wave emitted by the light source and traveling along the $\boldsymbol{e}_3$-direction may be formulated at some point $\boldsymbol{q} \in \mathbb{R}^3$ and time t as the integral of the monochromatic waves with respect to the wavenumber, *i.e.*,

$$E^{\mathrm{src}}(\boldsymbol{q}; t) = \int_0^{+\infty} E_0^{\mathrm{src}}(q_1, q_2; \nu, t) e^{i(2\pi q_3 \nu - \omega t)} \,\mathrm{d}\nu, \tag{4.3}$$

where[4] $E_0^{\mathrm{src}}(q_1, q_2; \nu; t)$ is the amplitude of light at $(q_1, q_2; \nu; t)$.

Assumption 4.2. *In this thesis, we consider that the illumination ensures that $E_0^{\mathrm{src}}(q_1, q_2; \nu; t)$ is constant in time and with respect to (q_1, q_2),* i.e., *$E_0^{\mathrm{src}}(q_1, q_2; \nu; t) = E_0^{\mathrm{src}}(\nu)$, so that the light source intensity per second and per unit area*

$$I^{\mathrm{src}} := \int_0^{+\infty} |E_0^{\mathrm{src}}(q_1, q_2; \nu, t)|^2 \,\mathrm{d}\nu = \int_0^{+\infty} |E_0^{\mathrm{src}}(\nu)|^2 \,\mathrm{d}\nu, \tag{4.4}$$

is constant in the same way.

Therefore, $c_0 I^{\mathrm{src}}$ is the total light exposure received per unit of time and per unit of surface on the biological specimen, with $c_0 > 0$ depending on the speed of light c_{v} in the vacuum, the medium refractive index and the vacuum permittivity. Hereafter, for simplicity and up to a general re-scaling of the intensities, we consider that $c_0 = 1$.

Neglecting the refraction of light induced by the transparent biological specimen, we can forget the representation of (q_1, q_2) in our next developments as well as in Fig. 4.2. This is equivalent to considering a single spatial location in the following equations. Moreover, the beam folding action of the optical elements in the MI (*i.e.*, the mirrors and the BS) are still compatible with the q_3-parametrization, provided that q_3

[4] Recall that in the application of the FTI in fluorescence spectroscopy the light frequency $\nu\mathsf{c}_{\mathrm{v}} \in$ [400 THz, 770 THz], with c_{v} the speed of light in the vacuum.

stands for the path length of the light propagation till the imaging sensor. Note that, from a one-to-one correspondence between the specimen plane and the imaging sensor, the coordinates (q_1, q_2) are associated with a 2-D pixel location on the imaging plane.

Since a coherent wide-band light is a collection of monochromatic waves, we proceed by considering a monochromatic wave $E^{\mathrm{src}}(q_3;\nu;t) = E_0^{\mathrm{src}}(\nu)e^{i(2\pi q_3\nu-\omega t)}$ of wavenumber ν that is incident to the biological specimen. After having traveled through the specimen, this beam reads

$$E^{\mathrm{smp}}(q_3;\nu;t) = E_0^{\mathrm{smp}}(\nu)e^{i(2\pi q_3\nu-\omega t)}. \tag{4.5}$$

Note that in non-fluorescent biological applications, the field intensity $|E_0^{\mathrm{smp}}(\nu)|^2$ is the result of the multiplication of $|E_0^{\mathrm{src}}(\nu)|^2$ by the absorption of the constituents. In fluorescent microscopy, however, the two field intensities are related through a more general model: in addition to absorption phenomena, a fluorescent material re-emits light at different wavelengths compared to the one of the incident light beam.

Assumption 4.3. *We consider here that E_0^{smp} and E_0^{src} are related through a general transfer operator $\mathcal{G}^{\mathrm{smp}}$*, i.e.,

$$E_0^{\mathrm{smp}}(\nu) = \mathcal{G}^{\mathrm{smp}}[E_0^{\mathrm{src}}](\nu), \tag{4.6}$$

which is linear with respect to the total light intensity, i.e., *for any $\lambda > 0$,*

$$I^{\mathrm{src}} \to \lambda I_0^{\mathrm{src}} \quad \Rightarrow \quad I^{\mathrm{smp}} \to \lambda I^{\mathrm{smp}}, \tag{4.7}$$

where $I^{\mathrm{smp}} := \int_0^{+\infty} |\mathcal{G}^{\mathrm{smp}}[E_0^{\mathrm{src}}](\nu)|^2 \mathrm{d}\nu$, which will identify the continuous HS data in this thesis; see below.

In words, if the illumination intensity is increased by a factor of λ, the same amplification is observed for the intensity of the light that has traveled though the specimen. As will be clear later, this assumption actually sustains a compressive FTI mode where the total light exposure is kept constant with respect to the Nyquist FTI acquisition (see § 4.6).

Next, from the field amplitude conversion (4.6) occurring in the specimen, the beam (4.5) enters the MI. Following the discussion in § 4.2.1 the intensity of the outgoing beam from the MI reads the relation (4.2) with $E_0^{\rm in} \leftarrow E_0^{\rm smp}$, *i.e.*,

$$I^{\rm MI}(\xi;\nu) = 2\,|E_0^{\rm smp}(\nu)|^2(1+\cos(2\pi\nu\xi)). \tag{4.8}$$

As illustrated in Fig. 4.2, there is a one-to-one correspondence between the time domain (or time slot), mirror position and the OPD value. Moreover, to fix the ideas, we consider throughout this chapter that each time slot (or the OPD sample) integrates light over a constant time duration $\tau_\xi > 0$. This time is actually associated with the frame-per-second (fps) rate of the imaging sensor.

Coming back to the general wide-band plane wave model (4.3) with constant amplitude and assuming that τ_ξ is much larger than the temporal range[5] $1/(\nu c_{\rm v})$, we quickly verify that the total temporal-averaged intensity recorded by the detector can be written as

$$I_{\rm T}(\xi) = \int_0^{+\infty} I^{\rm MI}(\xi;\nu)\,{\rm d}\nu = I_{\rm m} + I(\xi),$$

with $I_{\rm m} = 2\int_0^{+\infty}|E_0^{\rm smp}(\nu)|^2\,{\rm d}\nu = \frac{1}{2}I_{\rm T}(0)$. After removing the mean component, it is easy to show that the zero-mean part I, termed as *interferogram*, is the Fourier transform of $|E_0^{\rm smp}(|\nu|)|^2$, *i.e.*, the symmetrization of $\nu \in \mathbb{R}_+ \mapsto |E_0^{\rm smp}(\nu)|^2 \in \mathbb{R}_+$ around $\nu = 0$:

$$\begin{aligned} I(\xi) &= I_{\rm T}(\xi) - \frac{1}{2}I_{\rm T}(0) \\ &= 2\int_0^{+\infty}|E_0^{\rm smp}(\nu)|^2\cos(2\pi\nu\xi)\,{\rm d}\nu \\ &= \int_{-\infty}^{+\infty}|E_0^{\rm smp}(|\nu|)|^2 e^{-i2\pi\nu\xi}\,{\rm d}\nu. \end{aligned} \tag{4.9}$$

[5]Which is clearly the case for a visible light and near-infrared HS system with $\nu c_{\rm v}$ in the range of 400 THz to 770 THz, and τ_ξ of the order of a few hundredths of a second.

Consequently, computing the inverse Fourier transform of the interferogram and restricting it to the positive spectral axis yields the spectrum[6] $|E_0^{\rm smp}(\nu)|^2$ of the observed specimen on a given spatial location (q_1, q_2). Overall, reinserting these spatial coordinates, we defined the continuous HS volume and the volume of interferograms, respectively, as

$$\mathcal{X}(\nu; q_1, q_2) := |\mathcal{G}^{\rm smp}[E_0^{\rm src}](q_1, q_2; |\nu|)|^2 = |E_0^{\rm smp}(q_1, q_2; |\nu|)|^2, \tag{4.10}$$

$$\mathcal{Y}(\xi; q_1, q_2) := I(\xi) = I(q_1, q_2; \xi). \tag{4.11}$$

The continuous sensing model relating $\mathcal{X}$ and $\mathcal{Y}$ hence reads

$$\mathcal{Y}(\xi; q_1, q_2) = \mathcal{F}[\mathcal{X}](q_1, q_2; \xi), \tag{4.12}$$

with $\mathcal{F} : f(q_1, q_2; \nu) \to \int_{-\infty}^{+\infty} f(q_1, q_2; \nu) e^{-i2\pi\nu\xi}\, {\rm d}\nu$ representing the one-dimensional (1-D) Fourier transform in the ν domain.

Remark 4.4. *As represented through the action of the transfer operator* $\mathcal{G}^{\rm smp}$ *above, in biological imaging, the spectrum* $|E_0^{\rm smp}(\nu)|^2$ *is thus a combined signature of the absorption spectrum of the biological specimen, the emission spectra of the fluorescent dyes and the spectrum of the light source. In this thesis, we will not consider the question of unmixing these three parts and we thus focus on the acquisition of* $|E_0^{\rm smp}(\nu)|^2$.

Remark 4.5. *In practice, the location of* $\xi = 0$ *(or the zero-OPD point), as required by* (4.9)*, can be obtained through calibration with two methods:* (i) *by looking for the OPD point where the intensity of the recorded signal is twice the value of the empirical mean of the recorded signal, or* (ii) *by identifying the interferogram maximum, which occurs at the zero-OPD from* (4.9)*. Note that the knowledge of the zero-OPD is also critical to define*

[6]This inverse Fourier transform is actually proportional to the spectrum, with a multiplicative constant depending on, *e.g.*, the attenuation and reflection coefficients of the mirrors and the BS, and the sensor pixel efficiency.

our compressive FTI and SP-FTI schemes, as will be seen in § 4.4, § 4.5, and Chapter 4.

4.2.2.2 Discrete Sensing Model of the FTI

The FTI actually proceeds by first capturing discrete samples of the continuous volume $\mathcal{Y}$ in (4.12), both in the spatial and in the time domain according to the pixel grid and the number of frames per second recorded by the imaging sensor (see Fig. 4.2), and by processing them in order to reconstruct a discretized version of $\mathcal{X}$ compatible with the Shannon-Nyquist sampling theorem [54, page 374].

To fix the ideas, we consider the ideal context where:

(i) the moving mirror gives access to an OPD interval $(-\xi_{\max}, \xi_{\max}]$ (for some range $\xi_{\max} > 0$) that is evenly discretized over $N_\xi \in \mathbb{N}$ samples with an OPD step size $\Delta_\xi > 0$ (*i.e.*, $2\xi_{\max} = N_\xi \Delta_\xi$), and for each sample, the detector integrates the recorded intensity over the OPD step size Δ_ξ;

(ii) the spatial domain is evenly discretized over $\bar{N}_{\rm p} \times \bar{N}_{\rm p}$ pixels with a pixel length $\Delta_{\rm p} > 0$, *i.e.*, determined by the pixel pitch of the detector, and each pixel of the detector integrates the intensity over a square grid of length $\Delta_{\rm p} \times \Delta_{\rm p}$.

Accordingly, the time domain is regularly discretized with N_ξ samples related to a *time slot* of $\tau_\xi = \tau_{\rm hs}/N_\xi$, given the total acquisition time[7] $\tau_{\rm hs} > 0$.

Mathematically, the discrete FTI measurements are gathered in a cube $\boldsymbol{\mathcal{Y}} \in \mathbb{R}^{N_\xi \times \bar{N}_{\rm p} \times \bar{N}_{\rm p}}$ approximating $\mathcal{Y}$ over $N_{\rm p} := \bar{N}_{\rm p}^2$ pixels and N_ξ OPD points, *i.e.*, over $N_{\rm hs} = N_\xi N_{\rm p}$ voxels. Similarly, the discrete HS volume is represented by a data cube $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{N_\nu \times \bar{N}_{\rm p} \times \bar{N}_{\rm p}}$ approximating $\mathcal{X}$ over $N_{\rm p}$ pixels and $N_\nu = N_\xi$ wavenumber samples, and thus also over

[7] In practice, the FTI system is equivalently characterized by the frame-per-second rate of the imaging sensor, which determines τ_ξ, and by the speed and the extent of the mirror motion, $2\xi_{\max}$.

$N_{\rm hs}$ voxels. Therefore, assuming N_ξ even and positioning the zero-OPD point on the spectral index $l = N_\xi/2$, the sampling rules are thus

$$\begin{aligned}
&\boldsymbol{\mathcal{Y}}_{l_\xi,j_1,j_2}\\
&\quad:= \iiint \mathrm{rect}(\frac{\xi}{\Delta_\xi} - l_\xi + \frac{N_\xi}{2}, \frac{q_1}{\Delta_{\rm p}} - j_1, \frac{q_2}{\Delta_{\rm p}} - j_2)\mathcal{Y}(\xi; q_1, q_2)\mathrm{d}q_1\mathrm{d}q_2\mathrm{d}\xi,\\
&\boldsymbol{\mathcal{X}}_{l_\nu,j_1,j_2}\\
&\quad:= \iiint \mathrm{rect}(\frac{\nu}{\Delta_\nu} - l_\nu + \frac{N_\nu}{2}, \frac{q_1}{\Delta_{\rm p}} - j_1, \frac{q_2}{\Delta_{\rm p}} - j_2)\mathcal{X}(\nu; q_1, q_2)\mathrm{d}q_1\mathrm{d}q_2\mathrm{d}\nu,
\end{aligned} \tag{4.13}$$

with $l_\xi, l_\nu \in [\![N_\xi]\!]$, $j_1, j_2 \in [\![\bar{N}_{\rm p}]\!]$, $\Delta_\nu = 1/(N_\xi\Delta_\xi)$ and $\nu_{\max} := 1/(2\Delta_\xi)$. Throughout this thesis we call l_ξ (and l_ν) as the OPD index (respectively, wavenumber index). Moreover, we will use the 1-D and 2-D pixel index $l_{\rm p}$ and (j_1, j_2), respectively, related via the rule $l_{\rm p} \overset{\bar{N}_{\rm p}}{\Longleftrightarrow} (j_1, j_2)$. Note that $\boldsymbol{\mathcal{X}}_{l_\nu,j_1,j_2} = \boldsymbol{\mathcal{X}}_{N_\nu - l_\nu,j_1,j_2}$ from the symmetrization of the intensity around $\nu = 0$ in (4.9). This symmetric construction will be assumed throughout the thesis, for the Nyquist FTI model above as well as the SP-FTI and their CS versions presented in the next sections. Consequently, despite the complex nature of the Fourier transform, all entries of $\mathcal{Y}$, partially observed or not, are real, in agreement with the measurement process.

As a result, by increasing the number of the OPD samples the final HS volume contains spectral information with higher details. However, since for a light source with constant intensity the total light exposure of an observed biological specimen is proportional to N_ξ, this increase is limited in practice by the risk of photo-bleaching [31]. Thus, a trade-off between spectral resolution and photo-bleaching must be found.

Finally, if $\boldsymbol{Y}^{\rm fti} \in \mathbb{R}^{N_\xi \times N_{\rm p}}$ and $\boldsymbol{X} \in \mathbb{R}^{N_\xi \times N_{\rm p}}$ denote the matrix unfolding (see the notations) of the cube $\boldsymbol{\mathcal{Y}}$ and $\boldsymbol{\mathcal{X}}$, respectively, the acquisition process of Nyquist FTI can be formulated in matrix form as

$$\boldsymbol{Y}^{\rm fti} = \boldsymbol{F}^*\boldsymbol{X} + \boldsymbol{N}^{\rm fti}, \tag{4.14}$$

where $X_{l_\xi,l_\mathrm{p}} = X_{N_\xi-l_\xi,l_\mathrm{p}}$ for all $l_\xi \in [\![N_\xi/2]\!]$ and $l_\mathrm{p} \in [\![N_\mathrm{p}]\!]$, $\boldsymbol{N}^{\mathrm{fti}} = [\boldsymbol{n}_1^{\mathrm{fti}}, \cdots, \boldsymbol{n}_{N_\mathrm{p}}^{\mathrm{fti}}] \in \mathbb{R}^{N_\xi \times N_\mathrm{p}}$ models an additive noise, and $\boldsymbol{F} \in \mathbb{C}^{N_\xi \times N_\xi}$ is the 1-D discrete Fourier basis. Equivalently, this description can also be arranged into a vector form, *i.e.*,

$$\boldsymbol{y}^{\mathrm{fti}} = \boldsymbol{\Phi}_{\mathrm{fti}}^{*}\boldsymbol{x} + \boldsymbol{n}^{\mathrm{fti}}, \tag{4.15}$$

with $\boldsymbol{\Phi}_{\mathrm{fti}} := \boldsymbol{I}_{N_\mathrm{p}} \otimes \boldsymbol{F}$, $\boldsymbol{y}^{\mathrm{fti}} := \mathrm{vec}(\boldsymbol{Y}^{\mathrm{fti}})$, $\boldsymbol{x} := \mathrm{vec}(\boldsymbol{X})$, and $\boldsymbol{n}^{\mathrm{fti}} := \mathrm{vec}(\boldsymbol{N}^{\mathrm{fti}})$.

4.2.3 Noise Model Identification and Estimation in the FTI

In practice, the FTI and SP-FTI (see Chapter 4) experiments are of course corrupted by different noise sources. Let us list and assess the impact of the main ones.

First, there is the *observation noise* that mainly results from the combination of camera's electronic noise (such as thermal noise), quantization noise and photon noise. Electronic noise is commonly assumed additive, independent of the observations, and distributed as Gaussian and homoscedastic noise, *i.e.*, with constant variance for all measurements. Quantization noise is induced by the digitization of the observations. Strictly speaking, this is not noise but a deterministic distortion. However, when the bit-depth is large (*i.e.*, under the high-resolution assumption), each measurement is quantized over a large number of bits and the quantization distortion is also well-modeled by an additive, heteroscedastic Gaussian noise [49]. Finally, the photon noise (or *shot* noise) is signal dependent and associated with the quantized nature of light, *i.e.*, camera sensors record light intensity by integrating individual photons' energies over a given time interval. Photon noise is modeled by a Poisson distribution, whose variance is equal to its mean. Therefore, for high-photon counting rates, the ratio between the standard deviation and the mean of the light intensity vanishes, so that the impact of the photon noise is limited before electronic and quantization noises. We will assume this high-photon counting regime in all our analysis in this thesis (see § 4.4, § 4.5, and Chapter 4).

Second, a *modeling noise* is induced by the discretization of the sensing model (4.12), *i.e.*, by the discrepancy between the output of (4.14), given the discretization $\boldsymbol{X}$ of the continuous HS volume $\mathcal{X}$, and the true FTI observations in the absence of observation noises. In this chapter, we suppose that this modeling noise is limited by considering large resolutions N_ξ and $N_{\rm p}$. Note that in [7, 99], the authors directly solve the infinite dimensional inverse problem posed by the reconstruction of a continuous function, sparsely representable in a continuous basis, and observed from finite or countable linear observations. We leave for a future study the application of this framework to the proposed models in in this chapter and Chapter 5.

Third, there also exist the *instrument noises* corrupting the acquired data. For instance, bad instrument calibration induces an error in the OPD origin, supposed to lie on $\xi = 0$ in (4.12). We will see in the following that this point is important for applying the VDS strategies. We omit this error by assuming an accurate calibration process. Additionally, the model explained in § 4.2.2.1 does not consider light diffraction through the different optical elements of the MI and through the transparent biological specimen itself. This diffraction induces, however, spatial mixing of the recorded interferences, that would be modeled by convolving uncorrupted observation by an instrumental *point spread function*.

Consequently, since these noise sources are independent and the interferograms in (4.14) are already mean-removed, we will suppose that the discrete FTI observations are corrupted by a global additive, zero-mean, homoscedastic Gaussian noise $\boldsymbol{N}^{\rm fti} = (\boldsymbol{n}_1^{\rm fti}, \cdots, \boldsymbol{n}_{N_{\rm p}}^{\rm fti}) \in \mathbb{R}^{N_\xi \times N_{\rm p}}$, with $N^{\rm fti}_{l_\xi, l_{\rm p}} \sim_{\rm iid} \mathcal{N}(0, \sigma^2_{\rm fti})$ for all OPDs $l_\xi \in [\![N_\xi]\!]$ and pixels $l_{\rm p} \in [\![N_{\rm p}]\!]$. Given the recorded discrete and noisy FTI measurements $\boldsymbol{Y}^{\rm fti}$, the variance $\sigma^2_{\rm fti}$ can be estimated using, *e.g.*, the Robust Median (RM) estimator [33, 112] (see § 4.8.3).

As an important aspect in the analysis of the two compressive FTI schemes developed in § 4.4, § 4.5, we will consider that the noise corrupting these compressive observations is also an additive homoscedastic Gaussian noise with the same variance, *i.e.*, with the variance σ_{fti}^2 estimated from the Nyquist FTI acquisition.

4.3 Sparse Models for a Biological HS Volume

To solve problems of the form (2.9) or (2.16), it is crucial to select a sparsity basis $\boldsymbol{\Psi}$ that captures key information about the signal of interest with a minimum of basis elements: we must ensure that the first error term in the right-hand side of (2.10) or (2.17) decreases rapidly when the sparsity level increases.

Remark 4.6. *Beyond the sparsity models used in this thesis, one can exploit other HS priors. A specimen is comprised of few biological elements whose spectral signatures share common support set in a sparsity basis. Moreover, smoothness of the spectral signatures (see,* e.g.*, Fig. 4.6) results in highly correlated spectral bands. These priors can be imposed on the reconstruction problem using nuclear-norm, mixed-norm, or both [45, 46, 57].*

In this thesis, we focus on two HS sparsity priors both leveraging the Haar wavelet transform in the spectral or in the spatiospectral domains [74]. While these choices are perfectible, *e.g.*, by selecting other wavelet schemes (such as the Daubechies wavelets) or redundant systems [74], the two selected priors yield computable bounds on the local coherence of the corresponding sensing and sparsity bases. Following § 2.2, they thus provide guarantees on the quality of the reconstructed HS volume.

(i) Spectral sparsity model: In the context of fluorescence microscopy of still biological specimen, each spatial location of an HS volume (*i.e.*, $\boldsymbol{\mathcal{X}}$ in § 4.2.2.2) corresponds to the spectrum of the specimen on this position, *i.e.*, the mixture of the spectral signatures of the fluorescent dyes. In general, these spectral signatures are (piece-wise) smooth (see, *e.g.*, [1, 86], and Fig. 4.6 for the spectra of 3 common fluorochromes); we expect them to be sparsely approximable in a wavelet basis, such as the DHW basis [57, 74] defined in § 3.2.

Our first possible prior will thus leverage this fact and will be a purely spectral sparsity inducing prior. This will be crucial in § 4.4 for our first compressive FTI system, *i.e.*, CI-FTI, which subsamples the OPD domain of interferometric data according to an identical sampling pattern for all spatial locations; this scheme can thus be seen as the replication of $N_{\rm p}$ 1-D CS systems that can only be controlled by the HS spectral sparsity.

Mathematically, our purely spectral sparsity basis is defined by $\boldsymbol{\Psi} := \boldsymbol{I}_{N_{\rm p}} \otimes \boldsymbol{\Psi}_{\rm dhw}$. This basis is associated with the following representations of the HS volume $\boldsymbol{X}$:

$$\boldsymbol{X} = \boldsymbol{\Psi}_{\rm dhw}\boldsymbol{S},\ \boldsymbol{S} = \boldsymbol{\Psi}_{\rm dhw}^\top \boldsymbol{X}, \quad \text{or} \quad \boldsymbol{x} = \boldsymbol{\Psi}\boldsymbol{s},\ \boldsymbol{s} = \boldsymbol{\Psi}^\top \boldsymbol{x},$$

with $\boldsymbol{s} := \operatorname{vec}(\boldsymbol{S})$, and where $\boldsymbol{\Psi}_{\rm dhw} \in \mathbb{R}^{N_\xi \times N_\xi}$ is the 1-D DHW basis. From the above-mentioned considerations, the columns of $\boldsymbol{S} = [\boldsymbol{s}_1, \cdots, \boldsymbol{s}_{N_{\rm p}}] \in \mathbb{R}^{N_\xi \times N_{\rm p}}$ (and the vector $\boldsymbol{s}$) are expected to have small best K-term approximation errors $\sigma_K(\boldsymbol{s}_i)_1$ (respectively, $\sigma_K(\boldsymbol{s})_1$) for a relatively small $K \ll N_\xi$ (respectively, $K \ll N_{\rm hs} = N_\xi N_{\rm p}$). § 4.4 leverages this model for our analysis.

(ii) Joint spatiospectral sparsity model: The HS volume of a still specimen observed by the FTI system is made of a collection of monochromatic images describing the spatial configuration of fixed materials at a given wavenumber index. These images are expected to be sparsely approximable in a convenient 2-D basis, *e.g.*, the 2-D IDHW $\boldsymbol{\Psi}_{\rm idhw} \in \mathbb{R}^{N_{\rm p} \times N_{\rm p}}$ or the ADHW basis $\boldsymbol{\Psi}_{\rm adhw} \in \mathbb{R}^{N_{\rm p} \times N_{\rm p}}$ defined in § 3.2.

Therefore, combining this new representation with the spectral sparsity model above, our second sparsity prior is defined by the sparsity basis $\boldsymbol{\Psi} := \boldsymbol{\Psi}_{\rm idhw} \otimes \boldsymbol{\Psi}_{\rm 1D}$. However, our results extend to the 2-D ADHW basis (see § 4.9.2).

Thanks to this sparsity basis $\boldsymbol{\Psi}$, we expect a small best K-term approximation error for the vectorization $\boldsymbol{s}$ of the matrix $\boldsymbol{S} \in \mathbb{R}^{N_\xi \times N_{\mathrm{p}}}$ in the representation $\boldsymbol{X} = \boldsymbol{\Psi}_{\mathrm{dhw}} \boldsymbol{S} \boldsymbol{\Psi}_{\mathrm{idhw}}^\top$, or $\boldsymbol{x} = \boldsymbol{\Psi}\boldsymbol{s}$. We will use this second model for SI-FTI scheme in § 4.5, which allows for an easy characterization of this stronger sparsity prior.

4.4 Coded-Illumination FTI

In this section we focus on a simple modification of the FTI system introduced in § 4.2.2. We introduce a temporal coding of the specimen illumination, *e.g.*, by acting on the global activation of the light source (see Fig. 4.3-top). This is potentially an easy adaptation, specially when the FTI module is combined with, *e.g.*, a confocal microscope where the light illumination can be programmed. We support this modification with the VDS scheme introduced in [61] (see § 2.3). We first describe the principles of the CI-FTI system before explaining how to reconstruct the observed HS volume.

4.4.1 Acquisition Strategy

We here refer to *coded illumination* as a technique that activates the light source in a portion of time slots, *i.e.*, coding temporally the light distribution. Recall that in Nyquist FTI, the light source is illuminating the biological specimen during N_ξ time slots. As will be clear below, from the one-to-one correspondence between the time domain (or time slot), mirror position in the FTI system and OPD value (see § 4.2.2.1), this temporal illumination coding amounts thus to subsample the OPD domain.

Pursuing the context developed in § 4.2.2.1, we assume that the light source delivers constant, but controllable, intensity I^{src}, *i.e.*, defined in (4.4), per second and per unit area on any location of the biological specimen. If each OPD sample is associated with a time slot of τ_ξ second, then the total light exposure per unit area on each location of the

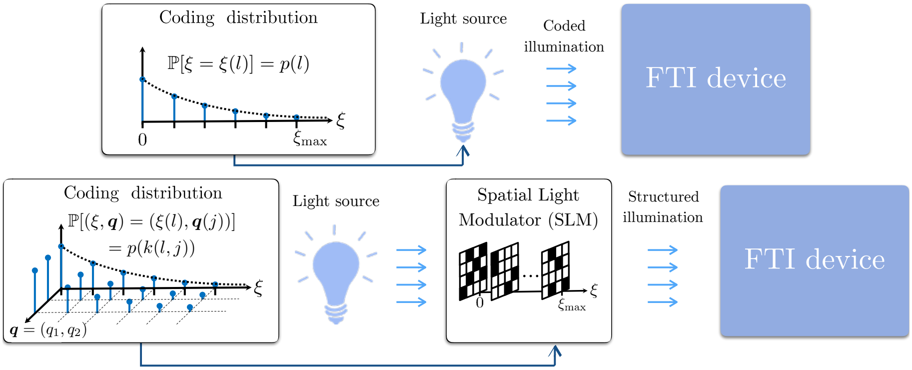

Figure 4.3: Illustration of (top) coded illumination-FTI and (bottom) structured illumination-FTI systems according to the coding distributions (*i.e.*, the pmfs of VDS) established in § 4.4 and § 4.5. For the sake of simplicity only the positive part of OPD axis is shown. In these figures, $\xi(l)$ represents the l^{th} OPD location, *i.e.*, $\xi(l) := (l - N_\xi/2)\Delta_\xi$, and $\boldsymbol{q}(j)$ the j^{th} spatial location in the ordering of the vectorized spatial coordinates.

biological specimen is equal to $N_\xi \tau_\xi I^{\text{src}}$. The idea here is to reduce this exposure to $M_\xi \tau_\xi I^{\text{src}}$ by activating the light source only over $M_\xi < N_\xi$ time slots; these being associated with a subset $\Omega^\xi := \{\omega_1^\xi, \cdots, \omega_{M_\xi}^\xi\}$ of M_ξ (possibly non-unique) OPD samples $\omega_l^\xi \in [\![N_\xi]\!]$ with $l \in [\![M_\xi]\!]$.

For this purpose, we decide to follow the VDS scheme explained in § 2.3; we generate Ω^ξ by randomly drawing its elements, *i.e.*, $\omega_l^\xi \sim_{\text{iid}} \omega^\xi$, for $l \in [\![M_\xi]\!]$, and $\beta_\xi \in [\![N_\xi]\!]$ is a r.v. whose pmf $p(l_\xi) := \mathbb{P}[\beta_\xi = l_\xi]$ is optimized in § 4.4.2 according to the FTI sensing and the HS sparsity prior.

According to the discrete FTI sensing model (see § 4.2.2.2), this selection of OPD samples results in recording the collection of interferograms $\mathcal{Y}\big((l_\xi - N_\xi/2)\Delta_\xi; j_1, j_2\big)$ at OPD indices $l_\xi \in \Omega^\xi$. The acquisition model of CI-FTI then reads

$$\boldsymbol{Y}^{\text{ci}} = \boldsymbol{P}_{\Omega^\xi} \boldsymbol{F}^* \boldsymbol{X} + \boldsymbol{N}, \quad \text{or} \quad \boldsymbol{y}^{\text{ci}} = \boldsymbol{P}_\Omega \boldsymbol{\Phi}_{\text{fti}}^* \boldsymbol{x} + \boldsymbol{n}, \tag{4.16}$$

where $\Omega = \bigcup_{l_{\text{p}}=1}^{N_{\text{p}}} \{N_\xi(l_{\text{p}} - 1) + \Omega^\xi\}$ contains the indices associated with all selected rows of $\boldsymbol{\Phi}_{\text{fti}}^*$, $\boldsymbol{n} = \text{vec}(\boldsymbol{N})$, and $\boldsymbol{N} \in \mathbb{R}^{M_\xi \times N_{\text{p}}}$ models an additive measurement noise.

Following the considerations of § 4.2.3, we assume that $\boldsymbol{N}$ is a random Gaussian noise with variance σ^2, *i.e.*, $N_{l,l'} \sim_{\text{iid}} \mathcal{N}(0,\sigma^2)$ for $l \in [\![M_\xi]\!]$ and $l' \in [\![N_\text{p}]\!]$. Moreover, we suppose that σ is independent of both the number of measurements and the observed HS volume. The validity of this assumption will be confirmed in § 4.8.3 by showing that σ only moderately grows when the intensity of the HS volume strongly increases. Consequently, we set hereafter $\sigma^2 = \sigma^2_\text{fti}$, with σ^2_fti being the variance of the measurement noise on each Nyquist FTI observation, *e.g.*, estimated from an RM estimator [33, 112], or by the calibration of the FTI system.

Remark 4.7. *CI-FTI is the simplest compressive variation of the FTI system. However, it entails some additional concerns, compared to the FTI system, that must be taken into account when considering an actual implementation. First, the subsampling of the OPD domain guided above requires the module controlling the intensity of the light source to be aware of the position of the (continuously) moving mirror inside the MI device, otherwise there will be an error in subsampling OPD domain. Second, the time intervals where the light source is activated must respect the discretized time domain characterized by the fps rate of the imaging sensor. However, these two issues can be solved by a careful calibration process.*

4.4.2 HS Reconstruction Method and Guarantee

Given the noisy CI-FTI measurements in (4.16), we leverage the VDS scheme supported by Thm. 2.4 and propose the following convex optimization problem for recovering the N_hs-voxel HS volume, *i.e.*, $\forall l_\text{p} \in [\![N_\text{p}]\!]$,

$$\hat{\boldsymbol{x}} = \underset{\boldsymbol{u}\in\mathbb{C}^{N_\text{hs}}}{\arg\min}\ \|\boldsymbol{\Psi}_\text{ci}^\top \boldsymbol{u}\|_1 \ \text{ s.t. } \ \|\boldsymbol{D}^\xi(\boldsymbol{y}^\text{ci}_{l_\text{p}} - \boldsymbol{P}_{\Omega^\xi}\boldsymbol{F}^*\boldsymbol{u}_{l_\text{p}})\| \le \varepsilon_\text{ci}\sqrt{M_\xi}, \quad (4.17)$$

where $\Omega^\xi = \{\omega^\xi_1, \cdots, \omega^\xi_{M_\xi}\}$ is randomly generated as described in § 4.4.1, $\boldsymbol{D}^\xi = \operatorname{diag}(\boldsymbol{d}^\xi) \in \mathbb{R}^{M_\xi\times M_\xi}$ is a diagonal matrix such that $d^\xi_l =$

$1/(p(\omega_l^\xi))^{1/2}$, $\boldsymbol{\Psi}_{\rm ci}$ is the sparsity basis specified below, and $\varepsilon_{\rm ci}$ is a bound on the weighted measurement noise level such that (with high probability) $\|\boldsymbol{D}^\xi \boldsymbol{n}_{l_{\rm p}}\| \leq \varepsilon_{\rm ci}\sqrt{M}$ for all $l_{\rm p} \in [\![N_{\rm p}]\!]$. We will characterize it momentarily.

Remark 4.8. *For the sake of consistency with the result of Krahmer and Ward in Thm. 2.4, we do not force the estimation of* (4.17) *to be real-valued, even, or non-negative. Adding those extra prior information do not worsen the theoretical guarantee in* (4.17)*, while, numerically, they improve the quality of the HS data recovery, as they reduce the set of feasible solutions in* (4.17)*.*

In the optimization problem (4.17), while a global fidelity constraint could have been set between the partial interferometric observations $\boldsymbol{y}^{\rm ci}$ and the candidate HS volume $\boldsymbol{u}$, we rather impose $N_{\rm p}$ individual fidelity constraints, one per pixel index $l_{\rm p} \in [\![N_{\rm p}]\!]$. Moreover, concerning the regularizer of (4.17), we consider only a spectral sparsity prior described in § 4.3, *i.e.*, $\boldsymbol{\Psi}_{\rm ci} = \boldsymbol{I}_{N_{\rm p}} \otimes \boldsymbol{\Psi}_{\rm dhw}$. The compressibility of the acquisition is indeed mainly brought on the spectral domain, with an OPD subsampling pattern shared for all spatial locations. As a result (expressed in the next lemma), the choice of these fidelity constraints together with this specific prior allows for the decomposition of (4.17) into small-size problems, for which we can find individual recovery guarantee.

Lemma 4.9. *Problem* (4.17) *can be decoupled into* $N_{\rm p}$ *subproblems*

$$\hat{\boldsymbol{x}}_{l_{\rm p}} = \mathop{\arg\min}_{\boldsymbol{u}\in\mathbb{C}^{N_\xi}} \|\boldsymbol{\Psi}_{\rm dhw}^\top \boldsymbol{u}\|_1 \ \text{ s.t. } \ \|\boldsymbol{D}^\xi(\boldsymbol{y}_{l_{\rm p}}^{\rm ci} - \boldsymbol{P}_{\Omega^\xi}\boldsymbol{F}^*\boldsymbol{u}_{l_{\rm p}})\| \leq \varepsilon_{\rm ci}\sqrt{M_\xi}, \tag{4.18}$$

with $1 \leq l_{\rm p} \leq N_{\rm p}$.

Proof. From Problem (4.17), the separability of the ℓ_1-prior, *i.e.*, $\|\boldsymbol{\Psi}_{\rm ci}^\top\boldsymbol{u}\|_1 = \sum_{l_{\rm p}=1}^{N_{\rm p}} \|\boldsymbol{\Psi}_{\rm dhw}^\top \boldsymbol{u}_{l_{\rm p}}\|_1$, and according to the argument in [98, page 337], the proof is straightforward. □

From N_p sub-problems in (4.18), we can develop a recovery guarantee for the reconstruction of any HS volume $\boldsymbol{X}$ from (4.17), according to the specified sparsity basis. According to Prop. 2.3, given the vector of bounds $\boldsymbol{\kappa}$ on the local coherence between the sparsity and sensing bases ($\boldsymbol{\Psi}_\mathrm{dhw}$ and $\boldsymbol{F}$, respectively), *i.e.*, with $\mu_l^\mathrm{loc}(\boldsymbol{F}^*\boldsymbol{\Psi}_\mathrm{dhw}) \leq \kappa_l$, by selecting $M_\xi \gtrsim \delta^{-2}\|\boldsymbol{\kappa}\|^2 K_\xi \log(\epsilon^{-1})$ OPD indices with respect to the pmf defined in (2.8), the RIP of order K_ξ holds for the matrix $(M_\xi)^{-1/2}\boldsymbol{D}^\xi \boldsymbol{P}_{\Omega^\xi}\boldsymbol{F}^*\boldsymbol{\Psi}_\mathrm{dhw}$ with probability exceeding $1-\epsilon$. Since Thm. 2.4 provides uniform guarantee in the sense that the recovery is ensured for all possible signals, the reconstruction of each column of the HS volume from (4.18) satisfies

$$\|\boldsymbol{x}_{l_\mathrm{p}} - \hat{\boldsymbol{x}}_{l_\mathrm{p}}\| \leq c_1 \frac{\sigma_{K_\xi}(\boldsymbol{\Psi}_\mathrm{dhw}^\top \boldsymbol{x}_{l_\mathrm{p}})_1}{\sqrt{K_\xi}} + c_2\varepsilon_\mathrm{ci}, \quad \forall l_\mathrm{p} \in [\![N_\mathrm{p}]\!],$$

with probability exceeding $1-\epsilon$ with constants c_1 and c_2 specified in Thm. 2.4. Consequently, since $(a+b)^2 \leq 2(a^2+b^2) \leq 2(a+b)^2$ for all $a, b > 0$, we can bound the estimation error of the whole HS image as follows

$$\|\boldsymbol{x} - \hat{\boldsymbol{x}}\| \leq \tfrac{c_1\sqrt{2}}{\sqrt{K_\xi}} \Big(\textstyle\sum_{l_\mathrm{p}=1}^{N_\mathrm{p}} \big(\sigma_{K_\xi}(\boldsymbol{\Psi}_\mathrm{dhw}^\top \boldsymbol{x}_{l_\mathrm{p}})_1\big)^2\Big)^{1/2} + c_2\sqrt{2N_\mathrm{p}}\varepsilon_\mathrm{ci}. \quad (4.19)$$

In order to adjust the noise level ε_ci as a function of known or estimable parameters, *e.g.*, M_ξ, N_ξ and σ_fti, we need to characterize the pmf responsible for the selection of the OPD indices.

Proposition 4.10. *For the context of CI-FTI explained above, we have*

$$\mu_{l_\xi}(\boldsymbol{F}^*\boldsymbol{\Psi}_\mathrm{dhw}) \;\leq\; \kappa_{l_\xi} := \sqrt{2}\min\big\{1, |l_\xi - \frac{N_\xi}{2}|^{-1/2}\big\}, \quad l_\xi \in [\![N_\xi]\!],$$

with $\|\boldsymbol{\kappa}\|^2 \leq 8 + 4\log(\frac{N_\xi}{2}) \lesssim \log(N_\xi)$. *Moreover, the corresponding pmf in* (2.8) *is given by*

$$p(l_\xi) = C_{N_\xi}\ \min\{1, |l_\xi - \frac{N_\xi}{2}|^{-1}\}, \quad l_\xi \in [\![N_\xi]\!], \tag{4.20}$$

where the normalization constants C_{N_ξ} *respects* $2\log(N_\xi/2) < C_{N_\xi}^{-1} < 4 + 2\log(N_\xi/2)$.

For the proof see § 4.9.1. Note that we computed a bound (not the exact values) for the local coherence. Even though this tight bound is not optimal, it defines a meaningful sampling pmf (4.20) and sample-complexity bound (4.22). We will see in § 4.8 that the performance of the proposed pmf is very close to the performance of the optimal pmf computed numerically. The same argument holds for Prop. 4.15 related to the SI-FTI system.

Remark 4.11. *Following this proposition, if the random quantities* $\Omega = \Omega^\xi$ *and* $\boldsymbol{D} = \boldsymbol{D}^\xi$ *defined above are specified by the pmf of Prop. 4.10, we can estimate the level of noise in CI-FTI at all pixels* $l_{\rm p} \in [\![N_{\rm p}]\!]$. *Since the pmf* p *in* (4.20) *corresponds to a VDS scheme of exponent* $\alpha = 1$, *offset* $l_0 = N_\xi/2$ *and constant* $c_\eta = C_{N_\xi}$ *in Remark 2.8, we find* $\rho = C_{N_\xi}^{-1}(N_\xi - (N_\xi/2))/N_\xi \leq 2 + \log(N_\xi/2)$. *Therefore, setting there* $\sigma = \sigma_{\rm fti}$ *and* $\eta(l) = p(l)$ *in Cor. 2.9 with that bound for* ρ*, and using a union bound over all pixels, we get* $\forall l_{\rm p} \in [\![N_{\rm p}]\!]$,

$$\begin{aligned} &\frac{1}{\sqrt{M_\xi}} \|\boldsymbol{D}^\xi(\boldsymbol{y}^{\rm ci}_{l_{\rm p}} - \boldsymbol{P}_{\Omega^\xi}\boldsymbol{F}^*\boldsymbol{x}_{l_{\rm p}})\| \\ &\qquad \leq \varepsilon_{\rm ci}(s) \coloneqq \varepsilon_{\sigma_{\rm fti},s}(N_\xi, M_\xi, 2 + \log(N_\xi/2)), \end{aligned} \tag{4.21}$$

with probability exceeding $1 - 3N_{\rm p}e^{-s/2}$ *and* $\varepsilon_{\sigma,s}$ *defined in* (2.21).

We are now ready to summarize all the analysis in this section and to provide the main result for CI-FTI in the following theorem.

Theorem 4.12 (Coded illumination-FTI). *Given $s > 0$, fix integers K_ξ, N_ξ, $N_{\rm p}$, M_ξ such that $K_\xi \gtrsim \log(N_\xi)$ and*

$$M_\xi \gtrsim K_\xi \log(N_\xi) \log(\epsilon^{-1}). \tag{4.22}$$

Generate M_ξ (possibly non-unique) OPD indices $\Omega^\xi = \{\omega_1^\xi, \cdots, \omega_{M_\xi}^\xi\} \subset [\![N_\xi]\!]$ such that $\omega_l^\xi \sim_{\rm iid} \beta_\xi$ for $l \in [\![M_\xi]\!]$, with β_ξ a r.v. with the pmf (4.20). *Then, given the corresponding noisy CI-FTI measurements $\boldsymbol{y}^{\rm ci}$ in* (4.16), *the HS volume $\boldsymbol{x}$ can be approximated by solving* (4.17) *with the bound $\varepsilon_{\rm ci}(s)$ in* (4.21), *up to an error*

$$\|\boldsymbol{x} - \hat{\boldsymbol{x}}\| \leq \tfrac{c_1\sqrt{2}}{\sqrt{K_\xi}} \Big(\textstyle\sum_{l_{\rm p}=1}^{N_{\rm p}} \big(\sigma_{K_\xi}(\boldsymbol{\Psi}_{\rm dhw}^\top \boldsymbol{x}_{l_{\rm p}})_1\big)^2\Big)^{1/2} + c_2\sqrt{2N_{\rm p}}\varepsilon_{\rm ci}(s), \tag{4.23}$$

and with probability exceeding $1 - \epsilon - 3N_{\rm p}e^{-s/2}$ with constants c_1 and c_2 specified in Thm. 2.4.

Proof. A combination of Thm. 2.4, Lem. 4.9, and Prop. 4.10 with (4.19) completes the proof. □

In words, this theorem states that the pmf in (4.20) associated with a near-optimal sampling strategy, follows a two-sided power-law decay centered at zero-OPD point $\xi = 0$. As discussed in § 4.2.2, zero-OPD point can be determined prior to FTI acquisition process and be used for developing the above-mentioned sampling strategy. Moreover, by leveraging this VDS strategy, the amount of required light exposure for successful HS recovery behaves of the order of K_ξ, *i.e.*, for typical HS volumes $K_\xi \ll N_\xi$. However, following the calculation of $\boldsymbol{\kappa}$ in S 4.9.1, it is seen that the use of a UDS strategy gives $M_\xi \gtrsim N_\xi K_\xi \log(\epsilon^{-1})$, meaning that the reduction in light exposure is not possible in this context.

4.5 Structured Illumination-FTI

SI-FTI is an alternative approach to further reduce the light exposure on a biological specimen compared to CI-FTI. In short, the light distribution is here coded (or *structured*) in both spatial and OPD (or time) domains. We explain hereafter SI-FTI acquisition procedure, the associated HS volume estimation problem, and we deduce theoretical guarantees on the quality of the reconstructed volume.

4.5.1 Acquisition Strategy

SI-FTI allows for the selection of different OPD samples at each specimen location, as opposed to CI-FTI where the selected OPD indices are shared for all locations. We assume in this thesis that this is achieved by spatially structuring (coding) the illumination of the system, *i.e.*, thanks to an SLM [34, 43] in between of the light source and the biological specimen (see Fig. 4.3-bottom).

Remark 4.13. *For the sake of simplicity, we suppose that the SLM and the imaging sensor have identical resolutions (*i.e.*, $N_{\rm p}$ pixels for both) and that there are perfect alignment and scaling between the discrete illumination pattern and the pixel grid of the camera,* e.g.*, using an appropriate optical magnification system (not represented in Fig. 4.2). Therefore, one coded location on a thin, still, 2-D biological specimen can be identified with one pixel location of the camera. In other words, masking one location of the biological specimen over an area δ_S fixed by the SLM pixel pitch corresponds to blocking the light received by a single-pixel of the camera. Additionally, we assume the SLM does not reduce the light intensity when its pixels are activated (*i.e.*, the system has 100 % light throughput).*

To understand the potential benefit of SI-FTI let us observe that the Nyquist FTI scheme can be recovered from such a system simply by activating all SLM pixels for all OPD samples. In this case, if we assume again that the light source delivers constant intensity $I^{\rm src}$ defined in (4.4) per second and per unit area on any location of the specimen, the total light exposure at every OPD point on the whole biological specimen is constant and equal to $N_{\rm p}\delta_S\tau_\xi I^{\rm src}$, where τ_ξ corresponds to the duration of each OPD sample and δ_S is the SLM pixel area. Using structured illumination in SI-FTI, if $\tilde{M}_{l_\xi} < N_{\rm p}$ spatial locations are exposed on the specimen at the $l_\xi^{\rm th}$ OPD sample with $l_\xi \in [\![N_\xi]\!]$, this exposure can be reduced to $\tilde{M}_{l_\xi}\delta_S\tau_\xi I^{\rm src}$ for this OPD sample. Therefore, the total light exposure undergone by the specimen during a Nyquist FTI acquisition, *i.e.*, $N_{\rm hs}\delta_S\tau_\xi I^{\rm src}$ with $N_{\rm hs} = N_\xi N_{\rm p}$, is decreased to $M\delta_S\tau_\xi I^{\rm src}$ in SI-FTI, where $M = \sum_{l_\xi=1}^{N_\xi}\tilde{M}_{l_\xi}$. The question is of course to be able to reconstruct the HS volume from such a partial FTI observations.

Let us now turn to SI-FTI sensing model. We follow the general random VDS scheme of § 2.3, with special care to integrate the spatiospectral geometry of SI-FTI. We denote by $p(l_{\rm p}, l_\xi) := \mathbb{P}[(\beta_{\rm p}, \beta_\xi) = (l_{\rm p}, l_\xi)]$ a bivariate pmf determining the random activation of the $l_{\rm p}^{\rm th}$ spatial location at the $l_\xi^{\rm th}$ OPD point, *i.e.*, the pmf of a bivariate r.v. $\boldsymbol{\beta} = (\beta_{\rm p}, \beta_\xi)$ for $\beta_{\rm p} \in [\![N_{\rm p}]\!]$ and $\beta_\xi \in [\![N_\xi]\!]$. In this context, we can generate a random set $\bar{\Omega}$ with M (possibly non-unique) elements from $\bar{\Omega} = \{\boldsymbol{\omega}_1, \cdots, \boldsymbol{\omega}_M\}$ with $\boldsymbol{\omega}_l = (\omega_l^{\rm p}, \omega_l^\xi) \sim_{\rm iid} \boldsymbol{\beta}$ and $l \in [\![M]\!]$.

The sensing model then reads

$$y_l = \boldsymbol{P}_{\{\omega_l^\xi\}}\boldsymbol{F}^*\boldsymbol{X}\boldsymbol{P}_{\{\omega_l^{\rm p}\}}^\top + n_l, \quad \forall l \in [\![M]\!],$$

with n_l modeling an additive measurement noise. As for CI-FTI, we assume this noise Gaussian, *i.e.*, $n_l \sim_{\rm iid} \mathcal{N}(0, \sigma_{\rm fti}^2)$ with variance $\sigma_{\rm fti}^2$ that can be estimated, *e.g.*, from the Nyquist measurements or by system calibration.

To reach more compact notation, we adopt a purely 1-D random sensing model. We consider the set $\Omega = \{\omega_1, \cdots, \omega_M\} \subset [\![N_{\rm hs}]\!]$ generated from M (scalar) r.v.s $\omega_l \sim_{\rm iid} \beta$ (with $l \in [\![M]\!]$), where the r.v. $\beta \in N_{\rm hs}$

is defined from the pmf

$$p(l_{\rm hs}) \coloneqq \mathbb{P}[\beta = l_{\rm hs}] = \mathbb{P}[N_\xi(\beta_{\rm p} - 1) + \beta_\xi = l_{\rm hs}], \quad \forall l_{\rm hs} \in [\![N_{\rm hs}]\!], \quad (4.24)$$

with $\boldsymbol{\beta} = (\beta_{\rm p}, \beta_\xi)$ the bivariate r.v. defined above from the pmf $p(l_{\rm p}, l_\xi)$. In this case, for each $l_{\rm p} \in [\![N_{\rm p}]\!]$,

$$\Omega^{l_{\rm p}} \;\coloneqq\; \big(\Omega \cap [N_\xi(l_{\rm p} - 1) + 1, N_\xi l_{\rm p}]\big) \;-\; N_\xi(l_{\rm p} - 1),$$

is the (possibly empty[8]) set of OPD samples selected at the $l_{\rm p}^{\rm th}$ pixel.

SI-FTI then amounts to

$$\boldsymbol{y}_{l_{\rm p}} = \boldsymbol{P}_{\Omega^{l_{\rm p}}} \boldsymbol{F}^* \boldsymbol{X} \boldsymbol{P}_{\{l_{\rm p}\}}^\top + \boldsymbol{n}_{l_{\rm p}}, \; \forall l_{\rm p} \in [\![N_{\rm p}]\!], \quad (4.25)$$

or equivalently,

$$\boldsymbol{y}^{\rm si} = [\boldsymbol{y}_1^\top, \cdots, \boldsymbol{y}_{N_{\rm p}}^\top]^\top = \boldsymbol{P}_\Omega \boldsymbol{\Phi}_{\rm fti}^* \boldsymbol{x} + \boldsymbol{n}, \quad (4.26)$$

where $\boldsymbol{x} \coloneqq \mathrm{vec}(\boldsymbol{X})$ and $\boldsymbol{n} \coloneqq [\boldsymbol{n}_1^\top, \cdots, \boldsymbol{n}_{N_{\rm p}}^\top]^\top \in \mathbb{R}^M$.

Recall that we can go back and forth between the 1-D and the 2-D index representations $l_{\rm hs} \in [\![N_{\rm hs}]\!]$ and $(l_\xi, l_{\rm p}) \in [\![N_\xi]\!] \times [\![N_{\rm p}]\!]$, respectively, using the rule $l_{\rm hs} \xrightleftharpoons{N_\xi, N_{\rm p}} (l_\xi, l_{\rm p})$.

The pmf p in (4.24) defining SI-FTI sensing is considered hereafter as a degree of freedom that we are going to relate to the VDS scheme formulated in Thm. 2.4. This will allow us to reach an optimized structured illumination strategy, as presented in Thm 4.16.

Remark 4.14. *SI-FTI brings the FTI to a more complex CS system, compared to CI-FTI, with four main practical challenges. First, the SLM module must be aware of the position of the (continuously) moving mirror inside the MI device, since at each OPD point one SLM is programmed as instructed above. Second, similar to the practical challenge in CI-FTI (see Remark 4.7),*

[8] With the convention that $\boldsymbol{u}_\emptyset = \emptyset$ for any vector $\boldsymbol{u}$.

the time intervals where the SLMs are programmed must respect the discretized time domain characterized by the fps rate of the imaging sensor. Third, misalignment and disposition of the pixel grid of the SLM with respect to the 2-D imaging sensor causes blurring effect in the reconstructed HS volumes. These three issues can be solved by a careful calibration process. Fourth, the issue that cannot be mitigated by calibration is the non-identical pixel size of the SLM and imaging sensor. Let us consider two cases: (i) the SLM's pixels are smaller than imaging sensors' pixels: each pixel of the imaging sensor integrates several intensities associated with a group of neighboring SLM's pixels, (ii) the SLM's pixels are larger than imaging sensors' pixels: a group of neighboring pixels on the imaging sensor will receive the same intensity associated with one SLM's pixel. We suggest to address this issue by introducing an integration operator in the sensing model (4.25).

4.5.2 HS Reconstruction Method and Guarantee

Given the noisy SI-FTI measurements as in (4.26), an HS volume $\boldsymbol{x}$ with $N_{\rm hs}$ voxels can be reconstructed via the convex optimization problem

$$\hat{\boldsymbol{x}} = \underset{\boldsymbol{u}\in\mathbb{C}^{N_{\rm hs}}}{\arg\min}\|\boldsymbol{\Psi}_{\rm si}^{\top}\boldsymbol{u}\|_1 \ \text{s.t.}\ \|\boldsymbol{D}(\boldsymbol{y}^{\rm si} - \boldsymbol{P}_{\Omega}\boldsymbol{\Phi}_{\rm fti}^{*}\boldsymbol{u})\| \leq \varepsilon_{\rm si}\sqrt{M}, \tag{4.27}$$

where $\Omega = \{\omega_1, \cdots, \omega_M\} \subset [\![N_{\rm hs}]\!]$ is randomly generated according to the pmf of (4.24), $\boldsymbol{\Psi}_{\rm si}$ is the sparsity basis, $\varepsilon_{\rm si}$ must be such that $\|\boldsymbol{D}\boldsymbol{n}\| \leq \varepsilon_{\rm si}\sqrt{M}$ with high probability, and $\boldsymbol{D} = \mathrm{diag}(\boldsymbol{d}) \in \mathbb{R}^{M\times M}$ with $d_l = 1/(p(\omega_l))^{1/2}$ for $l \in [\![M]\!]$.

For generality of our model, we regularize Problem (4.27) with the joint spatiospectral HS sparsity model described in § 4.3, *i.e.*, $\boldsymbol{\Psi}_{\rm si} = \boldsymbol{\Psi}_{\rm idhw} \otimes \boldsymbol{\Psi}_{\rm dhw}$. Contrary to the CI-FTI optimization Problem (4.17), Problem (4.27) cannot be decoupled into sub-problems since $\|\boldsymbol{\Psi}_{\rm si}^{\top}\boldsymbol{u}\|_1$ is not separable in $\{\boldsymbol{u}_{l_{\rm p}} : l_{\rm p} \in [\![N_{\rm p}]\!]\}$ and the random set Ω detailed above is not separable in the spatial and OPD domains. Therefore, we tackle the reconstruction problem of the full HS volume at once.

Having set the sparsity basis, we can now adjust the pmf (4.24) determining Ω from Thm. 2.4. According to this theorem, the preconditioned matrix $\frac{1}{\sqrt{M}}\boldsymbol{D}\boldsymbol{P}_\Omega\boldsymbol{\Phi}^*_{\rm fti}\boldsymbol{\Psi}_{\rm si}$ respects the RIP of order K with probability exceeding $1-\epsilon$ if

$$M \gtrsim \delta^{-2}\|\boldsymbol{\kappa}\|^2 K\log(\epsilon^{-1})$$

SI-FTI measurements are recorded with respect to the pmf $p(l_{\rm hs}) := \kappa^2_{l_{\rm hs}}/\|\boldsymbol{\kappa}\|^2$ for $l_{\rm hs} \in [\![N_{\rm hs}]\!]$, where the vector $\boldsymbol{\kappa} \in \mathbb{R}^{N_{\rm hs}}_+$ is a bound for the local coherence $\mu^{\rm loc}_{l_{\rm hs}}(\boldsymbol{\Phi}^*_{\rm fti}\boldsymbol{\Psi}_{\rm si})$ with $l_{\rm hs} \in [\![N_{\rm hs}]\!]$. The next proposition (proved in § 4.9.2) bounds this local coherence, and thus determines the pmf p.

Proposition 4.15. *For the context of SI-FTI explained above, we have*

$$\mu^{\rm loc}_{l_{\rm hs}}(\boldsymbol{\Phi}^*_{\rm fti}\boldsymbol{\Psi}_{\rm si}) \le \kappa_{l_{\rm hs}} := \frac{\sqrt{2}}{2}\min\Big\{1, \frac{1}{\sqrt{|l_\xi - (N_\xi/2)|}}\Big\},\ l_{\rm hs} \in [\![N_{\rm hs}]\!], \tag{4.28}$$

with $\|\boldsymbol{\kappa}\|^2 \le N_{\rm p}(2+\log(N_\xi/2)) \lesssim N_{\rm p}\log N_\xi$ *and the relation* $l_{\rm hs} \xrightleftharpoons{N_\xi, N_{\rm p}} (l_\xi, l_{\rm p})$. *In this case,* (2.8) *reads*

$$p(l_{\rm hs}) = \frac{C_{N_\xi}}{N_{\rm p}}\min\{1, |l_\xi - (N_\xi/2)|^{-1}\},\ l_{\rm hs} \in [\![N_{\rm hs}]\!], \tag{4.29}$$

where the normalization constant C_{N_ξ} *respects* $2\log(N_\xi/2) < C^{-1}_{N_\xi} < 4+2\log(N_\xi/2)$.

If $\frac{1}{\sqrt{M}}\boldsymbol{D}\boldsymbol{P}_\Omega\boldsymbol{\Phi}^*_{\rm fti}\boldsymbol{\Psi}_{\rm si}$ respects the RIP with the pmf specified in the previous proposition, Thm. 2.4 shows that the estimation error achieved by (4.27) can be bounded as

$$\|\boldsymbol{x} - \hat{\boldsymbol{x}}\| \le c_1\frac{\sigma_K(\boldsymbol{\Psi}^\top_{\rm si}\boldsymbol{x})_1}{\sqrt{K}} + c_2\varepsilon_{\rm si}. \tag{4.30}$$

Moreover, following Prop. 4.15, Cor. 2.9 and Remark 2.8, we can estimate the level $\varepsilon_{\rm si}$ of a Gaussian noise in the SI-FTI model (4.26). Indeed, we can compute the bound (2.18) associated with the pmf (4.29) since, for any integer $q \geq 1$, we easily prove that

$$\tfrac{1}{N_{\rm hs}} \big(\mathbb{E}_\beta\, p(\beta)^{-q}\big)^{1/q} = \tfrac{1}{N_\xi} \big(\mathbb{E}_\gamma\, p'(\gamma)^{-q}\big)^{1/q},$$

where $\beta \in [\![N_{\rm hs}]\!]$ is a r.v. with the pmf p of (4.29), and $\gamma \in [\![N_\xi]\!]$ is a r.v. with pmf $p'(l) := C_{N_\xi} \min\{1, |l - (N_\xi/2)|^{-1}\}$. From Remark 2.8, p' is thus a VDS scheme with exponent $\alpha = 1$ and offset $l_0 = N_\xi/2$. We can thus set $\rho = 2 + \log(N_\xi/2) > C_{N_\xi}^{-1}/2 = C_{N_\xi}^{-1}(N_\xi - l_0)/N_\xi$, so that

$$\frac{1}{M}\|\boldsymbol{D}\boldsymbol{n}\|^2 \leq \varepsilon_{\rm si}^2(s) := \varepsilon^2_{\sigma_{\rm fti},s}(N_{\rm hs}, M, 2 + \log(N_\xi/2)), \tag{4.31}$$

with probability exceeding $1 - 3e^{-s/2}$ and $\varepsilon_{\sigma,s}$ defined in (2.21).

We are now ready to summarize the complete analysis of SI-FTI in the following theorem.

Theorem 4.16 (Structured Illumination-FTI)**.** *Given $s > 0$, fix integers K, $N_{\rm hs} = N_\xi N_{\rm p}$ such that $K \gtrsim \log(N_{\rm hs})$ and*

$$M \gtrsim N_{\rm p} K \log(N_\xi) \log(\epsilon^{-1}). \tag{4.32}$$

Generate M random (non-unique) indices associated with a (1-D) index set $\Omega = \{\omega_1, \cdots, \omega_M\}$ such that $\omega_l \sim_{\rm iid} \beta$ for $l \in [\![M]\!]$, with β a r.v. with the pmf (4.29)*. Then, given the noisy SI-FTI measurements $\boldsymbol{y}^{\rm si}$ in* (4.26)*, the HS volume $\boldsymbol{x}$ can be approximated by solving* (4.27) *with the bound $\varepsilon_{\rm si}(s)$ in* (4.31)*, up to an error*

$$\|\boldsymbol{x} - \hat{\boldsymbol{x}}\| \leq c_1 \frac{\sigma_K(\boldsymbol{\Psi}_{\rm si}^\top \boldsymbol{x})_1}{\sqrt{K}} + c_2 \varepsilon_{\rm si}(s), \tag{4.33}$$

where the constants c_1 and c_2 are specified in Thm. 2.4, with probability exceeding $1 - \epsilon - 3e^{-s/2}$.

Proof. combination of Thm. 2.4 and Prop. 4.15 with (4.31), (4.30) completes the proof. □

Note that according to (4.29), the spatial locations are selected (and thus exposed) uniformly at random; while for a fixed spatial location the probability of exposing that location obeys a two-sided power-law decay distribution centered at OPD origin $N_\xi/2$. In addition, since we can ensure $N_{\rm p}K \ll N_{\rm hs} = N_{\rm p}N_\xi$ with a low best K-term approximation error $\sigma_K(\boldsymbol{\Psi}_{\rm si}^\top \boldsymbol{x})_1/\sqrt{K}$ under the hypotheses made on the observed HS volume, the light exposure on the biological specimen can be significantly reduced thanks to the VDS strategy (4.26); the UDS strategy would require $M \gtrsim N_{\rm hs}K\log(\epsilon^{-1})$, which is equivalent to overexposing the specimen.

4.6 Constrained-Exposure Coding

As aforementioned, in fluorescence spectroscopy, the type of the injected fluorescent dyes (and therefore their tolerance to light exposure) is known. In this section, assuming that photo-bleaching[9] for a fluorescent dye depends only on the total light exposure, which it has been subject to, and supposing that the maximum light exposure $I_{\rm max}$ that a fluorescent dye can tolerate is known, we adapt the proposed CI- and SI-FTI schemes by ensuring that the total light exposure $I_{\rm tot}$ on each spatial location of a biological specimen is constant and smaller than $I_{\rm max}$, whatever the number of compressive observations.

From the description in § 4.2.2.1, in the case of the Nyquist FTI system, if the total acquisition time is $\tau_{\rm hs} > 0$, we have $I_{\rm tot} = \tau_{\rm hs} I^{\rm src} = \tau_{\rm hs}\int_0^{+\infty} |E_0^{\rm src}(\nu)|^2\,{\rm d}\nu$. Since N_ξ OPD samples are recorded, each location thus receives an intensity of $I_{\rm opd} = I_{\rm tot}/N_\xi$ per OPD sample. If we fix the value $I_{\rm tot}$, we show below that for CI- and SI-FTI schemes, the light source intensity, and thus the intensity per OPD sample, can be increased.

[9] We suppose here that the intensity of the light illuminated by a fluorescent dye is constant during CI- and SI-FTI experiments. As a more realistic model, however, the intensity of the fluorescent light would exponentially decrease as a function of the received intensity.

This leads to an improved Measurement-to-Noise Ratio (MNR), *i.e.*,

$$\mathrm{MNR} := 10 \log_{10} \frac{\|\boldsymbol{y}\|^2}{\|\boldsymbol{n}\|^2}, \tag{4.34}$$

if the measurement noise is assumed to not depend on the light intensity (*i.e.*, under high-photon counting assumption, see § 4.2.3).

4.6.1 Constrained-exposure CI-FTI

We want to adjust the intensity $I'_{\rm opd}$ received per second and per unit area by each location of the biological specimen (when illuminated), assuming that the total light exposure of the specimen per unit area is fixed (constrained).

In CI-FTI, $I_{\rm tot} = M_\xi \tau_\xi I'_{\rm opd}$, with τ_ξ is the constant duration of each time slot fixed by the acquisition. In constrained-exposure CI-FTI, we keep $I_{\rm tot}$ constant so that it matches the total light exposure of the Nyquist FTI where $M_\xi = N_\xi$, *i.e.*, $I_{\rm tot} = N_\xi \tau_\xi I_{\rm opd}$. This thus imposes the relation $I'_{\rm opd} = (N_\xi/M_\xi) I_{\rm opd}$.

The light source intensity can thus be amplified by a factor of N_ξ/M_ξ while still preventing photo-bleaching if $I_{\rm tot} < I_{\rm max}$; from the assumption of linear intensity scaling made in (4.7), the intensity of the light outgoing from the biological specimen then undergoes the same amplification. As explained in § 4.4.1, we also assume that the additive measurement noise is independent of this increase of intensity. This assumption is experimentally well-verified (see § 4.8.3).

Therefore, under the constant light exposure constraint, the acquisition model of CI-FTI in (4.16) reads

$$\boldsymbol{Y}'^{\rm ci} = \frac{N_\xi}{M_\xi} \boldsymbol{P}_{\Omega^\xi} \boldsymbol{F}^* \boldsymbol{X} + \boldsymbol{N} \;=\; \frac{N_\xi}{M_\xi} (\boldsymbol{P}_{\Omega^\xi} \boldsymbol{F}^* \boldsymbol{X} + \frac{M_\xi}{N_\xi} \boldsymbol{N}) \;=:\; \frac{N_\xi}{M_\xi} \tilde{\boldsymbol{Y}}^{\rm ci}. \tag{4.35}$$

The last equality in (4.35) thus shows that $\boldsymbol{Y}'^{\rm ci} = N_\xi/M_\xi \tilde{\boldsymbol{Y}}^{\rm ci}$, with $\tilde{\boldsymbol{Y}}^{\rm ci}$ being the measurements that would be acquired from (4.16) by attenuating the noise $\boldsymbol{N}$ by a factor of M_ξ/N_ξ.

Therefore, in this constrained-exposure context, we can recover $\boldsymbol{X}$ from (4.17) by computing $\tilde{\boldsymbol{Y}}^{\rm ci} = (\tilde{\boldsymbol{y}}_1, \cdots, \tilde{\boldsymbol{y}}_{N_{\rm p}}) = (M_\xi/N_\xi)\boldsymbol{Y}'^{\rm ci}$ and replacing $\sigma^2_{\rm fti} \leftarrow (M_\xi/N_\xi)\sigma^2_{\rm fti}$ in the evaluation of $\varepsilon_{\rm ci}(s)$ in (4.21). In other words, as expected, by fixing the total light exposure, the MNR (4.34) is boosted when the number of measurements decreases. This effect is verified numerically in § 4.8.

Under the conditions of Thm. 4.12, compared to (4.23), we also get an improved error bound

$$\|\boldsymbol{x}-\hat{\boldsymbol{x}}\| \leq \tfrac{c_1\sqrt{2}}{\sqrt{K_\xi}} \Big(\textstyle\sum_{l_{\rm p}=1}^{N_{\rm p}} \big(\sigma_{K_\xi}(\boldsymbol{\Psi}_{\rm dhw}^\top \boldsymbol{x}_{l_{\rm p}})_1\big)^2\Big)^{1/2} + c_2 \tfrac{\sqrt{2N_{\rm p}}M_\xi}{N_\xi}\varepsilon_{\rm ci}(s). \quad (4.36)$$

4.6.2 Constrained-exposure SI-FTI

Compared to CI-FTI, the total light exposure received on each specimen location varies spatially. From the randomness of the structured illumination pattern described in § 4.5, the $l_{\rm p}^{\rm th}$ specimen location is indeed highlighted during $M_{l_{\rm p}} = |\Omega_{l_{\rm p}}|$ time slots, *i.e.*, a r.v. determined by the pmf $p(l_{\rm hs})$ in (4.29). Despite this variability, we can compute a tight (worst case) upper bound on all $\{M_{l_{\rm p}} : l_{\rm p} \in [\![N_{\rm p}]\!]\}$ that holds with arbitrarily high probability. This bound can then be used to constrain the light exposure, *i.e.*, to ensure that no specimen location will be over-exposed.

At each independent random draw of the r.v. β with pmf p, whose M draws populate Ω in Thm. 4.16, the probability of illuminating the $l_{\rm p}^{th}$ specimen location, irrespective of the OPD index, is the marginal pmf $\sum_{l_\xi \in [\![N_\xi]\!]} p(l_{\rm hs}) = 1/N_{\rm p}$; $M_{l_{\rm p}}$ is thus a Binomial r.v. with M trials and success probability $1/N_{\rm p}$. Consequently, $\mathbb{E}M_{l_{\rm p}} = M/N_{\rm p}$ and from Bernstein inequality, we have[10]

$$\mathbb{P}[M_{l_{\rm p}} > \frac{M}{N_{\rm p}} + t] \leq \exp(-\frac{1}{2}t^2(\frac{M}{N_{\rm p}} + \frac{t}{3})^{-1}),$$

[10]Remark that the event $M_{l_{\rm p}} < \frac{M}{N_{\rm p}} - t$, which is useless here, holds with the same probability bound.

for all $t > 0$.

Moreover, by union bound and applying the rescaling $t \leftarrow \frac{M}{N_{\rm p}} t$, this concentration is uniform for all spatial locations, *i.e.*,

$$\mathbb{P}[M_{l_{\rm p}} > \frac{M}{N_{\rm p}}(1+t)] \leq N_{\rm p} \exp(-\frac{3}{2}\, t^2 \frac{M}{N_{\rm p}}(3+t)^{-1}), \;\text{ for all } l_{\rm p} \in [\![N_{\rm p}]\!]. \tag{4.37}$$

Note that this holds true despite the dependence of the r.v.s $\{M_{l_{\rm p}} : l_{\rm p} \in [\![N_{\rm p}]\!]\}$ induced from the relation $\sum_{l_{\rm p}} M_{l_{\rm p}} = M$. Finally, with the change of variable $\zeta = N_{\rm p} \exp(-\frac{3}{2}\, t^2 \frac{M}{N_{\rm p}}(3+t)^{-1})$, *i.e.*, $t = \frac{1}{2}(t_0 + \sqrt{t_0^2 + 12 t_0})$ with $t_0 := \frac{2N_{\rm p}}{3M} \log(\frac{N_{\rm p}}{\zeta})$, (4.37) involves

$$\mathbb{P}[M_{l_{\rm p}} > \frac{M}{N_{\rm p}}(1 + \frac{t_0 + \sqrt{t_0^2 + 12 t_0}}{2})] \leq \zeta, \;\text{ for all } l_{\rm p} \in [\![N_{\rm p}]\!]. \tag{4.38}$$

Therefore, despite the spatial variability of $M_{l_{\rm p}}$, we can set a low failure probability ζ and adjust the light exposure by relying on the fact that

$$M_{l_{\rm p}} < \bar{M}(\zeta) := \frac{M}{N_{\rm p}}(1 + \frac{t_0 + \sqrt{t_0^2 + 12 t_0}}{2}), \tag{4.39}$$

for all $l_{\rm p} \in [\![N_{\rm p}]\!]$ with probability exceeding $1 - \zeta$.

Let $I'_{\rm opd}$ denotes the light intensity received per second and per unit area by each location of the biological specimen when illuminated. The light exposure per unit area on each location $l_{\rm p}$ does not exceed $I_{\rm tot} = \bar{M} I'_{\rm opd} \tau_\xi$, and photo-bleaching does not occur if this quantity is smaller than $I_{\max}$. Matching the constrained $I_{\rm tot}$ with the light exposure $N_\xi \tau_\xi I_{\rm opd}$ of the Nyquist FTI scenario (*i.e.*, full OPD sampling), we thus get the light amplification rule

$$I'_{\rm opd} = (N_\xi / \bar{M}) I_{\rm opd} = \frac{N_{\rm hs}}{M}(1 + \frac{t_0 + \sqrt{t_0^2 + 12 t_0}}{2})^{-1} I_{\rm opd}.$$

Thereby, we can increase the intensity of the light source by a factor of $N_\xi / \bar{M}$, for a fixed parameter ζ independent from other parameters.

From the linear intensity scaling assumption made in (4.7), and assuming that the measurement noise is independent of this scaling, the acquisition model of constrained-exposure SI-FTI, with respect to (4.26), then reads

$$\boldsymbol{y}'^{\rm si} = \frac{N_\xi}{\bar{M}}\,\boldsymbol{P}_\Omega \boldsymbol{\Phi}^*_{\rm fti}\boldsymbol{x} + \boldsymbol{n} = \frac{N_\xi}{\bar{M}}\big(\boldsymbol{P}_\Omega \boldsymbol{\Phi}^*_{\rm fti}\boldsymbol{x} + \frac{\bar{M}}{N_\xi}\boldsymbol{n}\big) := \frac{N_\xi}{\bar{M}}\tilde{\boldsymbol{y}}^{\rm si}. \tag{4.40}$$

Similarly to the discussion for constrained-exposure CI-FTI, (4.40) means that $\boldsymbol{y}'^{\rm si} = (N_\xi/\bar{M})\,\tilde{\boldsymbol{y}}^{\rm si}$, with $\tilde{\boldsymbol{y}}^{\rm si}$ being the measurements that would be acquired from (4.26) by attenuating the noise $\boldsymbol{n}$ by a factor of $(\bar{M}/N_\xi) = (M/N_{\rm hs})(1 + \frac{t_0+\sqrt{t_0^2+12t_0}}{2})$.

Finally, by union bound over the events ensuring (4.39) and the statement of Thm. 4.16, with probability exceeding $1-\epsilon-3e^{-s/2}-\zeta$, the recovery guarantee for the reconstruction of $\boldsymbol{x}$ from $\tilde{\boldsymbol{y}}^{\rm si}$ in constrained-exposure SI-FTI model is

$$\|\boldsymbol{x}-\hat{\boldsymbol{x}}\| \le \frac{c_1\sigma_K(\boldsymbol{\Psi}_{\rm si}^\top \boldsymbol{x})_1}{\sqrt{K}} + c_2\frac{M}{N_{\rm hs}}(1+\frac{t_0+\sqrt{t_0^2+12t_0}}{2})\,\varepsilon_{\rm si}(s), \tag{4.41}$$

where $t_0 = t_0(\zeta)$ is defined above, and replacing $\sigma^2_{\rm fti} \leftarrow (\bar{M}/N_\xi)\sigma^2_{\rm fti}$ in the evaluation of $\varepsilon_{\rm si}(s)$ in (4.31).

Note that, the second term of the right-hand side of (4.41) in SI-FTI, as well as the second term of the right-hand side of (4.36) in CI-FTI, increase linearly with the number of compressive FTI measurements. Thus, given the constrained-exposure budget, recording full FTI measurements does not yield the best recovery quality; the optimum number of measurements is the outcome of a trade-off between the best K-term approximation of the HS volume and the noise level.

4.7 An Improved Subsampling Strategy for CI-FTI in Fluorescence Spectroscopy

Having introduced a provably efficient illumination coding for CI-FTI in § 4.4, we here aim at further improvement of that coding strategy specifically for fluorescence spectroscopy application. The VDS strategy in Thm. 4.12 provides a uniform recovery guarantee, *i.e.*, it holds for recovering all HS data. In fluorescence spectroscopy, however, we are not dealing with all types of HS data. Since the type (or at least the class) of the fluorochromes used in an experiment are known, one can question:

Can we extract the structure of the fluorochromes spectral signatures and adjust the subsampling strategy accordingly?

To answer this question we here leverage the MDS framework described in § 2.4, which has been successfully applied in the context of MRI and also other fluorescence spectroscopy experiments [99]. Our approach is similar to the MDS for Hadamard-Haar system in § 3.3, except that we here study the MDS for Fourier-Haar (and Fourier-Fourier) where the sensing bases is set to the 1-D discrete Fourier basis and the sparsity basis is set to the DHW (respectively, discrete Fourier) basis.

Our goal also entails studying the fluorochrome dataset [1] and the spectra therein; see § 4.8.4 for the details. Conversely to [99] where the spectral dimension is scanned sequentially and the MDS is applied on the spatial dimension, our approach applies the MDS on the Fourier transform of the spectra.

Recall from § 4.4 that the idea of CI-FTI is to activate the light source only over $M_\xi < N_\xi$ time slots associated with a subset $\Omega^\xi = \{\omega_1^\xi, \cdots, \omega_{M_\xi}^\xi\}$. Unlike the VDS-based CI-FTI, in the MDS-based CI-FTI we are not concerned by the pmf associated with selecting the elements of Ω^ξ. Moreover, since the MDS scheme does not allow repetition in the subsampling set, the subsampling set Ω^ξ will have M_ξ unique elements.

Accordingly, we follow the same acquisition model as in (4.16), except that Ω^ξ is now a $(\mathcal{W}, \boldsymbol{m})$-MDS subsampling, which will be designed momentarily.

4.7.1 HS Reconstruction Method and Guarantee

Given the noisy CI-FTI measurements in (4.16) with Ω^ξ a $(\mathcal{W}, \boldsymbol{m})$-MDS subsampling set, we leverage the MDS scheme supported in Thm. 2.6 and propose the following convex optimization problem for recovering the $N_{\rm hs}$-voxel HS volume, *i.e.*, $\forall l_{\rm p} \in [\![N_{\rm p}]\!]$,

$$\hat{\boldsymbol{x}} = \operatorname*{arg\,min}_{\boldsymbol{u} \in \mathbb{C}^{N_{\rm hs}}} \|\boldsymbol{\Psi}_{\rm ci}^\top \boldsymbol{u}\|_1 \ \text{s.t.}\ \|\boldsymbol{y}_{l_{\rm p}}^{\rm ci} - \boldsymbol{P}_{\Omega^\xi} \boldsymbol{F}^* \boldsymbol{u}_{l_{\rm p}}\| \le \varepsilon_{\rm ci}, \tag{4.42}$$

where, compared to (4.17), $\varepsilon_{\rm ci}$ is now a bound on the measurement noise level, *i.e.*, $\|\boldsymbol{n}_{l_{\rm p}}\| \le \varepsilon_{\rm ci}$ for all $l_{\rm p} \in [\![N_{\rm p}]\!]$. We assume that $\boldsymbol{N} = [\boldsymbol{n}_1^\top, \cdots, \boldsymbol{n}_{N_{\rm p}}^\top]^\top \in \mathbb{R}^{M_\xi \times N_{\rm p}}$ is a random iid Gaussian noise with variance $\sigma_{\rm fti}^2$, *i.e.*, $n_{l,l'} \sim_{\rm iid} \mathcal{N}(0, \sigma^2)$ for $l \in [\![M_\xi]\!]$ and $l' \in [\![N_{\rm p}]\!]$. Moreover, in this section we assume $\boldsymbol{\Psi}_{\rm ci} = \boldsymbol{I}_{N_{\rm p}} \otimes \boldsymbol{\Psi}_\nu$, where $\boldsymbol{\Psi}_\nu \in \{\boldsymbol{F}_{N_\xi}, \boldsymbol{\Psi}_{\rm dhw}\}$ is set to either the 1-D discrete Fourier or DHW basis.

Remark 4.17. *Since, for $l \in [\![M_\xi]\!]$ and $l' \in [\![N_{\rm p}]\!]$, $n_{l,l'}^2$ is a χ^2 distribution, we have $\|\boldsymbol{n}_l\|^2 \le \sigma_{\rm fti}^2 (M_\xi + \sqrt{M_\xi}\sqrt{s/2} + s)$ with probability exceeding $1 - \exp(-s/2)$ (see,* e.g.*, [64, Lem. 1]). Therefore, in view of the noise bound $\varepsilon_{\rm ci}$ in* (4.42)*, by using a union bound over all pixels, we get $\forall l_{\rm p} \in [\![N_{\rm p}]\!]$,*

$$\|\boldsymbol{y}_{l_{\rm p}}^{\rm ci} - \boldsymbol{P}_{\Omega^\xi} \boldsymbol{F}^* \boldsymbol{u}_{l_{\rm p}}\| \le \varepsilon_{\rm ci}(s) := \sigma_{\rm fti} (M_\xi + \sqrt{M_\xi}\sqrt{s/2} + s)^{1/2}, \tag{4.43}$$

with probability exceeding $1 - N_{\rm p} \exp(-s/2)$.

Similar to the discussion in § 4.4.2, the choice of the fidelity term in (4.42) allows for the decomposition of (4.42) into small-sized problems.

Lemma 4.18. *Problem* (4.42) *can be decoupled into* N_p *subproblems*

$$\hat{\boldsymbol{x}}_{l_\mathrm{p}} = \underset{\boldsymbol{u}\in\mathbb{C}^{N_\xi}}{\arg\min} \ \|\boldsymbol{\Psi}_\nu^\top \boldsymbol{u}\|_1 \ \text{s.t.} \ \|\boldsymbol{y}_{l_\mathrm{p}}^\mathrm{ci} - \boldsymbol{P}_{\Omega^\xi}\boldsymbol{F}^*\boldsymbol{u}_{l_\mathrm{p}}\| \le \varepsilon_\mathrm{ci}, \tag{4.44}$$

with $1 \le l_\mathrm{p} \le N_\mathrm{p}$.

Proof. It follows from the proof of Lemma 4.9, by change of variables $\boldsymbol{\Psi}_\mathrm{dhw} \leftarrow \boldsymbol{\Psi}_\nu$ and $\boldsymbol{D}^\xi \leftarrow \boldsymbol{I}_{M_\xi}\sqrt{M_\xi}$. □

From N_p sub-problems in (4.44), we can derive a recovery guarantee for the reconstruction of an HS volume $\boldsymbol{X}$ from (4.42), according to the specified sparsity basis.

Following the setup in § 2.4 and similar to the steps in § 3.3, we need to assign proper sparsity and sampling levels to the sparsity basis $\boldsymbol{\Psi}_\nu$ and sensing basis $\boldsymbol{F}$.

For $\boldsymbol{\Psi}_\nu = \boldsymbol{\Psi}_\mathrm{dhw}$, we leverage the convention of Adcock *et al.* in [9], *i.e.*, assigning the sparsity levels to the 1-D dyadic wavelet levels $\mathcal{T}^\mathrm{1d}$, defined in (3.5) with $N = N_\xi$, and assigning the sampling levels to the 1-D dyadic Fourier bands $\bar{\mathcal{T}}^\mathrm{df} := \{\bar{\mathcal{T}}_l^\mathrm{df}\}_{l=0}^r$, with $r = \log_2(N_\xi)$ defined as

$$\begin{aligned} &\bar{\mathcal{T}}_l^\mathrm{df} := \{-2^{l-1}+1,\cdots,2^{l-1}\}\backslash\{-2^{l-2}+1,\cdots,2^{l-2}\}, \ \text{for } l \ge 2,\\ &\bar{\mathcal{T}}_0^\mathrm{df} := \{0\}, \ \text{and } \bar{\mathcal{T}}_1^\mathrm{df} := \{1\}. \end{aligned} \tag{4.45}$$

Notice that the elements of the levels $\bar{\mathcal{T}}_l^\mathrm{df}$, for $l \ge 2$ are symmetric around the zeroth index, which corresponds to the zero frequency component. However, in this thesis we assume that the rows and the columns of a matrix are indexed with positive integers, meaning that the zero Fourier frequency is identified at the $N_\xi/2$-th row of the Discrete Fourier Transform (DFT) matrix $\boldsymbol{F}^*$. Therefore, it is more convenient to consider a shifted version of the dyadic Fourier bands in (4.45) by defining $\mathcal{T}^\mathrm{df} := \{\mathcal{T}_l^\mathrm{df}\}_{l=0}^r$, where

$$\mathcal{T}_l^\mathrm{df} := \{j + N_\xi/2 \ \text{s.t.} \ j \in \bar{\mathcal{T}}_l^\mathrm{df}\} \subset [\![N_\xi]\!], \ \forall l \in [\![r]\!]_0, \tag{4.46}$$

with $|\mathcal{T}_l^{\rm df}| = 2^{(l-1)_+}$. With these conventions, Adcock *et al.* [9] formulated the recovery guarantee of the Fourier-Haar system in the MDS context of Thm. 2.6, as follows.

Proposition 4.19 (Non-uniform recovery guarantee for the 1-D Fourier-Haar system, adapted from [9]). *Given $N_\xi = 2^r$ for some integer $r \in \mathbb{N}$, for $\boldsymbol{\Phi} = \boldsymbol{F}_{N_\xi} \in \mathbb{C}^{N_\xi \times N_\xi}$, $\boldsymbol{\Psi} = \boldsymbol{\Psi}_{\rm dhw} \in \mathbb{R}^{N_\xi \times N_\xi}$, $\mathcal{S} = \mathcal{T}^{\rm 1d}$, and $\mathcal{W} = \mathcal{T}^{\rm df}$, if we fix*

$$m_t \gtrsim \Big(\sum_{l=0}^{r} 2^{-|t-l|/2}\, k_l\Big)\, \log(K\epsilon^{-1})\, \log(N_\xi), \text{ for } t \in [\![r]\!]_0, \tag{4.47}$$

then (2.15) *and* (2.17) *in Thm. 2.6 are satisfied.*

We now consider $\boldsymbol{\Psi}_\nu = \boldsymbol{F}_{N_\xi}$. In this case, we assign the sparsity and sampling levels to the 1-D identical-cardinality levels $\bar{\mathcal{T}}^{\rm id} := \{\bar{\mathcal{T}}_l^{\rm id}\}_{l=1}^{\bar{r}}$, with $\bar{r} = 2^{q_0}$ for some integer $q_0 \geq 0$, defined as

$$\begin{aligned}
&\bar{\mathcal{T}}_l^{\rm id} := \{-\tfrac{lN_\xi}{2\bar{r}} + 1, \cdots, \tfrac{lN_\xi}{2\bar{r}}\} \backslash \{-\tfrac{(l-1)N_\xi}{2\bar{r}} + 1, \cdots, \tfrac{(l-1)N_\xi}{2\bar{r}}\}, \text{ for } l \geq 2,\\
&\bar{\mathcal{T}}_1^{\rm id} := \{-\tfrac{N_\xi}{2\bar{r}} + 1, \cdots, \tfrac{N_\xi}{2\bar{r}}\}.
\end{aligned} \tag{4.48}$$

Similar to the above, we pursue our analysis by considering a shifted version of the levels in (4.48), *i.e.*, we define

$$\mathcal{T}_l^{\rm id} := \{j + N_\xi/2 \text{ s.t. } j \in \bar{\mathcal{T}}_l^{\rm id}\} \subset [\![N_\xi]\!],\ \forall l \in [\![\bar{r}]\!], \tag{4.49}$$

with $|\mathcal{T}_l^{\rm id}| = N/\bar{r}$. With these definitions in hand, in the following proposition, we present the recovery guarantee of the Fourier-Fourier system for the MDS scheme.

Proposition 4.20 (Non-uniform recovery guarantee for the 1-D Fourier-Fourier system). *Given $N_\xi = 2^r$ and $\bar{r} = 2^{q_0} \leq N_\xi$ for some integers*

$r, q_0 \in \mathbb{N}$, for $\boldsymbol{\Phi} = \boldsymbol{\Psi} = \boldsymbol{F}_{N_\xi} \in \mathbb{C}^{N_\xi \times N_\xi}$ and $\mathcal{S} = \mathcal{W} = \mathcal{T}^{\mathrm{id}}$, if we fix

$$m_t \gtrsim \frac{N_\xi}{\bar{r}} \min\big(1, k_t \log(K\epsilon^{-1}) \log(N_\xi)\big), \text{ for } t \in [\![\bar{r}]\!], \tag{4.50}$$

then (2.15) *and* (2.17) *in Thm. 2.6 are satisfied.*

Proof. See § 4.9.3. □

This proposition enforces full sampling for the levels where $k_t > 0$, which might be undesirable from theoretical point of view. However, when the majority of local sparsity values k_t are zero, the Fourier-Fourier system will require very few number of measurements (see § 4.8.4).

According to Prop. 4.19 and Prop. 4.20, by selecting $M_\xi = \sum_t m_t$ OPD indices with m_t satisfying the sample-complexity bounds (4.47) and (4.50) for the settings $\{\boldsymbol{\Psi}_\nu, \mathcal{S}\} = \{\boldsymbol{\Psi}_{\mathrm{dhw}}, \mathcal{T}^{\mathrm{df}}\}$ and $\{\boldsymbol{\Psi}_\nu, \mathcal{S}\} = \{\boldsymbol{F}_{N_\xi}, \mathcal{T}^{\mathrm{id}}\}$, respectively, the reconstruction of a fixed column of $\boldsymbol{X}$, say $\boldsymbol{x}_{l_{\mathrm{p}}}$, from (4.44) satisfies

$$\|\boldsymbol{x}_{l_{\mathrm{p}}} - \hat{\boldsymbol{x}}_{l_{\mathrm{p}}}\| \le c_1\, \sigma_{\mathcal{S},\boldsymbol{k}}(\boldsymbol{\Psi}_\nu^* \boldsymbol{x}_{l_{\mathrm{p}}}) + c_2(1 + C\sqrt{K_\xi})\, \varepsilon_{\mathrm{ci}} \sqrt{q},$$

where $K_\xi = \sum_l k_l$, with probability exceeding $1 - \epsilon$ and with constants c_1, c_2, C and q specified in Thm. 2.6. Consequently, since $(a+b)^2 \le 2(a^2 + b^2) \le 2(a+b)^2$ for all $a, b > 0$, using a union bound over all the columns of $\boldsymbol{X}$, the reconstruction of the whole HS data from (4.42) satisfies

$$\begin{aligned}\|\boldsymbol{x} - \hat{\boldsymbol{x}}\| \le c_1 \sqrt{2}\, \Big(\sum_{l_{\mathrm{p}}=1}^{N_{\mathrm{p}}} (\sigma_{\mathcal{S},\boldsymbol{k}}(\boldsymbol{\Psi}_\nu^* \boldsymbol{x}_{l_{\mathrm{p}}}))^2 \Big)^{1/2} \\ + c_2 \sqrt{2N_{\mathrm{p}}}(1 + C\sqrt{K_\xi})\, \varepsilon_{\mathrm{ci}} \sqrt{q},\end{aligned} \tag{4.51}$$

with probability $1 - N_{\mathrm{p}}\epsilon$.

We are now ready to summarize all the analysis in this section and to provide the main result for the MDS-based CI-FTI in the following theorem.

Theorem 4.21 (MDS for Coded illumination-FTI)**.** *Given* $s > 0, 0 < \epsilon \leq \exp(-1)$*, and integers* $N_{\rm p}$ *and* $N_\xi = 2^r$ *for some* $r \in \mathbb{N}$*, if we fix*
(i) for $t \in [\![r]\!]_0,\ \boldsymbol{\Psi}_\nu = \boldsymbol{\Psi}_{\rm dhw} \in \mathbb{R}^{N_\xi \times N_\xi}, \mathcal{W} = \mathcal{T}^{\rm df}, \mathcal{S} = \mathcal{T}^{\rm 1d}$;

$$m_t \gtrsim \big(\sum_{l=0}^{r} 2^{-|t-l|/2}\, k_l\big)\, \log(K_\xi \epsilon^{-1})\, \log(N_\xi),$$

(ii) for $t \in [\![\bar{r}]\!],\ \boldsymbol{\Psi}_\nu = \boldsymbol{F}_\xi \in \mathbb{C}^{N_\xi \times N_\xi}, \mathcal{W} = \mathcal{S} = \mathcal{T}^{\rm id}$;

$$m_t \gtrsim \frac{N_\xi}{\bar{r}} \min\big(1, k_t \log(K_\xi \epsilon^{-1})\, \log(N_\xi)\big),$$

where $K_\xi = \sum_l k_l$*, then, given the corresponding noisy CI-FTI measurements* $\boldsymbol{y}^{\rm ci}$ *in* (4.16) *with* Ω^ξ *a* $(\mathcal{W}, \boldsymbol{m})$*-MDS subsampling set, the HS volume* $\boldsymbol{x}$ *can be approximated by solving* (4.42) *with the bound* $\varepsilon_{\rm ci}(s)$ *in* (4.43)*, up to an error*

$$\begin{aligned}\|\boldsymbol{x} - \hat{\boldsymbol{x}}\| \leq c_1\sqrt{2}\,\Big(\sum_{l_{\rm p}=1}^{N_{\rm p}} (\sigma_{\mathcal{S},\boldsymbol{k}}(\boldsymbol{\Psi}_\nu^* \boldsymbol{x}_{l_{\rm p}}))^2\Big)^{1/2} \\ + c_2\sqrt{2N_{\rm p}}(1 + C\sqrt{K_\xi})\, \varepsilon_{\rm ci}\sqrt{q},\end{aligned}$$

and with probability exceeding $1 - N_{\rm p}\epsilon - N_{\rm p} e^{-s/2}$ *with constants* c_1, c_2, C, L *and* q *specified in Thm. 2.6.*

Proof. A combination of Thm. 2.6, Prop. 4.19, and Prop. 4.20 with (4.43) completes the proof. □

This theorem states that, *e.g.*, for $\boldsymbol{\Psi}_\nu = \boldsymbol{\Psi}_{\rm dhw}$, the local number of measurements m_t scales as a linear combination of the local sparsity values k_l; and interestingly, the dependence of m_t on k_l, for $t \neq l$, is exponentially vanishing in $|t - l|$. Moreover, when $k_l \to 0$ as l increases, the dependence of m_t on k_l will be further diminished. Fortunately, this is true for the spectral signatures of the fluorochromes; see Fig. 4.13. Hence, we expect that an application of the MDS in CI-FTI would yield higher signal recovery quality, compared to the VDS-based CI-FTI supported in Thm. 4.12. Let us now consider $\boldsymbol{\Psi}_\nu = \boldsymbol{F}$. According to Thm. 4.21, when $k_l > 0$ for all $l \in [\![\bar{r}]\!]$, full sampling of the OPD domain is sufficient for an HS recovery. Indeed, this statement is not informative. However, as we will see in § 4.8.4 and Fig. 4.13, the local sparsity values of the fluorochromes spectra in the Fourier domain are mostly zero, *i.e.*, $k_l = 0$ for $l \geq l_0$ with $l_0 \ll \bar{r}$; in our experiments in § 4.8.4, $l_0 \leq 6$, compared to $\bar{r} = 64$.

4.8 Numerical Results

We conduct four simulations to verify the performance of the proposed compressive FTI methods. In the first scenario, we examine the efficiency of the proposed VDS strategies for CI-FTI and SI-FTI in (4.20) and (4.29) for any arbitrary sparse HS data by tracing the phase transition curves of successful recovery in a noiseless setting. The second scenario consists in following up the performance of the proposed FTI frameworks with constrained- and unconstrained-exposure budget on simulated biological HS volumes. In the third setup, constrained-exposure CI-FTI is simulated from the real FTI measurements. In the fourth simulation, we illustrate the improvement of the HS data recovery in CI-FTI, when the illumination coding is designed based on the MDS scheme proposed in Thm. 4.21. For all the experiments we report the SRE defined in (3.18). The HS volume reconstructions (4.17) and (4.27) are numerically performed with the SPGL1 solver [117, 118].

Concerning the value $\varepsilon_{\rm ci}$ and $\varepsilon_{\rm si}$ of the noise power in the optimization problems (4.17) and (4.27), we observed numerically that the estimations in (4.21) and (4.31) are not tight enough at small number of measurements. While tightening those bounds constitutes an interesting future work, we have rather numerically estimated $\varepsilon_{\rm ci}$ and $\varepsilon_{\rm si}$ in our simulations. For this, we have computed the empirical 95^{th} percentile curve of the weighted noise power over 100 Monte-Carlo realizations of both a Gaussian random noise (with unit variance) and the index set Ω

(for the selected sensing scenario, *i.e.*, CI- or SI-FTI) over the considered interval of measurement number. This experimental curve has then be multiplied by the noise standard deviation, known (for synthetic examples) or estimated (for experimental data).

4.8.1 Impact of Different VDS Sstrategies in Compressive FTI

We here challenge VDS densities defined in Thm. 4.12 and Thm. 4.16 by comparing their performances (in terms of reconstruction error) with those reached by other density functions with different decaying power laws in the OPD domain. We also show that the prescribed schemes achieve near-optimal results with respect to a pmf set to the value of the local coherence $\boldsymbol{\mu}^{\rm loc}$ between the sparsity and the sensing bases.

In details, following § 4.4.1 and § 4.5.1, the compressive FTI scenarios are associated with different subsampling of measurement indices, as recorded in the measurement index set Ω. In particular, the (possibly non-unique) random elements of Ω are iid according to the following pmfs generalizing (4.20) and (4.29) to variable power laws:

$$\begin{array}{|l|l|}\hline (\mathrm{CI-FTI}) & p^{\alpha}(l_\xi) := C_{N_\xi,\alpha}\min\big\{1, \frac{1}{|l_\xi - N_\xi/2|^\alpha}\big\}, \\ & p^{\rm opt}(l_\xi) := \|\boldsymbol{\mu}^{\rm loc}(\boldsymbol{F}^*\boldsymbol{\Psi}_{\rm dhw})\|^{-2}\mu^{\rm loc}_{l_\xi}(\boldsymbol{F}^*\boldsymbol{\Psi}_{\rm dhw})^2, \\ \hline (\mathrm{SI-FTI}) & p^{\alpha}(l_{\rm hs}) := \frac{C_{N_\xi,\alpha}}{N_{\rm p}}\min\big\{1, \frac{1}{|l_\xi - N_\xi/2|^\alpha}\big\}, \\ & p^{\rm opt}(l_{\rm hs}) := \|\boldsymbol{\mu}^{\rm loc}(\boldsymbol{\Phi}^*_{\rm si}\boldsymbol{\Psi}_{\rm si})\|^{-2}\mu^{\rm loc}_{l_{\rm hs}}(\boldsymbol{\Phi}^*_{\rm si}\boldsymbol{\Psi}_{\rm si})^2, \\ & \text{with } l_{\rm hs} \xrightleftharpoons{N_\xi, N_{\rm p}} (l_\xi, l_{\rm p}), \\ \hline \end{array} \tag{4.52}$$

where $C_{N_\xi,\alpha}$ ensures that the probabilities sum to one, and the specific pmf parameterizations are defined in § 4.4.1 and § 4.5.1. In (4.52), the parameter $\alpha \in \{0, 1, 1.5, 2, 8\}$ controls the decaying power of the VDS strategy; for $\alpha = 0$ it reduces to the UDS strategy and for $\alpha = 1$ it matches the pmfs of (4.20) and (4.29). We also consider the optimal VDS strategies, *i.e.*, the two pmfs $p^{\rm opt}$ in (4.52), where each pmf — for CI- and SI-FTI — is set to the exact values of local coherence between the sensing

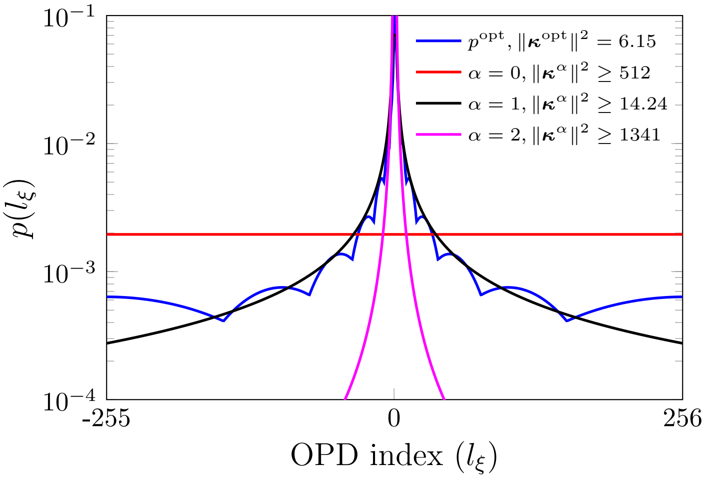

Figure 4.4: Representation of different CI-FTI sampling strategies for $N_\xi = 512$ associated with pmfs $p^\alpha(l_\xi)$ and $p^{\mathrm{opt}}(l_\xi)$ in (4.52). The lower value of $\|\boldsymbol{\kappa}^\alpha\|^2$ is an indicator of sampling optimality. Smaller values amount to a tighter sample-complexity bound in (2.7). The same trend is expected for the phase transition curves as well, which is confirmed in Fig. 4.5.

and the sparsity bases. While this setting is numerically computable, it is, however, hardly integrable to our estimation of the noise power, and thus to our analysis of the reconstruction error of the HS volume.

Note that in the case of CI-FTI, from Prop. 2.3, the pmf p^α in (4.52) is admissible — *i.e.*, it allows for HS volume reconstruction in CI-FTI — only if it is proportional to a vector $\boldsymbol{\kappa}^\alpha \in \mathbb{R}_+^{N_\xi}$ that bounds the local coherence $\mu_{l_\xi}^{\mathrm{loc}}(\boldsymbol{F}^*\boldsymbol{\Psi}_{\mathrm{dhw}})$. Thus, we must have $\|\boldsymbol{\kappa}^\alpha\|^2 p^\alpha(l_\xi) = (\kappa_{l_\xi}^\alpha)^2 \geq \mu_l^{\mathrm{loc}}(\boldsymbol{F}^*\boldsymbol{\Psi}_{\mathrm{dhw}})^2$ for all $l_\xi \in [\![N_\xi]\!]$. This involves in particular $\|\boldsymbol{\kappa}^\alpha\|^2 p^\alpha(N_\xi/2) = \|\boldsymbol{\kappa}^\alpha\|^2 C_{N_\xi,\alpha} \geq 1$, since $\mu_{N_\xi/2}^{\mathrm{loc}}(\boldsymbol{F}^*\boldsymbol{\Psi}_{\mathrm{dhw}}) = 1$ (see § 4.9.2), *i.e.*, we necessarily have

$$\|\boldsymbol{\kappa}^\alpha\|^2 \geq 1/C_{N_\xi,\alpha} = \sum_{l_\xi} \min\{1, \frac{1}{|l_\xi - N_\xi/2|^\alpha}\}.$$

In the case of p^{opt}, we can directly set $\kappa_l^{\mathrm{opt}} = \mu_{l_\xi}^{\mathrm{loc}}(\boldsymbol{F}^*\boldsymbol{\Psi}_{\mathrm{dhw}})$ by construction.

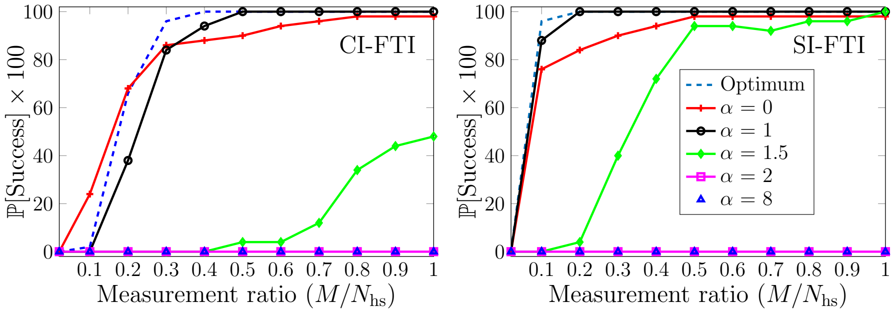

Figure 4.5: Phase transition comparison of proposed VDS ($\alpha = 1$) against UDS ($\alpha = 0$) and other VDS strategies in the frameworks of CI-FTI and SI-FTI. For CI-FTI, $M = M_\xi N_{\mathrm{p}}$.

We provide an illustration of the corresponding pmfs in Fig. 4.4 for CI-FTI. It is clear that the pmf for $\alpha = 1$ is close to the optimal pmf curve p^{opt}. This is in agreement with the bound $\|\boldsymbol{\kappa}^1\|^2 \geq 14.24$, *i.e.*, a value closer to $\|\boldsymbol{\kappa}^{\mathrm{opt}}\|^2 \simeq 6.15$ than the lower bounds of $\|\boldsymbol{\kappa}^\alpha\|^2$ for the other values of α (see Fig. 4.4). The CI- and SI-FTI measurements are simulated by restricting the Nyquist measurements to a subset Ω specified by the pmfs in (4.52) and the procedures defined in § 4.4.1 and § 4.5.1. At each trial of our experiments, we generate randomly both Ω and a synthetic HS volume from $\boldsymbol{X} = \boldsymbol{\Psi S}$, where $\boldsymbol{\Psi} = \boldsymbol{\Psi}_{\mathrm{idhw}} \otimes \boldsymbol{\Psi}_{\mathrm{dhw}}$ for SI-FTI, and $\boldsymbol{\Psi} = \boldsymbol{I}_{N_{\mathrm{p}}} \otimes \boldsymbol{\Psi}_{\mathrm{dhw}}$ for CI-FTI. The matrix $\boldsymbol{S} \in \mathbb{R}^{N_\xi \times N_{\mathrm{p}}}$ is sparse, its dimensions are $N_\xi = 512$ and $N_{\mathrm{p}} = 8^2$ (*i.e.*, $N_{\mathrm{hs}} = 2^{15}$), and its row and column sparsity levels are $K_\xi = 4$ and $K_{\mathrm{p}} = 4$, respectively. The indices of the K_ξ rows and K_{p} columns are chosen uniformly at random and the non-zero random coefficients follow a normal distribution. We finally simulate noiseless Nyquist FTI measurements of $\boldsymbol{x} = \mathrm{vec}(\boldsymbol{X})$ from $\boldsymbol{y}^{\mathrm{fti}} = \boldsymbol{\Phi}^*_{\mathrm{fti}}\boldsymbol{x}$.

For a fixed measurement ratio M/N_{hs} this procedure is repeated for 50 independent trials. We trace the phase transition curves in Fig. 4.5 which shows the probability of successful recovery against the number of measurements by solving (4.17) and (4.27) with a zero noise level (*i.e.*, $\varepsilon^{\mathrm{ci}} = \varepsilon^{\mathrm{si}} = 0$). We count a recovery as successful if $\|\hat{\boldsymbol{x}} - \boldsymbol{x}\|^2 \leq 10^{-10}\|\boldsymbol{x}\|^2$.

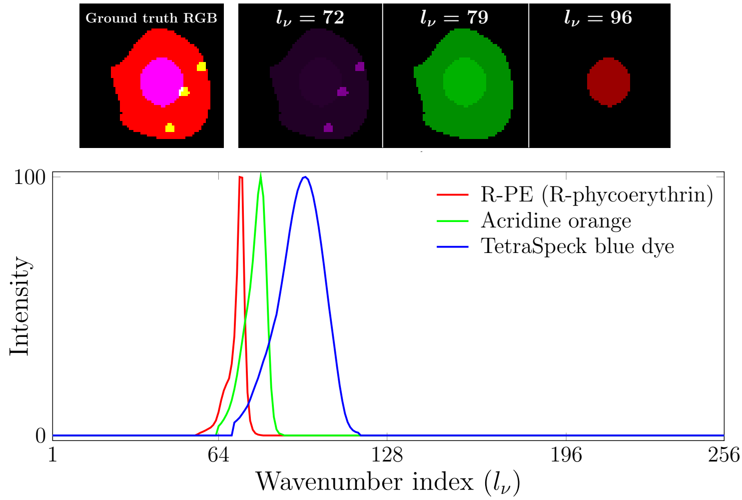

Figure 4.6: A synthetic biological RGB image (top-left); three spectral bands of the generated ground truth HS volume (top-right); the known spectral signatures of three fluorochromes (bottom).

The plot confirms that the value $\alpha = 1$ (*i.e.*, the proposed VDS strategy) is very close to the optimal sampling strategy p^{opt} and can reach 100 % chance of successful recovery from $M/N_{\mathrm{hs}} > 0.2$ and $M/N_{\mathrm{hs}} > 0.5$ for SI-FTI and CI-FTI, respectively; while for the UDS strategy ($\alpha = 0$) we observe that this cannot be achieved even when $M/N_{\mathrm{hs}} = 1$. Hereafter, the rest of our experiments are restricted to the case $\alpha = 1$.

4.8.2 Reconstruction Performances on a Synthetic Biological HS Volume

We now test the performance of the proposed compressive FTI frameworks in a more realistic context including measurement noise. We consider sensing scenarios with both constrained and unconstrained light exposure.

As illustrated in Fig. 4.6, we simulate a biological HS volume of size $(N_\xi, N_{\mathrm{p}}) = (512, 64^2)$ by mixing the coefficients of three spectral bands of a synthetic biological RGB image (selected from the benchmark images [101]) with the known spectra of three common fluorochromes.

Remark 4.22. *The RGB image in Fig. 4.6 is a simulated cell including three different components, such as nuclei, cytoplasm, and small intercellular elements. In the construction of the HS volume here we simply assumed that the nuclei, cytoplasm, and intercellular objects are marked by TetraSpeck blue dye, Acridine orange, and R-phycoerythrin fluorochromes, respectively. Although the spatial configuration of the constructed HS volume is synthetic, its spectral content contains real fluorochrome spectra. However, note that our results in this chapter and Chapter 5 are not limited only to this specific HS volume. In Thms. 4.12, 4.16, and 5.10, we provide recovery guarantees for a class of HS volumes whose spectral (and spatial) content have sparse or compressible representation in 1-D (respectively, 2-D) Haar wavelet domain. Therefore, replacing the synthetic HS volume generated here with a real volume will still be supported by our results, provided that the real HS volume has a compressible representation (in the sense of what is described above).*

The Nyquist measurements are formed as $\boldsymbol{y}^{\rm fti} = \boldsymbol{\Phi}_{\rm fti}^{*}\boldsymbol{x} + \boldsymbol{n}^{\rm fti}$ where $n_{l,l'}^{\rm fti} \sim_{\rm iid} \mathcal{N}(0, \sigma_{\rm fti})$ and $\sigma_{\rm fti}$ is fixed such that SNR $= 20$ dB. In the unconstrained-exposure context, CI-FTI and SI-FTI observations are formed according to (4.16) and (4.26), where the sets Ω^{ξ} and Ω are randomly generated from the pmfs (4.20) and (4.29), respectively, and the variance of each component of the associated additive Gaussian measurement noises is also set to $\sigma_{\rm fti}^2$. In the context of constrained-exposure simulations, the values of ground truth HS volume are multiplied by N_{ξ}/M_{ξ} and $N_{\xi}/\bar{M}$ for CI-FTI and SI-FTI, respectively, according to the linear intensity scaling assumption explained in (4.35) and (4.40); see § 4.6.

This sensing context is repeated over 10 random realizations of the noise, and the sets Ω^{ξ} or Ω. Fig. 4.7 depicts the SRE in dB as a function of the number of measurements. The HS volumes are reconstructed by solving two types of problems: *(i)* the CS-based ℓ_1 minimization

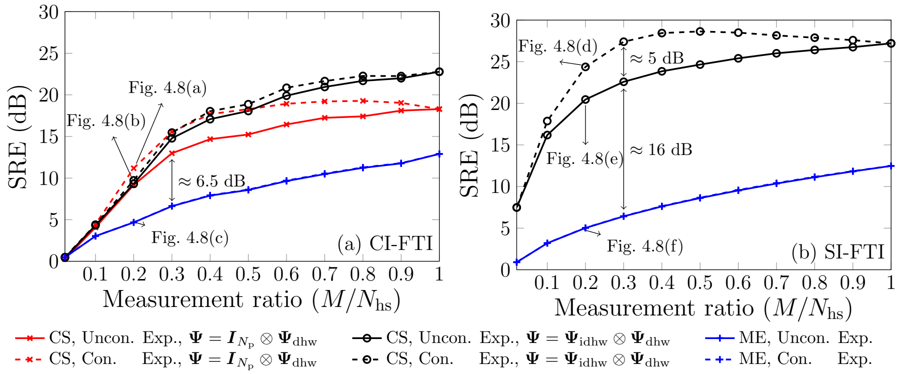

Figure 4.7: The reconstruction performance of CI- and SI-FTI systems solved by ℓ_1 minimization (CS) and ME problem with the constrained and unconstrained light exposure budget. Note that for a fixed measurement ratio, the amount of light exposure in the unconstrained exposure scenario is less than the one in constrained exposure scenario, by a factor M/N_{hs}.

problem (4.17) or (4.27); and *(ii)* the ME problem [22], *i.e.*, a standard reconstruction method that amounts to applying the pseudo-inverse of the sensing operator on the noisy measurements.

In both cases, ME reconstructions (the blue curves) do not reach CS reconstruction qualities, even though they benefit of the VDS strategy. The poor performances of the ME solutions advocate the necessity of promoting sparsity prior in the reconstruction problem. The increased SRE value of SI-FTI over CI-FTI is induced by both the 3-D wavelet sparsity model and the greater diversity of the compressive sensing matrix in SI-FTI where spatial information of the HS volume is partially captured at each OPD index in $[\![N_\xi]\!]$; while in CI-FTI, this spatial information is fully observed only on the M_ξ selected OPD indices. We recall that since there are repeated indices in subsampled sets Ω^ξ and Ω, even for $M/N_{\text{hs}} = 1$, we cannot sample all the distinct elements. As a consequence, ME reconstructions does not reach Nyquist quality at $M/N_{\text{hs}} = 1$.

We now question whether the superior performance of SI-FTI is the consequence of the sensing diversity or of the selected 3-D wavelet

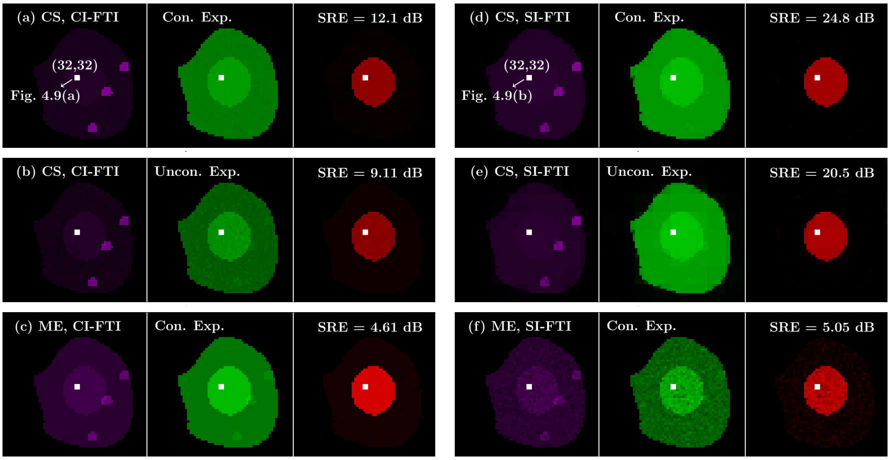

Figure 4.8: An example of the reconstructed HS volumes from 20 % of the total measurements (or light exposure). The spectral content at the spatial location indicated by a white square is shown in Fig. 4.9.

sparsity model. For this, deviating from what is guaranteed by CS literature (see § 4.4.2), we test the recovery of HS data from CI-FTI measurements using the recovery scheme (4.27) regularized by the same prior as in SI-FTI, *i.e.*, with the 3-D wavelet sparsity basis $\boldsymbol{\Psi}_{\rm si} = \boldsymbol{\Psi}_{\rm idhw} \otimes \boldsymbol{\Psi}_{\rm dhw}$. We thus solve this optimization with the changes $\boldsymbol{D} \leftarrow \boldsymbol{I}_{N_{\rm p}} \otimes \sqrt{M_\xi}\boldsymbol{D}^\xi$, $M \leftarrow M_\xi N_{\rm p}$, and $\boldsymbol{y}^{\rm si} \leftarrow \boldsymbol{y}^{\rm ci}$, and assuming that the level of the noise is unaffected between the CI-FTI and the SI-FTI sensing schemes. The results are displayed in Fig. 4.7-left (the black curves). We observe that this 3-D wavelet sparsity model increases the SRE, up to 4.5 dB, when $M/N_{\rm hs}$ is close to 1. This improvement, however, does not reach the SRE of SI-FTI displayed in Fig. 4.7-right. We thus conclude that the SI-FTI scheme benefit more from the diversity of its measurements than from its regularization compared to CI-FTI. In practice, HS data in CI-FTI should be reconstructed with this 3-D wavelet sparsity basis to reach the highest SRE when $M/N_{\rm hs}$ is large. Keeping this in mind, the next CI-FTI experiments are, however, run with $\boldsymbol{\Psi}_{\rm ci} = \boldsymbol{I}_{N_{\rm p}} \otimes \boldsymbol{\Psi}_{\rm dhw}$ for simplicity and consistency with our analysis in § 4.4.2 and § 4.6.

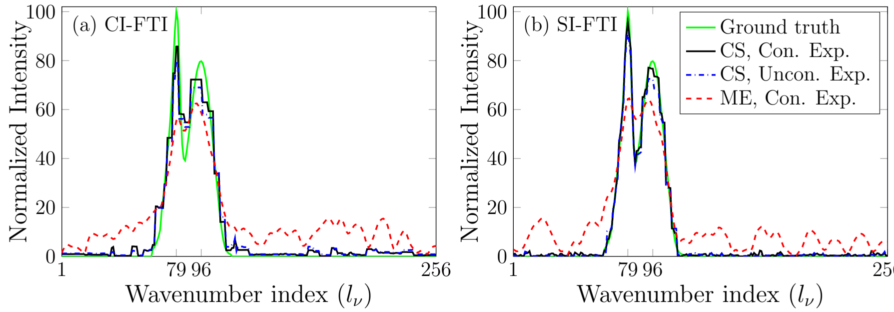

Figure 4.9: The spectrum of the reconstructed HS volumes in Fig. 4.8 at spatial location $(l_1, l_2) = (32, 32)$.

For constrained-exposure scenarios, we observe that — as opposed to the common belief — subsampling 40 percent of the FTI measurements yields better HS recovery quality than the one obtained from maximal number of FTI measurements[11]. However, this effect vanishes for large values of M and the SRE decreases to reach the quality performances at the maximal number of measurements, which parallels with the behavior of the second error term in (4.36) and (4.41). This phenomenon is due to the fact that the best K-term approximation error in the right-hand side of (4.36) or (4.41) is dominated by the noise term, which can be reduced by taking less number of measurements. Besides, ME reconstruction cannot leverage constrained-exposure budget, since the measurement noise is the same as the one in unconstrained-exposure scenario.

In Fig. 4.8, we illustrate three spatial maps associated with wavenumber indices $l_\nu \in \{72, 79, 96\}$ of the reconstructed HS volumes. This figure is one instance of the Monte-Carlo trials in the experiment of Fig. 4.7, for $M/N_{\rm hs} = 0.2$. It can be seen that, for both CI- and SI-FTI systems, the HS volume recovered with CS-based formulation preserves the spatial

[11]We note that in the case of CI-FTI regularized with 3-D wavelet basis, there is still a 1 dB gain (around $M/N_{\rm hs} = 0.5$) to use the constrained-exposure scenario over the unconstrained one. However, a deeper analysis of this effect for this setting goes beyond the scope of this thesis.

configuration of the specimen, although the SRE for CI-FTI may not be satisfactory. However, as mentioned in § 4.1, the main motivation for studying FTI is to acquire HS volumes with high spectral resolution, *i.e.*, limited by the life-time of the fluorochromes. Note that the spatial resolution is determined by physical characteristics of the 2-D imaging sensor in Fig. 4.2. In order to illustrate the ability of the proposed approaches in preserving the spectral information, Fig. 4.9 also compares the spectral content of the reconstructed HS volumes at the center spatial location indicated by a white square. The artifacts in the solution of ME problem are obvious and they could be sufficiently prominent for causing false detection in the decomposition of a biological HS volume into its spectral constituents. Despite minor disturbances, the CS-based solution in CI-FTI successfully follows the ground truth spectrum. However, the solution of SI-FTI almost perfectly matches the ground truth signal.

4.8.3 Simulation of the Constrained-exposure CI-FTI from Actual Experimental data

To prove the concept of CI-FTI, we have simulated a CI-FTI setup in a constrained-exposure context by subsampling the data recorded by an actual (Nyquist) FTI acquisition. We have scanned a thin layer of a biologic cell (*i.e.*, Convallaria, lily of the valley, cross section of rhizome with concentric vascular bundles) stained on a glass slide, using a confocal microscope (with a 40× objective) equipped with the FTI and a LUXEON light source [2], *i.e.*, the *HYPE* device used in [65]. By changing the current level of the light source from $I_{\rm low} = 100$ mA to $I_{\rm high} = 700$ mA (with $\Delta_I = 25$ mA) we acquired 25 sets of Nyquist FTI measurements of size $(N_\xi, \bar{N}_{\rm p}, \bar{N}_{\rm p}) = (1024, 128, 128)$, with $\xi_{\max} = 2.537$ μm, *i.e.*, each set corresponds to one light intensity level[12]. This exhaustive acquisi-

[12]The light source specifications [2] specifies that the relative luminous flux is proportional to the forward current.

tion enables us to simulate constrained-exposure compressive FTI (see below).

Concerning the level of $\sigma_{\rm fti}$ of the measurement noise, we have estimated it using an RM estimator [33] applied on each of the 25 HS data cubes. We observed that, over the range $[I_{\rm low}, I_{\rm high}]$ with $I_{\rm high}/I_{\rm low} = 7$, $\sigma_{\rm fti} \in [0.1408, 0.2272]$, *i.e.*, a linear regression provides $\sigma_{\rm fti}(I) \approx a(I/100) + b$ with $a = 1.44\ 10^{-2}$ and $b = 1.26\ 10^{-1}$. This limited variation of σ when I varies confirms that the dependence of the Nyquist noise on the light intensity can be neglected (see § 4.2.3 and § 4.6).

The simulated CI-FTI sensing model of this section contains one crucial difference compared to the subsampling procedure studied in § 4.4.1 and tested in the two previous sections. Since the subsampling set Ω^ξ contains possible repetition of the same OPD indices (*i.e.*, according to the pmf (4.20) that is used to draw M_ξ iid OPD indices with repetition), a restriction of the (noisy) Nyquist FTI observations above to Ω^ξ at each pixel location copies the same realizations of the noise on repeated indices. This differs from the model (4.16) where the noise samples vary on possible multiple instances of the same OPD index.

Consequently, rather than artificially copying the same observations for each repeated index in Ω^ξ, we have adopted the following more realistic experimental sensing scenario. For a given M_ξ, the subsampling set Ω^ξ is randomly generated according to the pmf (4.20). We then construct a set $\bar\Omega^\xi := \{\bar\omega_1, \cdots, \bar\omega_{M_{\rm eff}}\}$ with cardinality $M_{\rm eff} < M_\xi$ which is a replica of Ω^ξ with only unique indices[13]. Finally, CI-FTI measurements are built by restricting the Nyquist FTI measurements $\boldsymbol{y}^{\rm fti}$ to the subset $\bar\Omega^\xi$.

By following this procedure, the data fidelity term in (4.18) must be adapted; the weighting matrix $\boldsymbol{D}^\xi$ must stay constant across its diagonal elements. Indeed, we can easily show that each index l presents in

[13]For $l \in [\![N_\xi]\!]$, define the Bernoulli random variable X_l being equal to 1 if the $l^{\rm th}$ OPD element is picked at least once over M_ξ trials, and zero otherwise, then $\mathbb{E}X_l = \mathbb{P}(X_l = 1) = 1 - (1 - p(l))^{M_\xi}$ and $\mathbb{E}M_{\rm eff} = \sum_l 1 - (1 - p(l))^{M_\xi}$, since $M_{\rm eff} = \sum_l X_l$.

Ω^ξ is repeated according to a binomial r.v. γ_l of M_ξ trials and success probability $p(l)$, with p the pmf defined in (4.20). Therefore, $\mathbb{E}\gamma_l = M_\xi p(l)$, which is exactly the inverse of each entry $\frac{1}{M_\xi} d_{kk}^2 = \frac{1}{M_\xi p(\omega_k)}$ of $\frac{1}{M_\xi}(\boldsymbol{D}^\xi)^2$ for which $\omega_k = l$. In other words, in expectation, *the matrix $\boldsymbol{D}^\xi$ accounts for a normalization of the multiplicity of each index present in Ω^ξ.* While a careful mathematical analysis of this effect is postponed to a future work, we conclude that, for each optimization vector $\boldsymbol{u}_{l_\mathrm{p}}$ and each $l_\mathrm{p} \in [\![N_\mathrm{p}]\!]$, $\|\boldsymbol{P}_{\bar{\Omega}^\xi}(\boldsymbol{y}_{l_\mathrm{p}}^{\mathrm{fti}} - \boldsymbol{F}^*\boldsymbol{u}_{l_\mathrm{p}})\|^2 \approx \frac{1}{M_\xi}\|\boldsymbol{D}^\xi \boldsymbol{P}_{\Omega^\xi}(\boldsymbol{y}_{l_\mathrm{p}}^{\mathrm{fti}} - \boldsymbol{F}^*\boldsymbol{u}_{l_\mathrm{p}})\|^2 \leq \varepsilon_{\mathrm{ci}}^2$ is an appropriate fidelity term in (4.18) in the replacement of Ω^ξ by $\bar{\Omega}^\xi$.

Back to our simulation setup, from the recorded Nyquist FTI measurements, a CI-FTI system with constrained-exposure budget is simulated as follows. Given a reference current level of the light source, *e.g.*, $I_\mathrm{ref} = 100$ mA as for the black curve in Fig. 4.10, without loss of generality, we assume[14] $I_\mathrm{opd} = I_\mathrm{ref}/\tau_\xi$ and thus a constrained-exposure budget is fixed as $I_\mathrm{tot} = I_\mathrm{ref} N_\xi \leq I_\mathrm{max}$. Therefore, for every other current levels, $I(\mathrm{mA})$, we are allowed to set $M_\xi(I, I_\mathrm{ref}) = \frac{I_\mathrm{ref}}{I} N_\xi$, *e.g.*, $M_\xi(700, 100)/N_\xi = 14.28\,\%$. Besides, in view of the above-mentioned sampling approach, considering $M_\mathrm{eff} \leq M$ unique subsampled indices does not violate the requirements of the constrained-exposure budget scenario. Hereafter, we report effective measurement ratio (M_eff/N_ξ) as it will be the actual exposure reduction ratio.

Since the Nyquist FTI measurements recorded at $I(\mathrm{mA}) = 700$ mA has the highest MNR (see (4.34)), the reconstructed HS volume[15] from this data is taken as the ground truth volume, *e.g.*, for the purpose of computing SRE. For the sake of obtaining fair SRE values, all the spectra are normalized with respect to their ℓ_2-norm, meaning that in SRE formula we replace $\boldsymbol{x} = [\boldsymbol{x}_1^\top, \cdots, \boldsymbol{x}_{N_\mathrm{p}}^\top]^\top$ with $\boldsymbol{x}^\circ := [\boldsymbol{x}_1^\top/\|\boldsymbol{x}_1\|, \cdots, \boldsymbol{x}_{N_\mathrm{p}}^\top/\|\boldsymbol{x}_{N_\mathrm{p}}\|]^\top$ and replace $\hat{\boldsymbol{x}} = [\hat{\boldsymbol{x}}_1^\top, \cdots, \hat{\boldsymbol{x}}_{N_\mathrm{p}}^\top]^\top$ with $\hat{\boldsymbol{x}}^\circ := [\hat{\boldsymbol{x}}_1^\top/\|\hat{\boldsymbol{x}}_1\|, \cdots, \hat{\boldsymbol{x}}_{N_\mathrm{p}}^\top/\|\hat{\boldsymbol{x}}_{N_\mathrm{p}}\|]^\top$. This allows for the computation of the SRE independently of the light intensity — the intensity of the reconstructed

[14] Remind that the flux of the light source is proportional to its current level.
[15] To do so, in (4.18), we set $\Omega^\xi = [\![N_\xi]\!]$, $\boldsymbol{D}^\xi = \sqrt{N_\xi}\boldsymbol{I}_{N_\xi}$, and $M_\xi = N_\xi$.

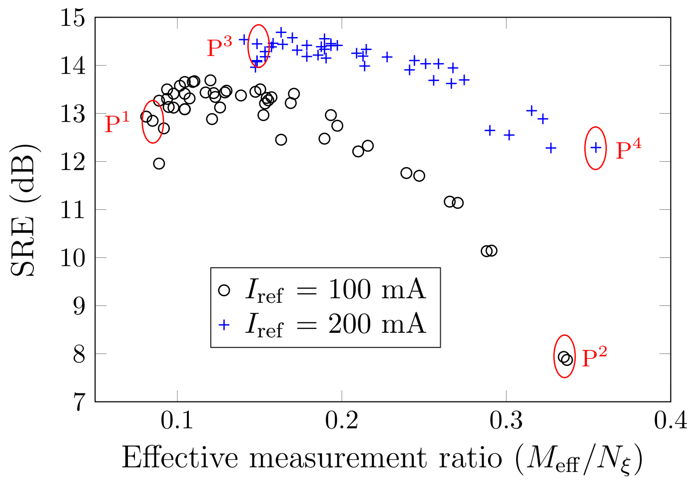

Figure 4.10: The reconstruction quality of CI-FTI system with constrained-exposure budget. For each curve, the (relative) light exposure budget is constrained to $I_{\text{ref}}N_\xi$; and thus $M_\xi(I, I_{\text{ref}})/N_\xi = I_{\text{ref}}/I$ for $I \in \{I_{\text{ref}}, I_{\text{ref}} + 25, \cdots, 700\}$ mA. Subsampling the OPD axis at higher light intensity results always in superior reconstruction.

spectrum being proportional to the light intensity — and it reduces the effect of various data corruptions, *e.g.*, saturated pixels, light diffraction due to the biological specimen, or motion artifact during the experiment.

The SRE values of CI-FTI system for two constrained-exposure budgets, *i.e.*, $I_{\text{tot}} \in \{100N_\xi, 200N_\xi\}$ mA, and for two random generations of Ω^ξ is illustrated in Fig. 4.10. These results stress that for a fixed light exposure budget, an application of the proposed CI-FTI method always results in a superior quality of the HS volume reconstruction. For instance, for $I_{\text{tot}} = 100N_\xi$ mA exposure budget, subsampling 8 % of the OPD axis (point P^1), approximately yields a 5 dB gain in the SRE, in comparison to subsampling 34 % of the OPD axis (point P^2). The reconstructed HS volumes corresponding to four circled points in Fig. 4.10 are shown in Fig. 4.11 (for the spectra observed at the center spatial location) and Fig. 4.12 (for the spatial maps at wavenumber index $l_\nu = 70$, *i.e.*, 594 nm wavelength, at which the spectra maximum occurs). In Fig. 4.11, the noise amplitudes in the spectra recovered from P^2 (34 % of FTI mea-

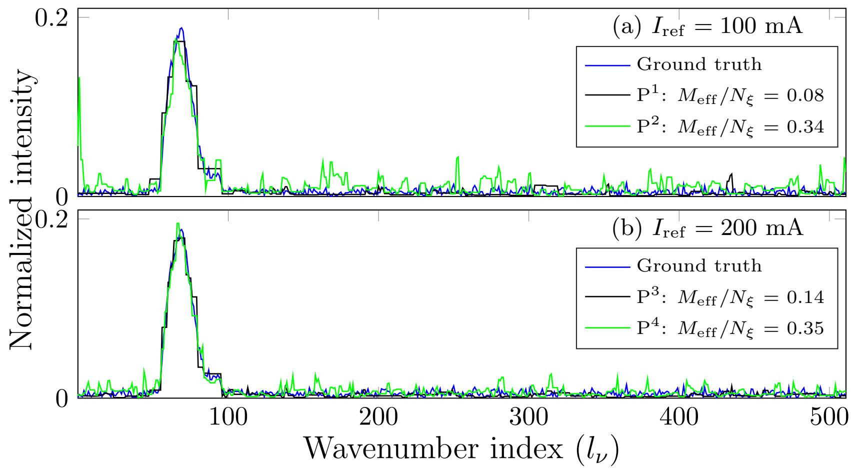

Figure 4.11: The spectra of the reconstructed HS volumes, associated with four marked instances in Fig. 4.10, at the center spatial location. Performing the proposed VDS strategy, even at the maximum compression ratio, significantly improves the quality of the reconstructed spectrum, especially in terms of denoising.

surements) and P^4 (35%) are significantly larger than in the spectra of P^1 (8 %) and P^3 (14%), respectively. In Fig. 4.12, as expected from the low light condition (*e.g.*, at 100 mA and 200 mA), the spatial map quality is significantly degraded due to noise (Fig. 4.12-c and Fig. 4.12-e). Lower subsampling rates allows for improved spatial map qualities (Fig. 4.12-b and Fig. 4.12-d) associated with an increased light intensity of 700 mA where the MNR peaks. Note that the parallel frontiers of the biological cells in Fig. 4.12 are an effect of the 3-D specimen transparency (also observed in the panchromatic microscope).

We conclude this section by mentioning that an SI-FTI framework could have also been simulated from the recorded data, *i.e.*, by randomly subsampling spatiotemporally the volume of recorded (Nyquist) interferograms. However, such a simulation would be too ideal with respect to an actual SI-FTI implementation integrating an SLM, *e.g.*, a semi-transparent liquid Crystal display, liquid crystal on silicon [88], or Digital Micro-mirror Devices (DMD), as used in the single-pixel cam-

(a) Ground truth
Intensity = 700 mA
$M/N = 100\%$

(b) P^1
Intensity = 700 mA
$M/N = 8\ \%$
SRE = 12.93 dB

(c) P^2
Intensity = 100 mA
$M/N = 32\ \%$
SRE = 7.86 dB

(d) P^3
Intensity = 700 mA
$M/N = 14\ \%$
SRE = 14.53 dB

(e) P^4
Intensity = 200 mA
$M/N = 35\ \%$
SRE = 12.30 dB

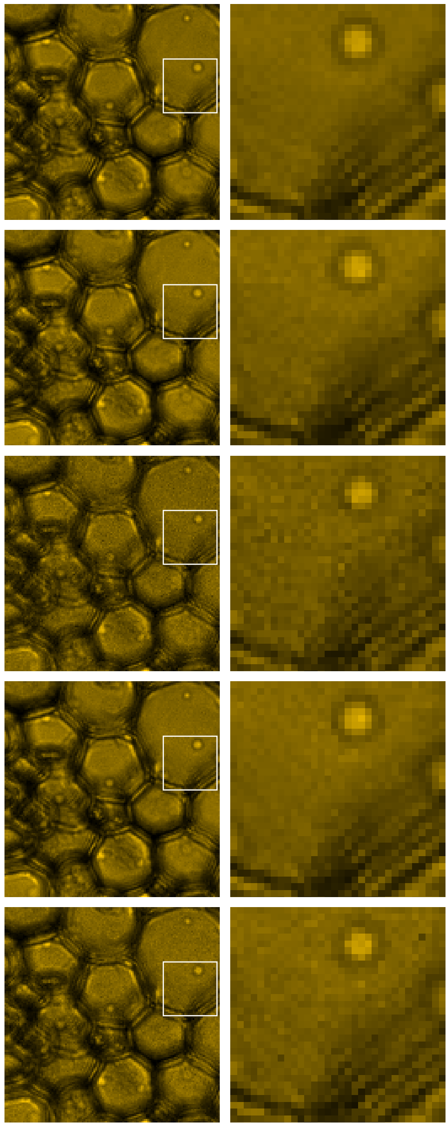

Figure 4.12: The spatial maps of the reconstructed HS volumes at $l_\nu = 70$ (equivalent to 594 nm wavelength). Recall that in CI-FTI only the OPD dimension is subsampled. Therefore, the spatial configuration of the reconstructed images are preserved. However, the images in the third and fifth rows have lower quality, since they are reconstructed from CI-FTI measurements with lower SNR.

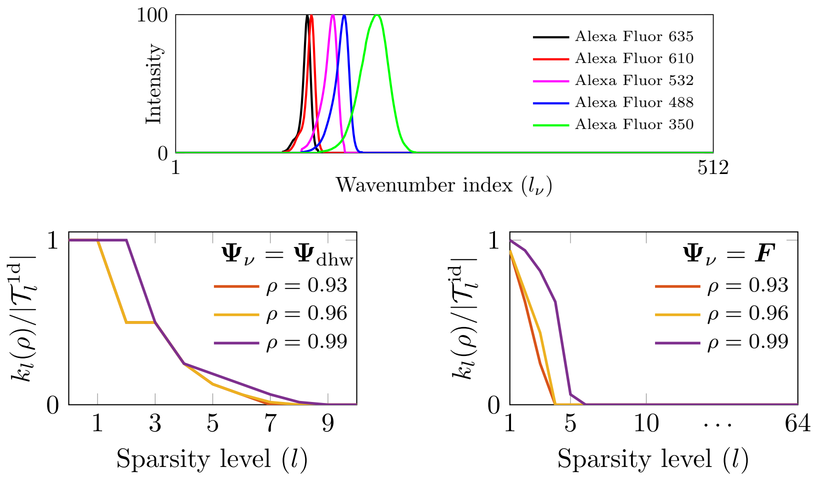

Figure 4.13: The spectra of five fluorochromes (top); The estimated local sparsity ratio for a collection of 38 fluorochromes spectra (bottom).

era [34]. While their inclusion in a compressive imaging procedure is often beneficial, these devices are also known to induce non-negligible light diffraction associated with the small pixel pitch of the SLMs' elements. We therefore postpone the analysis of an actual SI-FTI to future study where the impact of light diffraction (*e.g.*, modeled by a spatial convolution with a calibrated point-spread-function) must be carefully integrated in the sensing model.

4.8.4 Improvement of the Signal Recovery Performance in CI-FTI using MDS

Let us now evaluate the performance of CI-FTI using the two driven MDS approaches in Thm. 4.21.

To design the MDS scheme, we first form a dictionary $\boldsymbol{D}^{\rm f} \in \mathbb{R}^{N_\xi \times N_{\rm f}}$ by collecting the spectra of $N_{\rm f}$ fluorochromes commonly used in fluorescence spectroscopy [65]; see Fig. 4.13-top which includes the spectra of Alexa Fluors. To estimate the values of local sparsity $k_l = k_l(\rho)$ we follow the same procedure as in § 3.4. We conducted that procedure

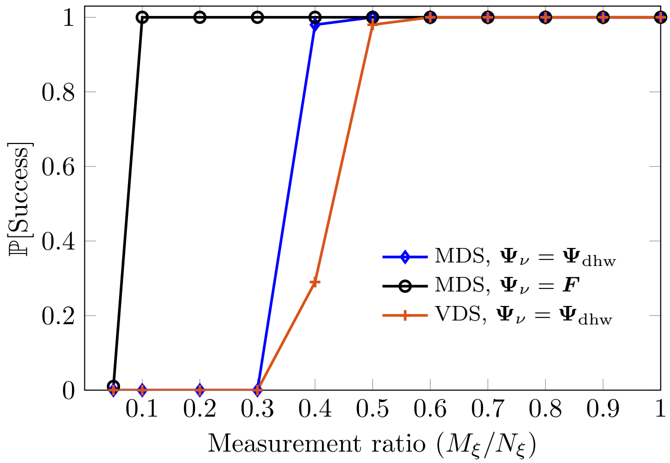

Figure 4.14: Comparison of the successful recovery rate between the MDS- and VDS-based CI-FTI systems supported in Thm. 4.21 and Thm. 4.12, respectively, with $(N_\xi, N_\mathrm{p}) = (1024, 1)$.

with $\rho \in \{0.93, 0.96, 0.99\}$, $\boldsymbol{\Psi}_\nu \in \{\boldsymbol{\Psi}_{\mathrm{dhw}}, \boldsymbol{F}_{N_\xi}\}$, $\mathcal{S} \in \{\mathcal{T}^{\mathrm{1d}}, \mathcal{T}^{\mathrm{id}}\}$ defined in § 4.7, $N_\xi = 1024$, $N_\mathrm{f} = 38$, and $q_0 = 6$. We observe in Fig. 4.13-bottom that fluorochromes spectra display a structured sparsity in both 1-D discrete Fourier and DHW bases. In the DHW basis, local sparsity ratio $(k_l(\rho)/|\mathcal{T}_l^{\mathrm{1d}}|)$ is decreasing in the wavelet level. Moreover, in the discrete Fourier basis, even for a high accuracy, *i.e.*, $\rho = 0.99$, all the non-zero coefficients are located in the first six levels, *i.e.*, in less than 10% of the coefficients. This compact representation in the Fourier basis is the main reason for superior HS data reconstruction that will follow.

We now compare the phase transition curves of successful signal recovery, between the two MDS schemes given by Thm. 4.21 and the VDS scheme given by Thm. 4.12 with $N_\mathrm{p} = 1$, in the absence of noise. The measurements are formed as $\boldsymbol{y} = \boldsymbol{P}_{\Omega^\xi}\boldsymbol{F}^*\boldsymbol{x} \in \mathbb{R}^{M_\xi}$, where $\boldsymbol{x} = \boldsymbol{\Psi}\boldsymbol{H}\boldsymbol{g}$ is a synthetic spectrum resulting from the linear mixing of N_f spectra, as realized by $\boldsymbol{g} \in [0,1]^{N_f}$ with $g_i \sim_{\mathrm{iid}} \mathcal{U}([0,1])$. The $(\mathcal{W}, \boldsymbol{m})$-MDS subsampling sets Ω^ξ are generated depending on the local sparsity values shown in Fig. 4.13. For the VDS scheme, Fig. 4.14 shows the probability

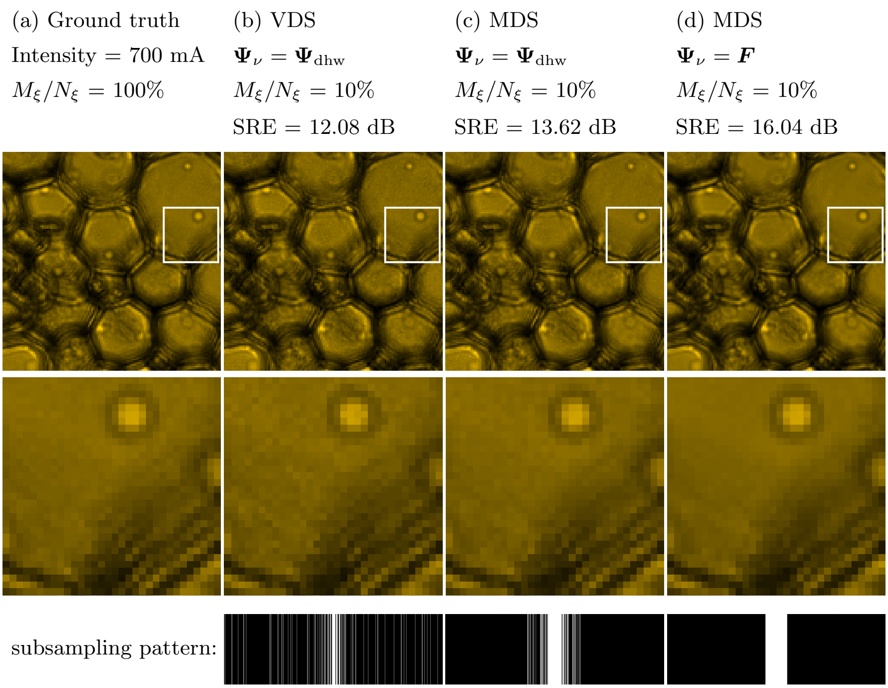

Figure 4.15: The spatial maps (at 594 nm wavelength) of the reconstructed HS volumes from 10% of the measurements. The coding pattern is shown in the bottom row.

of successful reconstruction over 100 independent realizations of Ω^ξ and $\boldsymbol{g}$. We count a recovery as successful if $\|\hat{\boldsymbol{x}} - \boldsymbol{x}\|^2 \leq 10^{-4}\|\boldsymbol{x}\|^2$.

The improvement of signal recovery with the two MDS schemes, over the VDS scheme, is due to the structured sparsity exploited for the sampling strategy. In addition, since the spectra are highly compressible in the Fourier basis, as seen in Fig. 4.13, the successful recovery rate is boosted when $\boldsymbol{\Psi}_\nu = \boldsymbol{F}$.

The Next experiment is an extension of the simulations in Fig. 4.12. We here test the performance of the proposed MDS schemes for CI-FTI on the real Nyquist FTI measurements used in § 4.8.3. We here form CI-FTI measurement by subsampling 10% of the Nyquist FTI measurements. The estimated local sparsity values in Fig. 4.13 are used for the design of the MDS schemes.

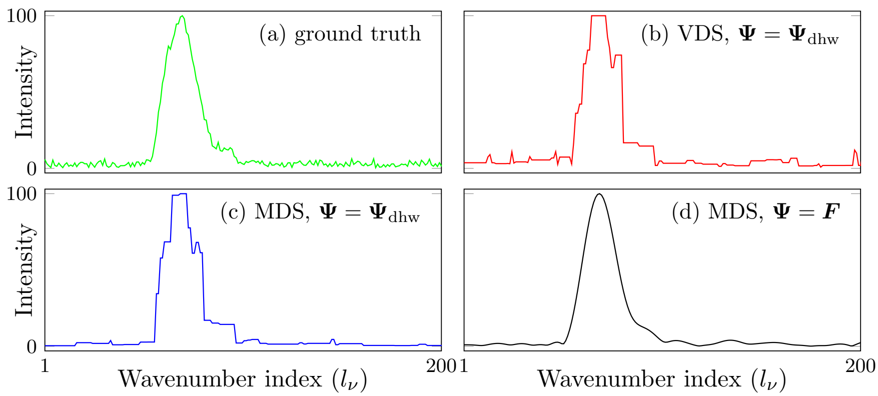

Figure 4.16: The spectra at the center pixel of the HS data shown in Fig. 4.15, here restricted to the first 200 indices of wavenumber axis.

The reconstructed HS volumes are illustrated in Fig. 4.15. Recall that in CI-FTI, the spatial dimension of the Nyquist FTI measurements is not subsampled; hence, the spatial configuration of the specimen is preserved. We have thus to assess the quality of the reconstructed spectra. As evident in Fig. 4.16, the proposed two MDS approaches yield superior reconstruction quality, compared to the VDS approach. Neglecting the noise elements in the ground truth spectrum, we notice that the shape of the reference spectrum is accurately preserved when we set $\boldsymbol{\Psi}_\nu = \boldsymbol{F}$; this choice of sparsity basis gives more compressible representation of the smooth spectra, as seen in Fig. 4.13.

4.9 Proofs

4.9.1 Proof of Prop. 4.10

The proof of Prop. 4.10 requires us to compute here the local coherence between the 1-D discrete Fourier and DHW bases. We follow the developments of [61, Cor. 6.4], improving them by a factor of 9π, which leads to a smaller hidden constant in the sample-complexity bound. Let us first recall the definition of the 1-D discrete Fourier basis of $\mathbb{C}^N$, for some dimension $N \in \mathbb{N}$.

Definition 4.23 (1-D discrete Fourier basis). *Fix $N = 2^r$ for some $r \in \mathbb{N}$. The 1-D discrete Fourier basis of $\mathbb{C}^N$ consists of the functions*

$$\left\{\phi_j^{\mathrm{dft}}(\tau) := \frac{1}{\sqrt{N}} e^{-2\pi i \frac{j}{N}\tau} : -\frac{N}{2}+1 \le j \le \frac{N}{2},\ \tau \in [\![N-1]\!]_0\right\},$$

with j and τ integers.

To get simpler notation, in this proof we write N and l for N_ξ and l_ξ, respectively. Recall the definition of the DHW basis in (3.1). Let us now bound the local coherence

$$\mu_l^{\mathrm{loc}}(\boldsymbol{F}^*\boldsymbol{\Psi}_{\mathrm{dhw}}) = \max_{1\le j\le N} |\langle \phi_{l-(N/2)}^{\mathrm{dft}}, \psi_j^{\mathrm{1d}}\rangle|,$$

for $l \in [\![N]\!]$. Given $l' = l - N/2$ with $-\frac{N}{2}+1 \le l' \le \frac{N}{2}$, we have to compute three cases: *(i)* $|\langle \phi_{l'}^{\mathrm{dft}}, \bar{h}\rangle|$ for all l', *(ii)* $|\langle \phi_0^{\mathrm{dft}}, h_{s,p}^{(1)}\rangle|$ for all s, p, and *(iii)* $|\langle \phi_{l'}^{\mathrm{dft}}, h_{s,p}^{(1)}\rangle|$ for all non-zero l' and all s, p. For the first two cases, we observe that $|\langle \phi_{l'}^{\mathrm{dft}}, \bar{h}\rangle|$ equals one if $l' = 0$ and zero otherwise, while $|\langle \phi_0^{\mathrm{dft}}, h_{s,p}^{(1)}\rangle| = 0$, for all s and p. For the third case, if $l' \neq 0$, a direct computation provides $z := \langle \phi_{l'}, h_{s,p}\rangle = z_1 - z_2$ with

$$z_1 := \sum_{j=p2^{r-s}}^{(p+\frac{1}{2})2^{r-s}-1} 2^{\frac{s-r}{2}} 2^{-\frac{r}{2}} e^{-2\pi i 2^{-r} l' j},$$

$$z_2 := \sum_{j=(p+\frac{1}{2})2^{r-s}}^{(p+1)2^{r-s}-1} 2^{\frac{s-r}{2}} 2^{-\frac{r}{2}} e^{-2\pi i 2^{-r} l' j}.$$

Separate computations of z_1 and z_2 yields

$$z_1 = 2^{\frac{s}{2}-r} e^{-2\pi i l' p 2^{-s}} \sum_{j=0}^{2^{r-s-1}-1} e^{-2\pi i 2^{-r} l' j},$$
$$z_2 = 2^{\frac{s}{2}-r} e^{-2\pi i l'(p+\frac{1}{2})2^{-s}} \sum_{j=0}^{2^{r-s-1}-1} e^{-2\pi i 2^{-r} l' j}.$$

Therefore, we get

$$\begin{aligned}
|z| &= |z_1 - z_2| \\
&= 2^{\frac{s}{2}-r} \Big| \sum_{j=0}^{2^{(r-s-1)}-1} e^{-2\pi i 2^{-r} l' j} \Big| \left| 1 - e^{-2\pi i l' 2^{-(s+1)}} \right| \\
&= 2^{\frac{s}{2}-r} \frac{2\sin^2(\pi l' 2^{-s-1})}{|\sin(\pi l' 2^{-r})|} \leq 2^{\frac{s}{2}} \frac{\sin^2(\pi l' 2^{-s-1})}{|l'|}.
\end{aligned}$$

In the last inequality we used the fact that $|\sin(\pi x)| \geq 2|x|$ for $|x| \leq 1/2$. Let us consider two cases. First, if $\pi|l'|/2^{-s} \geq \pi/2$, then $2^{s/2} < \sqrt{2|l'|}$ and $s \leq \sqrt{2/|l'|}$. Second, if $\pi|l'|/2^{-s} < \pi/2$, then using the fact that $\cos|x| \geq 1 - 2/\pi|x|$ for $|x| < \pi/2$ we get $\sin^2 \pi l' 2^{-s-1} = 1 - \cos \pi l' 2^{-s} \leq |l'|2^{-s}$ which leads to $s \leq 2^{-s/2} \leq 1/\sqrt{2|l'|}$. Combining the two cases leads to

$$|\langle \phi_{l'}^{\mathrm{dft}}, h_{s,p}^{(1)} \rangle| \leq \frac{\sqrt{2}}{\sqrt{|l'|}}.$$

Gathering all results, we thus find $\max_j |\langle \phi_{l'}^{\mathrm{dft}}, \psi_j^{\mathrm{1d}} \rangle| \leq \min\left\{1, \frac{\sqrt{2}}{\sqrt{|l'|}}\right\}$, and

$$\mu_l^{\mathrm{loc}}(\boldsymbol{F}^* \boldsymbol{\Psi}_{\mathrm{dhw}}) \leq \kappa_l := \sqrt{2} \min\left\{1, \frac{1}{\sqrt{|l - N/2|}}\right\}. \tag{4.53}$$

Therefore, (2.8) gives $p(l) := C_N \min\left\{1, \frac{1}{|l-N/2|}\right\}$ for $l \in [\![N]\!]$, which implies (4.20) for $N = N_\xi$, with C_N ensuring $\sum_{l=1}^{N} p(l) = 1$. Concerning this constant, we find

$$C_N^{-1} = \sum_l \min\{1, \frac{1}{|l - \frac{N}{2}|}\} = 1 + \sum_{l=1}^{\frac{N}{2}-1} \frac{1}{\frac{N}{2} - l} + \sum_{l=\frac{N}{2}+1}^{N} \frac{1}{l - \frac{N}{2}}$$
$$= 1 + \sum_{l=1}^{\frac{N}{2}-1} \frac{1}{l} + \sum_{l=1}^{\frac{N}{2}} \frac{1}{l} = 1 + \frac{2}{N-2} + 2\sum_{l=1}^{\frac{N}{2}-1} \frac{1}{l}.$$

However, for any integer $D \geq 2$, $\sum_{l=1}^{D-1} \frac{1}{l} \leq 1 + \int_2^D \frac{1}{s-1}\mathrm{d}s = 1 + \log(D-1)$ and $\sum_{l=1}^{D-1} \frac{1}{l} \geq \int_1^D \frac{1}{s}\mathrm{d}s = \log(D)$. Therefore,

$$\begin{aligned} 2\log(\frac{N}{2}) &\leq 1 + \frac{2}{N-2} + 2\log(\frac{N}{2}) \\ &\leq C_N^{-1} \leq 3 + \frac{2}{N-2} + 2\log(\frac{N}{2} - 1) \\ &< 4 + 2\log(\frac{N}{2}), \end{aligned} \tag{4.54}$$

which is true for $N \geq 2$. Finally, from the definition of $\boldsymbol{\kappa}$, we have $\|\boldsymbol{\kappa}\|^2/2 \leq C_N^{-1}$, so that $\|\boldsymbol{\kappa}\|^2 \leq 8 + 4\log(\frac{N}{2}) \lesssim \log N$, and (2.7) then explains the sufficient condition (4.22) for $N = N_\xi$.

4.9.2 Proof of Prop. 4.15

To get simpler notation, we here write k, j, and l for $l_{\mathrm{hs}}, l_{\mathrm{p}}$, and l_ξ respectively. This proof requires us to compute a bound on the local coherence $\mu_k^{\mathrm{loc}}(\boldsymbol{\Phi}_{\mathrm{fti}}^* \boldsymbol{\Psi}_{\mathrm{si}})$ for $k \in [\![N_{\mathrm{hs}}]\!]$.

We first note that, from the Kronecker product properties, $\boldsymbol{\Phi}_{\mathrm{fti}}^* \boldsymbol{\Psi}_{\mathrm{si}} = (\boldsymbol{I}_{N_{\mathrm{p}}} \boldsymbol{\Psi}_{\mathrm{idhw}}) \otimes (\boldsymbol{F}^* \boldsymbol{\Psi}_{\mathrm{dhw}})$. Using the relation $k \xrightleftharpoons{N_\xi, N_{\mathrm{p}}} (l, j)$ between the 1D "k" and the 2D "(l, j)" index representations of interferometric data and (3.20b), the definition of local coherence (2.6) gives

$$\mu_k^{\mathrm{loc}}(\boldsymbol{\Phi}_{\mathrm{fti}}^* \boldsymbol{\Psi}_{\mathrm{si}}) = \underbrace{\max_{l'} |(\boldsymbol{F}^* \boldsymbol{\Psi}_{\mathrm{dhw}})_{l,l'}|}_{=: \mu_l^{\mathrm{loc}}(\boldsymbol{F}^* \boldsymbol{\Psi}_{\mathrm{dhw}})} \cdot \underbrace{\max_{j'} |(\boldsymbol{I}_{N_{\mathrm{p}}} \boldsymbol{\Psi}_{\mathrm{idhw}})_{j,j'}|}_{=: \mu_j^{\mathrm{loc}}(\boldsymbol{I}_{N_{\mathrm{p}}} \boldsymbol{\Psi}_{\mathrm{idhw}})}. \tag{4.55}$$

Let us now bound the two terms of right-hand side of (4.55). For the first one, (4.53) provides

$$\mu_l^{\mathrm{loc}}(\boldsymbol{F}^*\boldsymbol{\Psi}_{\mathrm{dhw}}) \leq \sqrt{2}\,\min\Big\{1, \frac{1}{\sqrt{|l-N_\xi/2|}}\Big\}. \tag{4.56}$$

Concerning the second term, we now show that $\mu_j^{\mathrm{loc}}(\boldsymbol{I}_{N_\mathrm{p}}\boldsymbol{\Psi}) = 1/2$ for the two possible types of 2-D wavelet constructions, *i.e.*, $\boldsymbol{\Psi} \in \{\boldsymbol{\Psi}_{\mathrm{idhw}}, \boldsymbol{\Psi}_{\mathrm{adhw}}\}$, defined in § 3.2.

Identifying in these definitions N with $\bar{N}_\mathrm{p}$ (and thus N^2 with $N_\mathrm{p} = \bar{N}_\mathrm{p}^2$), we can now proceed and bound the local coherence of these two bases when the sensing basis is the Dirac basis $\boldsymbol{I}_{N_\mathrm{p}}$.

(i) ADHW case: In this case, from the definition of the 2-D ADHW matrix in § 3.2, *i.e.*, $\boldsymbol{\Psi}_{\mathrm{adhw}} = \boldsymbol{\Psi}_{\mathrm{dhw}} \otimes \boldsymbol{\Psi}_{\mathrm{dhw}}$, and since $\boldsymbol{I}_{N_\mathrm{p}} = \boldsymbol{I}_{\bar{N}_\mathrm{p}} \otimes \boldsymbol{I}_{\bar{N}_\mathrm{p}}$, using (3.20b) we find

$$\begin{aligned}
\mu_j^{\mathrm{loc}}(\boldsymbol{I}_{N_\mathrm{p}}\boldsymbol{\Psi}_{\mathrm{2D}}) &= \mu_j^{\mathrm{loc}}\big((\boldsymbol{I}_{\bar{N}_\mathrm{p}}\boldsymbol{\Psi}_{\mathrm{dhw}}) \otimes (\boldsymbol{I}_{\bar{N}_\mathrm{p}}\boldsymbol{\Psi}_{\mathrm{dhw}})\big)\\
&= \underbrace{\max_{j_1'}|(\boldsymbol{I}_{\bar{N}_\mathrm{p}}\boldsymbol{\Psi}_{\mathrm{dhw}})_{j_1,j_1'}|}_{=:\mu_{j_1}^{\mathrm{loc}}(\boldsymbol{I}_{\bar{N}_\mathrm{p}}\boldsymbol{\Psi}_{\mathrm{dhw}})} \cdot \underbrace{\max_{j_2'}|(\boldsymbol{I}_{\bar{N}_\mathrm{p}}\boldsymbol{\Psi}_{\mathrm{dhw}})_{j_2,j_2'}|}_{=:\mu_{j_2}^{\mathrm{loc}}(\boldsymbol{I}_{\bar{N}_\mathrm{p}}\boldsymbol{\Psi}_{\mathrm{dhw}})},
\end{aligned}$$

where we invoked the rule $j \xrightleftharpoons{\bar{N}_\mathrm{p}} (j_1, j_2)$ between the 1-D pixel indexing $j \in [\![N_\mathrm{p}]\!]$ and the 2-D pixel indexing $j_1, j_2 \in [\![\bar{N}_\mathrm{p}]\!]$. Let $\boldsymbol{e}_k$ be the k^{th} element of the Dirac basis of $\mathbb{C}^{\bar{N}_\mathrm{p}}$. Setting $N = \bar{N}_\mathrm{p}$ in Def. 3.1, we find $|(\boldsymbol{I}_{\bar{N}_\mathrm{p}}\boldsymbol{\Psi}_{\mathrm{dhw}})_{j_1,j_1'}| = |\langle e_{j_1}, \psi_{j_1'}^{\mathrm{1d}}\rangle|$ and

$$|\langle e_{j_1}, \psi_{j_1'}^{\mathrm{1d}}\rangle| = \begin{cases} 2^{\frac{-r}{2}}, & \text{if } \psi_{j_1'}^{\mathrm{1d}} = \bar{h},\\ 2^{\frac{n-r}{2}}, & \text{if } \psi_{j_1'}^{\mathrm{1d}} = h_{s,p}^{(1)},\ 0 \leq s \leq r-1,\ 0 \leq p \leq 2^s - 1, \end{cases}$$

where $r = \log_2(\bar{N}_\mathrm{p})$. Therefore, the maximum of the expression above is reached for $s = r-1$ and

$$\mu_{j_1}^{\mathrm{loc}}(\boldsymbol{I}_{\bar{N}_\mathrm{p}}\boldsymbol{\Psi}_{\mathrm{dhw}}) = \mu_{j_2}^{\mathrm{loc}}(\boldsymbol{I}_{\bar{N}_\mathrm{p}}\boldsymbol{\Psi}_{\mathrm{dhw}}) = 1/\sqrt{2}$$

and $\mu_j^{\mathrm{loc}}(\boldsymbol{I}_{N_\mathrm{p}}\boldsymbol{\Psi}_{\mathrm{adhw}}) = 1/2$.

(ii) IDHW case: Similarly to the developments above, by setting $N = \bar{N}_\mathrm{p}$ and $r = \log_2(\bar{N}_\mathrm{p})$ in the definition of the IDHW basis in § 3.2, we find

$$|(\boldsymbol{I}_{N_\mathrm{p}}\boldsymbol{\Psi}_{\mathrm{idhw}})_{j,j'}| = |\langle e_j, \psi_{j'}^{\mathrm{iso}}\rangle| \;=\; \begin{cases} 2^{-r} & \text{if } \psi_{j'}^{\mathrm{iso}} = \phi^{(0,0)}, \\ 2^{s-r} & \text{if } \psi_{j'}^{\mathrm{iso}} = \phi_{s,(p_1,p_2)}^{(ab)}, \end{cases}$$

for some $(a,b) \in \{0,1\}^2 \backslash \{0.0\}$, $0 \leq s \leq r-1$, and $0 \leq p_1, p_2 \leq 2^s - 1$. Therefore, its maximum is reached for $s = r-1$ and $\mu_j^{\mathrm{loc}}(\boldsymbol{I}_{N_\mathrm{p}}\boldsymbol{\Psi}_{\mathrm{idhw}}) = 1/2$.

Combining the last two cases with (4.56), and (4.55) results in

$$\mu_k^{\mathrm{loc}}(\boldsymbol{\Phi}_{\mathrm{fti}}^*\boldsymbol{\Psi}_{\mathrm{si}}) \leq \kappa_k := \frac{\sqrt{2}}{2}\min\Big\{1, \frac{1}{\sqrt{|l - N_\xi/2|}}\Big\}, \tag{4.57}$$

which provides (4.28).

The pmf associated with SI-FTI can be then formulated from (4.57) and (4.58) as

$$p(k) = p(j,l) = \frac{C_{N_\xi}}{N_\mathrm{p}}\min\Big\{1, \frac{1}{|l - N_\xi/2|}\Big\},$$

where the normalizing constant C_{N_ξ}, *i.e.*, such that $\sum_k p(k) = 1$, is formulated in § 4.9.1. Moreover,

$$\|\boldsymbol{\kappa}\|^2 = \frac{1}{2}N_\mathrm{p}\sum_{l=1}^{N_\xi}\min\Big\{1, \frac{1}{|l - N_\xi/2|}\Big\} = \frac{1}{2}N_\mathrm{p}C_{N_\xi}^{-1},$$

so that, from (4.54),

$$\|\boldsymbol{\kappa}\|^2 \leq N_\mathrm{p}(2 + \log(N_\xi/2)) \lesssim N_\mathrm{p}\log N_\xi. \tag{4.58}$$

4.9.3 Proof of Prop. 4.20

We first note $\boldsymbol{F}_{N_\xi}^* \boldsymbol{F}_{N_\xi} = \boldsymbol{I}_{N_\xi}$ and $\boldsymbol{P}_{\mathcal{T}_t^{\mathrm{id}}} \boldsymbol{I}_{N_\xi} \boldsymbol{P}_{\mathcal{T}_l^{\mathrm{id}}}^\top = \boldsymbol{I}_{|\mathcal{T}_t^{\mathrm{id}}|}$ if $t = l$, and $\boldsymbol{0}$ otherwise. From the definitions of the relative sparsity in (2.11), and using Lemma 2.5, we find

$$\begin{aligned} K_t^{\mathcal{T}^{\mathrm{id}},\mathcal{T}^{\mathrm{id}}}(\boldsymbol{I}_{N_\xi}, \boldsymbol{k})^{\frac{1}{2}} &\leq \sum_{l=1}^{\bar{r}} \|\boldsymbol{P}_{\mathcal{T}_t^{\mathrm{id}}} \boldsymbol{P}_{\mathcal{T}_l^{\mathrm{id}}}^\top\|_{2,2}\, k_l^{\frac{1}{2}} \\ &= \|\boldsymbol{I}_{|\mathcal{T}_t^{\mathrm{id}}|}\|_{2,2}\, \delta_{t,l}\, k_l^{\frac{1}{2}} = k_t^{\frac{1}{2}}. \end{aligned} \tag{4.59}$$

Moreover, $\mu(\boldsymbol{P}_{\mathcal{T}_t^{\mathrm{id}}} \boldsymbol{I}_{N_\xi} \boldsymbol{P}_{\mathcal{T}_l^{\mathrm{id}}}^\top) = \mu(\boldsymbol{I}_{|\mathcal{T}_t^{\mathrm{id}}|})\, \delta_{t,l} = \delta_{t,l}$ and $\mu(\boldsymbol{P}_{\mathcal{T}_t^{\mathrm{id}}} \boldsymbol{I}_{N_\xi}) = 1$. Therefore, from the definitions of the multilevel coherence in (2.12), we find

$$\mu_{t,l}^{\mathcal{T}^{\mathrm{id}},\mathcal{T}^{\mathrm{id}}}(\boldsymbol{I}_{N_\xi}) = \delta_{t,l}. \tag{4.60}$$

By setting $\mathcal{W} = \mathcal{S} = \mathcal{T}^{\mathrm{id}}$ in Thm. 2.6, and by using the estimations in (4.59) and (4.60), the sample-complexity bounds in (2.15), (2.13), and (2.14) reduce to

$$m_t \gtrsim \frac{N}{\bar{r}} k_t \log(K\epsilon^{-1})\, \log(N_\xi), \text{ for } t \in [\![r]\!].$$

Noting that we must have $m_t \leq N/\bar{r}$ the proof is complete.

4.10 Discussion

In this chapter, we have proposed two versions of compressive FTI (*i.e.*, CI-FTI and SI-FTI) where the light exposure can be compressed temporally and spatially in order to maximize the spectral resolution of the resulting hyperspectral volumes, *i.e.*, in a process that minimizes the photo-bleaching phenomenon. These methods are practically plausible without any modification of MI but only of the optical setup to which it is associated (*e.g.*, the light system of a confocal microscope).

The first proposed system, called Coded illumination-FTI, consists in binary modulation of illuminating light before exposing the biological specimen; while the second system, referred to as Structured Illumination-FTI, involves a spatial light-modulation that allows spatiotemporal light coding. By invoking the theory of compressive sensing and by deriving a VDS strategy guided by [61], the two proposed frameworks are proved to reach efficient uniform recovery guarantees. Furthermore, the impact of promoting fluorochrome life-time as a constraint has been analyzed.

We have also presented two improved coding designs for CI-FTI that are adapted to fluorescence spectroscopy experiment. These schemes are derived from the notions of sparsity-in-levels, MDS, and multilevel coherence introduced in [7].

Our theoretical analyses were verified via exhaustive numerical tests for four cases: *(i)* recovery of synthetic sparse HS data in the absence of noise, *(ii)* recovery of simulated biological HS volume in the presence of noise, *(iii)* recovery of an actual HS data from CI-FTI measurements, which are simulated from experimental Nyquist FTI measurement, and *(iv)* designing an MDS strategy for CI-FTI based on the common sparsity structure among different fluorochrome spectra.

Future investigations can be followed in several directions. For instance, assuming other discrepancy sources, *e.g.*, Poisson noise, quantization, instrumental response. Tracing other low-complexity prior models, *e.g.*, low-rankness [38], group sparsity, dictionary-based sparsity, shearlet [62], or TV sparsity models will be the scope of a future work. An other line of research would be to improve the sampling strategy of SI-FTI for its application in fluorescence spectroscopy using an MDS scheme, as we did for CI-FTI in § 4.7.

Chapter 5

Compressive Single-Pixel Fourier Transform Interferometry

Having discussed about a conventional FTI system, consisting of a 2-D pixel array, and its two CS variations in Chapter 4, we now turn our attention to a single-pixel FTI system.

Single-pixel imaging is now a reality in many applications, like biomedical ultra-thin endoscope and fluorescence spectroscopy. In this context, many schemes exist to improve the light throughput of these devices, *e.g.,* using structured illumination driven by CS theory. In this chapter, we consider the combination of single-pixel imaging with Fourier Transform Interferometry (SP-FTI), to reach high-resolution HS imaging, as desirable, *e.g.,* in fluorescence spectroscopy. While this association is not new, we here focus on an efficient spatial illumination, structured as Hadamard patterns, during the optical path progression. We follow a VDS strategy for space-time coding of the light illumination, and show theoretically and numerically that this scheme allows us to reduce the number of measurements and light-exposure of the observed object compared to conventional SP-FTI.

The content of this chapter has been published in [81, 82, 84].

5.1 Definition of the Problem

Nowadays, single-pixel imaging has become an emerging paradigm for capturing high-quality images using a single photo-detector [28, 34, 50, 104] in a low-cost and high light throughput process, with low memory requirement. These advantages make single-pixel imaging an excellent candidate for recording HS volumes [99, 106], in particular, when combined with the FTI [59].

Similar to the FTI described in § 4.2, SP-FTI also works on the principle of the MI; hence its ability to provide high-resolution HS data, *e.g.*, for biomedical imaging applications, is limited by the tolerance of the biological elements against the light exposure. Fortunately, as discussed in Chapter 4, the theory of CS has shown successful results in reducing the amount of light exposure during the interferometry acquisition, while preserving the spectral resolution.

Unlike the CI-FTI and SI-FTI models, we here propose an FTI-based HS acquisition that is constrained to use single-pixel imaging. In SP-FTI, the light distribution is coded (or structured) in both OPD (or time) and spatial domains before being integrated into a single beam (unlike SI-FTI), using, *e.g.*, collimating optics (see § 5.2). While the first SP-FTI has been implemented in food monitoring application [59], its theoretical analysis is not covered in the literature.

In this chapter, adopting the stable and robust sampling strategies of Krahmer and Ward [61] (see § 2.3) for compressive imaging, we develop an efficient SP-FTI sensing by following a VDS of both the OPD domain and the Hadamard transformation of the spatial domain. It is worth mentioning that in the proof of concept study[1] [81] we followed an MDS scheme applied separately on the OPD and Hadamard domains. In particular, our sampling strategy here relies on the estimation of tight bounds on the local coherence between the 3-D sensing and sparsity

[1]Deriving the stable and robust sampling strategy for the MDS-based SP-FTI is not covered in this thesis.

bases. The sensing basis is obtained from the Kronecker product of the Fourier (imposed by the MI) and Hadamard (an aspect of our sensing design) bases, and the sparsity basis is achieved by the Kronecker product of the 1-D DHW and 2-D IDHW bases. The genuine results in this chapter are built upon our analysis in Chapter 3.

In the following, after describing the principles of SP-FTI, our stable and robust compressive SP-FTI is proposed in § 5.3. Numerical simulations are provided in § 5.4. We refer the reader to § 4.1.1 for the related works.

5.2 Acquisition Model in Nyquist SP-FTI

We now describe the acquisition procedure of SP-FTI originated in [59]. By an abuse of terminology, SP-FTI refers in this thesis to the Nyquist implementation of SP-FTI, *i.e.*, collecting as many observations as the number of HS volume voxels so that HS reconstruction can be achieved by a linear (inverse) transform. In contrast, in the literature of imaging science, the term SP-FTI (or in general single-pixel imaging system [34]) implies a modality which operates with sub-Nyquist sampling rate.

SP-FTI operates on the principle of the FTI explained in § 4.2, except that *(i)* the 2-D sensor in the FTI is replaced by a single-pixel sensor and *(ii)* SP-FTI includes a collimator and an SLM. For the sake of succinctness in this chapter, we neglect the explanation of the parts that are identical to the FTI model explained in § 4.2.2.

As shown in Fig. 5.1, a parallel light beam is first spatially structured using an SLM before traveling through a biological specimen. After being integrated into a single beam by means of an optical collimator, the obtained beam enters the MI and undergoes a transformation explained in § 4.2.1. The intensity of the outgoing beam from the MI is next acquired by a single-pixel sensor; each measurement captured by this sensor corresponds to one position of the moving mirror and one SLM

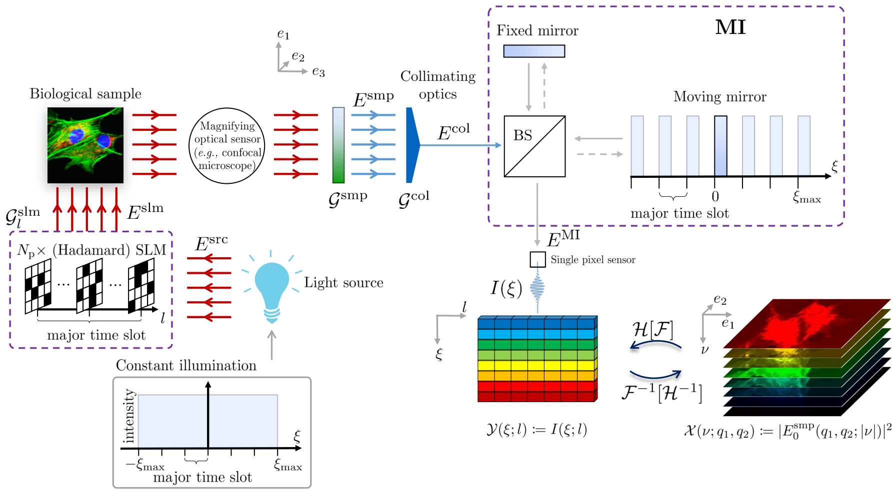

Figure 5.1: Operating principle of SP-FTI.

pattern. At each position of the moving mirror, this process is repeated for multiple SLM patterns.

5.2.1 Continuous Observation Model of SP-FTI

Let us now develop an idealized optical SP-FTI model for acquiring an HS volume. We follow the same conventions as in § 4.2. Our assumptions here about the light source, biological sample, and other optical elements follow Assumptions 4.1, 4.2, and 4.3.

Similar to the FTI, we start by considering a monochromatic plane wave $E^{\text{src}}(\boldsymbol{q};\nu;t) = E_0^{\text{src}}(q_1,q_2;\nu)e^{i(2\pi q_3\nu-\omega t)}$, *i.e.*, traveling along $\boldsymbol{e}_3$-direction, at some point $\boldsymbol{q} \coloneqq [q_1,q_2,q_3]^\top \in \mathbb{R}^3$ and time t. This beam is next spatially structured by an SLM.

Remark 5.1. *For the sake of simplicity and a fair comparison of different MI-based systems, we assume that the SLM here and the 2-D imaging sensor in the FTI system in Chapter 4 have identical resolutions,* i.e., $\bar{N}_{\text{p}} \times \bar{N}_{\text{p}}$ *pixels of length* Δ_{p}.

Let $\mathcal{G}_l^{\rm slm} : f(q_1, q_2; \nu) \mapsto G_l^{\rm slm}(q_1, q_2) \cdot f(q_1, q_2; \nu)$ be the transfer function of the SLM, where

$$G_l^{\rm slm}(q_1, q_2) \coloneqq \sum_{j_1, j_2} h^{(l)}(j_1, j_2) \cdot \mathrm{rect}(\frac{q_1}{\Delta_{\rm p}} - j_1, \frac{q_2}{\Delta_{\rm p}} - j_2). \tag{5.1}$$

In (5.1), l indexes the SLM patterns, $j_1, j_2 \in [\![N_{\rm p}]\!]$ represent the pixel index on the SLM, and $h^{(l)}(j_1, j_2) \in \{0, 1\}$ denotes the value of the $l^{\rm th}$ SLM pattern at location (j_1, j_2). Consequently, the beam after the SLM reads

$$E^{\rm slm}(\boldsymbol{q}; \nu; t; l) = E_0^{\rm slm}(q_1, q_2; \nu; l) e^{i(2\pi q_3 \nu - \omega t)},$$

where

$$\begin{aligned} E_0^{\rm slm}(q_1, q_2; \nu; l) &= \mathcal{G}_l^{\rm slm}[E_0^{\rm src}](q_1, q_2; \nu) \\ &= G_l^{\rm slm}(q_1, q_2) \cdot E_0^{\rm src}(q_1, q_2; \nu). \end{aligned} \tag{5.2}$$

After traveling through the specimen with transfer function $\mathcal{G}^{\rm smp}$ (see Assumption 4.3), this beam writes

$$E^{\rm smp}(\boldsymbol{q}; \nu; t; l) = E_0^{\rm smp}(q_1, q_2; \nu; l) e^{i(2\pi q_3 \nu - \omega t)}.$$

From Assumption 4.3, with $\lambda \leftarrow G_l^{\rm slm}(q_1, q_2)$, and using (5.2), the field intensities $E_0^{\rm slm}$ and $E_0^{\rm smp}$ are related as

$$\begin{aligned} E_0^{\rm smp}(q_1, q_2; \nu; l) &= \mathcal{G}^{\rm smp}[E_0^{\rm slm}](q_1, q_2; \nu; l) \\ &= G_l^{\rm slm}(q_1, q_2) \cdot \mathcal{G}^{\rm smp}[E_0^{\rm src}](q_1, q_2; \nu). \end{aligned} \tag{5.3}$$

Similar to the FTI system, we neglect the refraction of the light induced by the specimen in (5.3). The monochromatic plane wave in (5.3) is then integrated into a monochromatic single-wave by collimating optics. The resulting single-beam reads

$$E^{\rm col}(q_3; \nu; t; l) = E_0^{\rm col}(\nu; l) e^{i(2\pi q_3 \nu - \omega t)}. \tag{5.4}$$

We assume that the field intensity E_0^{col} is related to E_0^{smp} through a general transfer function $\mathcal{G}^{\mathrm{col}} : f(q_1, q_2; \nu; l) \mapsto \bar{f}(\nu; l)$ of the collimator, *i.e.*,

$$E_0^{\mathrm{col}}(\nu; l) = \mathcal{G}^{\mathrm{col}}[E_0^{\mathrm{smp}}](\nu; l). \tag{5.5}$$

Assumption 5.2. *Similar to the works [34, 106, 119] and [18, page 291], we also assume that the field intensity after the ideal collimator is the integral of the input filed intensities with respect to the spatial coordinates,* i.e.,

$$|E_0^{\mathrm{col}}(\nu; l)|^2 = |\mathcal{G}^{\mathrm{col}}[E_0^{\mathrm{smp}}](\nu; l)|^2 = \iint |E_0^{\mathrm{smp}}(q_1, q_2; \nu; l)|^2 \mathrm{d}q_1 \mathrm{d}q_2. \tag{5.6}$$

Next, the beam (5.4) enters the MI. Recall that (4.2) relates the intensity of the outgoing beam from the MI to the intensity of beams entering thereto. By setting $|E^{\mathrm{in}}(q_1, q_2; \nu; t)|^2$ to $|\mathcal{G}^{\mathrm{col}}[E^{\mathrm{smp}}](\nu; l)|^2$ in (4.2), the intensity of the beam after the MI reads

$$I^{\mathrm{MI}}(\xi; \nu; l) = 2\,|\mathcal{G}^{\mathrm{col}}[E_0^{\mathrm{smp}}](\nu; l)|^2(1 + \cos(2\pi\nu\xi)). \tag{5.7}$$

This intensity is later recorded by a single-pixel sensor. With the same reasoning we made for modeling FTI acquisition in § 4.2.2 from (4.8) to (4.9), we develop (5.7) into an interferogram. After removing the mean component of (5.7) and symmetrization of the intensity around $\nu = 0$, the resulting interferogram at a given OPD ξ and associated with the l^{th} SLM pattern can be written as

$$\begin{aligned} I(\xi; l) &= \int_{-\infty}^{+\infty} |\mathcal{G}^{\mathrm{col}}[E^{\mathrm{smp}}](|\nu|; l)|^2 e^{-i2\pi\nu\xi}\,\mathrm{d}\nu \\ &= \int_{-\infty}^{+\infty} \iint |E_0^{\mathrm{smp}}(q_1, q_2; |\nu|; l)|^2 e^{-i2\pi\nu\xi}\,\mathrm{d}q_1\mathrm{d}q_2\mathrm{d}\nu, \end{aligned} \tag{5.8}$$

where in the second line we used (5.6). Similar to the FTI model (4.10), the continuous HS volume is defined as

$$\mathcal{X}(\nu; q_1, q_2) := |\mathcal{G}^{\mathrm{smp}}[E_0^{\mathrm{src}}](q_1, q_2; |\nu|)|^2, \tag{5.9}$$

and unlike (4.11), we define the collection of continuous interferograms as

$$\mathcal{Y}(\xi; l) := I(\xi; l). \tag{5.10}$$

In order to relate $\mathcal{X}$ and $\mathcal{Y}$ we first need to define an auxiliary variable

$$\mathcal{Z}(\nu; l) := \iint |E_0^{\mathrm{smp}}(q_1, q_2; |\nu|; l)|^2 \,\mathrm{d}q_1 \mathrm{d}q_2 \tag{5.11a}$$

$$= \iint G_l^{\mathrm{slm}}(q_1, q_2) \cdot |\mathcal{G}^{\mathrm{smp}}[E_0^{\mathrm{src}}](q_1, q_2; |\nu|)|^2 \,\mathrm{d}q_1 \mathrm{d}q_2, \tag{5.11b}$$

where in the second line we used (5.3). Inserting (5.8) in (5.10) and using (5.11a) gives

$$\mathcal{Y}(\xi; l) := \int_{-\infty}^{+\infty} \mathcal{Z}(\nu; l) e^{-i2\pi\nu\xi} \mathrm{d}\nu = \mathcal{F}[\mathcal{Z}](\xi; l), \tag{5.12}$$

where $\mathcal{F}$ denotes the 1-D Fourier transform. Furthermore, inserting (5.9), (5.2), and (5.1) in (5.11b) yields

$$\mathcal{Z}(\nu; l) = \mathcal{H}[\mathcal{X}](\nu; l), \tag{5.13}$$

where

$$\mathcal{H} : f(q_1, q_2) \mapsto \sum_{j_1, j_2} h^{(l)}(j_1, j_2) \cdot \iint \mathrm{rect}(\frac{q_1}{\Delta_{\mathrm{p}}} - j_1, \frac{q_2}{\Delta_{\mathrm{p}}} - j_2) \cdot f(q_1, q_2) \,\mathrm{d}q_1 \mathrm{d}q_2. \tag{5.14}$$

In words, the function $\mathcal{H}$ is jointly modeling the act of SLM and collimating optics on the light beam. The former is associated with the summation term in (5.14) and the latter is associated with the integral term.

Finally, from (5.12) and (5.13) the SP-FTI model relating $\mathcal{X}$ and $\mathcal{Y}$ reads

$$\mathcal{Y}(\xi; l) = \mathcal{H}[\mathcal{F}[\mathcal{X}]](\xi; l). \tag{5.15}$$

Since $\mathcal{F}$ and $\mathcal{H}$ are a function of only ξ and l, respectively, their order in (5.15) can be flipped. We notice that the $\mathrm{rect}(\cdot)$ function in (5.15) has already discretized the spatial domain, which is due to the pixelated SLM.

Remark 5.3. *There exist different schemes for programming the SLM patterns* $h^{(l)} \in \{0,1\}^{\bar{N}_\mathrm{p} \times \bar{N}_\mathrm{p}}$, e.g., *according to a random Bernoulli distribution [34] or Hadamard patterns [59]. In addition to practical advantages of the Hadamard matrix (*e.g., *it needs less memory and computational resources) several observations in the literature of CS promotes the use of orthonormal bases (*e.g., *Hadamard) over random matrices (*e.g., *Bernoulli); see,* e.g., *the discussion in [99].*

Remark 5.4. *As presented in Fig. 5.1, we hereafter assume that the SLM patterns* $h^{(l)}$*, up to some scaling operations* (5.18)*, are set to the Hadamard patterns,* i.e., *the rows of the Hadamard matrix defined in § 3.2, and* $l \in [\![N_\mathrm{p}]\!]$. *Leveraging the orthogonality of the Hadamard matrix, by programming* N_p *Hadamard patterns at each OPD point* ξ*, the spatially discretized HS volume in* (5.15) *can be obtained by (interchangeably) computing the inverse Fourier transform of* $\mathcal{Y}(\xi; l)$ *with respect to* ξ *followed by the inverse Hadamard transform with respect to* l.

5.2.2 Discrete Sensing Model of SP-FTI

SP-FTI proceeds by capturing discrete samples of the semi-continuous data $\mathcal{Y}$ in (5.15) in the OPD domain (see Fig. 5.1) and by processing them in order to reconstruct a discretized version of $\mathcal{X}$.

To fix the ideas, we consider that

(i) similar to the FTI model, the time domain is regularly discretized with N_ξ "major" samples (or time slots), as shown in Fig. 5.1, of duration $\tau_\xi^{\rm maj} = \tau_{\rm hs}/N_\xi$ related to a time slot of each major OPD point, given the total acquisition time $\tau_{\rm hs} > 0$. The moving mirror thus gives access to an OPD axis $(-\xi_{\max}, \xi_{\max}]$ (for some range $\xi_{\max} > 0$) that is evenly discretized over N_ξ samples with an OPD step size $\Delta_\xi > 0$ (*i.e.*, $2\xi_{\max} = N_\xi \Delta_\xi$), and the detector integrates the recorded intensity during the OPD step size Δ_ξ.

(ii) Unlike the FTI model, we assume that each major time slot is further discretized with $N_{\rm p}$ "minor" samples (not shown in Fig. 5.1) of duration $\tau_\xi^{\rm min} = \tau_\xi^{\rm maj}/N_{\rm p}$, related to a time slot of each SLM pattern; resulting in $N_{\rm hs}$ total minor slots.

(iii) Variation of interferograms during a time slot associated with a major OPD point is assumed very small.

(iv) The spatial domain is evenly discretized over $\bar{N}_{\rm p} \times \bar{N}_{\rm p}$ pixels with a pixel length $\Delta_{\rm p} > 0$, *i.e.*, determined by the pixel pitch of the detector, and each pixel of the detector integrates the intensity over a square grid of length $\Delta_{\rm p} \times \Delta_{\rm p}$.

Accordingly, the time domain is regularly discretized with $N_{\rm hs}$ minor samples related to a time slot of $\tau_\xi^{\rm min} = \tau_{\rm hs}/N_{\rm hs}$. These assumptions will allow us to have a comparison between the compressive FTI and SP-FTI systems; see Remark 5.12 and Table. 5.1.

Mathematically, discrete SP-FTI measurements are gathered in a matrix $\boldsymbol{Y} \in \mathbb{R}^{N_\xi \times N_{\rm p}}$ approximating $\mathcal{Y}$ over N_ξ OPD points. Likewise, discrete HS volume is gathered in a cube $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{N_\nu \times \bar{N}_{\rm p} \times \bar{N}_{\rm p}}$ approximating $\mathcal{X}$ over $N_{\rm p}$ pixels and $N_\nu = N_\xi$ wavenumber samples, and thus over $N_{\rm hs}$ voxels. Therefore, according to the Shannon-Nyquist theorem, assuming N_ξ even and positioning the zero-OPD point on the spectral index $l_\xi = N_\xi/2$, the sampling rules of SP-FTI reads

$$\begin{aligned}
\boldsymbol{Y}_{l_\xi,l_\mathrm{h}} &\coloneqq \int \mathrm{rect}(\frac{\xi}{\Delta_\xi} - l_\xi + \frac{N_\xi}{2})\mathcal{Y}(\xi; l_\mathrm{h})\mathrm{d}\xi, \\
\boldsymbol{\mathcal{X}}_{l_\nu,j_1,j_2} &\coloneqq \\
&\iiint \mathrm{rect}(\frac{\nu}{\Delta_\nu} - l_\nu + \frac{N_\nu}{2}, \frac{q_1}{\Delta_\mathrm{p}} - j_1, \frac{q_2}{\Delta_\mathrm{p}} - j_2)\mathcal{X}(\nu; q_1, q_2)\mathrm{d}q_1\mathrm{d}q_2\mathrm{d}\nu,
\end{aligned} \tag{5.16}$$

where $l_\mathrm{h} \in [\![N_\mathrm{p}]\!]$ denotes the index of (Hadamard) SLM pattern, and as in (4.13) $l_\xi \in [\![N_\xi]\!]$, $l_\nu \in [\![N_\nu]\!]$, and $j_1, j_2 \in [\![\bar{N}_\mathrm{p}]\!]$ denote the OPD, wavenumber, and spatial indices, receptively. In (5.16) we recalled the definition of $\boldsymbol{\mathcal{X}}$ from (4.13) for the sake of self-containment.

Recall that $\boldsymbol{X} \in \mathbb{R}^{N_\xi \times N_\mathrm{p}}$ denotes the matrix form of $\boldsymbol{\mathcal{X}}$. Hence, the acquisition process of the Nyquist SP-FTI can be formulated in a matrix form as

$$\boldsymbol{Y} = \boldsymbol{F}^*_{N_\xi}\boldsymbol{X}\boldsymbol{\Phi}^\top_\mathrm{slm} + \boldsymbol{N}^\mathrm{fti} \in \mathbb{R}^{N_\xi \times N_\mathrm{p}}, \tag{5.17}$$

where $\boldsymbol{\Phi}_\mathrm{slm} \in \{0,1\}^{N_\mathrm{p} \times N_\mathrm{p}}$ collects in its columns the SLM patterns $\{h^{(l)}\}_{l=1}^{N_\mathrm{p}}$. In (5.17) we assume that the noise corrupting SP-FTI measurements is the same noise as in the FTI measurements (4.14), and follows the model described in § 4.2.3: it is a zero-mean random iid Gaussian noise with variance σ^2_fti.

However, as aforementioned, in this chapter we limit our model to the SLM patterns that follow the Hadamard patterns. Recall from § 3.2 that the Hadamard matrix $\boldsymbol{\Phi}_\mathrm{2had} \in \mathbb{R}^{N_\mathrm{p} \times N_\mathrm{p}}$ contains $\pm N_\mathrm{p}^{-1/2}$ entries. Therefore, to obtain $\boldsymbol{\Phi}_\mathrm{slm}$ with 0 and 1 entries we assume

$$\boldsymbol{\Phi}_\mathrm{slm} = (\sqrt{N_\mathrm{p}}\boldsymbol{\Phi}_\mathrm{2had} + \boldsymbol{1}_{N_\mathrm{p}}\boldsymbol{1}^\top_{N_\mathrm{p}})/2 \in \{0,1\}^{N_\mathrm{p} \times N_\mathrm{p}}, \tag{5.18}$$

which implies

$$\boldsymbol{\Phi}_\mathrm{slm}\boldsymbol{P}^\top_{\{1\}} = \boldsymbol{1}_{N_\mathrm{p}}. \tag{5.19}$$

Combining (5.18) with (5.19) and inserting the result in (5.17) gives

$$2\boldsymbol{Y} = \sqrt{N_\mathrm{p}}\boldsymbol{F}^*\boldsymbol{X}\boldsymbol{\Phi}_{2\mathrm{had}} + \boldsymbol{F}^*\boldsymbol{X}\boldsymbol{\Phi}_{\mathrm{slm}}^\top\boldsymbol{P}_{\{1\}}^\top\boldsymbol{1}_{N_\mathrm{p}}^\top + 2\boldsymbol{N}^{\mathrm{fti}}. \tag{5.20}$$

Moreover, for $\boldsymbol{Y} := [\boldsymbol{y}_1, \cdots, \boldsymbol{y}_{N_\mathrm{p}}] \in \mathbb{R}^{N_\xi \times N_\mathrm{p}}$ and $\boldsymbol{N}^{\mathrm{fti}} := [\boldsymbol{n}_1^{\mathrm{fti}}, \cdots, \boldsymbol{n}_{N_\mathrm{p}}^{\mathrm{fti}}] \in \mathbb{R}^{N_\xi \times N_\mathrm{p}}$, from (5.17) we get

$$\boldsymbol{F}^*\boldsymbol{X}\boldsymbol{\Phi}_{\mathrm{slm}}^\top\boldsymbol{P}_{\{1\}}^\top\boldsymbol{1}_{N_\mathrm{p}}^\top = \boldsymbol{y}_1\boldsymbol{1}_{N_\mathrm{p}}^\top - \boldsymbol{n}_1^{\mathrm{fti}}\boldsymbol{1}_{N_\mathrm{p}}^\top. \tag{5.21}$$

Inserting (5.21) in (5.20) and defining

$$\boldsymbol{Y}^{\mathrm{sp}} := \frac{2\boldsymbol{Y} - \boldsymbol{y}_1\boldsymbol{1}_{N_\mathrm{p}}^\top}{\sqrt{N_\mathrm{p}}}, \text{ and } \boldsymbol{N}^{\mathrm{sp}} := \frac{2\boldsymbol{N}^{\mathrm{fti}} - \boldsymbol{n}_1^{\mathrm{fti}}\boldsymbol{1}_{N_\mathrm{p}}^\top}{\sqrt{N_\mathrm{p}}}, \tag{5.22}$$

the discrete acquisition model of SP-FTI with Hadamard patterns reads

$$\boldsymbol{Y}^{\mathrm{sp}} = \boldsymbol{F}_{N_\xi}^*\boldsymbol{X}\boldsymbol{\Phi}_{2\mathrm{had}} + \boldsymbol{N}^{\mathrm{sp}} \in \mathbb{R}^{N_\xi \times N_\mathrm{p}}, \tag{5.23}$$

or, equivalently,

$$\boldsymbol{y}^{\mathrm{sp}} = \boldsymbol{\Phi}_{\mathrm{sp}}^*\boldsymbol{x} + \boldsymbol{n}^{\mathrm{sp}} \in \mathbb{R}^{N_{\mathrm{hs}}}, \tag{5.24}$$

where $\boldsymbol{\Phi}_{\mathrm{sp}} := \boldsymbol{\Phi}_{2\mathrm{had}} \otimes \boldsymbol{F}_{N_\xi} \in \mathbb{R}^{N_\xi \times N_\mathrm{p}}$, $\boldsymbol{y}^{\mathrm{sp}} := \mathrm{vec}(\boldsymbol{Y}^{\mathrm{sp}}) \in \mathbb{R}^{N_{\mathrm{hs}}}$, $\boldsymbol{x} := \mathrm{vec}(\boldsymbol{X}) \in \mathbb{R}^{N_{\mathrm{hs}}}$, and $\boldsymbol{n}^{\mathrm{sp}} := \mathrm{vec}(\boldsymbol{N}^{\mathrm{sp}}) \in \mathbb{R}^{N_{\mathrm{hs}}}$.

Corollary 5.5. *In view of* (5.22)*, let us consider the Gaussian random matrix* $\boldsymbol{N}^{\mathrm{fti}} \in \mathbb{R}^{N_\xi \times N_\mathrm{p}}$ *with* $n_{l_\xi, l_\mathrm{h}}^{\mathrm{fti}} \sim_{\mathrm{iid}} \mathcal{N}(0, \sigma_{\mathrm{fti}}^2)$*,* $l_\xi \in [\![N_\xi]\!]$ *and* $l_\mathrm{p} \in [\![N_\mathrm{p}]\!]$*. For* $l_\xi \in [\![N_\xi]\!]$*, we have*

$$\begin{cases} n_{l_\xi, l_\mathrm{h}}^{\mathrm{sp}} \sim_{\mathrm{iid}} \mathcal{N}(0, N_\mathrm{p}^{-1}\sigma_{\mathrm{fti}}^2), & \text{if } l_\mathrm{h} = 1, \\ n_{l_\xi, l_\mathrm{h}}^{\mathrm{sp}} \sim_{\mathrm{iid}} \mathcal{N}(0, 5N_\mathrm{p}^{-1}\sigma_{\mathrm{fti}}^2), & \text{if } l_\mathrm{h} \in [\![N_\mathrm{p}]\!]\backslash\{1\}. \end{cases}$$

Moreover, for $l_\xi \in [\![N_\xi]\!]$*, the random variable* $n_{l_\xi, 1}^{\mathrm{sp}}$ *is not independent from* $N_\mathrm{p} - 1$ *random variables* $n_{l_\xi, l_\mathrm{h}}^{\mathrm{sp}}$ *with* $l_\mathrm{h} \in [\![N_\mathrm{p}]\!]\backslash\{1\}$*.*

Proof. See § 5.5.2. □

Remark 5.6. *In view of* (5.23)*, we note that:*

(i) similar conversion argument to (5.18) *can be found in [99, 106, 107].*

(ii) In order to obtain post-processed measurements (5.23) *from SP-FTI measurements* (5.17)*, we implicitly assumed in* (5.20) *that the set of measurements associated with the first row of the Hadamard matrix is already included in the recorded data. This assumption holds for the Nyquist SP-FTI; while for the compressive SP-FTI explained in § 5.3, due to the random subsampling of the Hadamard matrix, it does not necessarily hold. To mitigate this issue, in the rest of this thesis, we assume that the first row of the Hadamard matrix is always sampled, no matter which subsampling strategy is performed.*

5.3 Compressive SP-FTI

With the Nyquist SP-FTI model introduced in the previous section, we here propose a compressive SP-FTI model.

5.3.1 Acquisition Strategy

Recall that in Nyquist SP-FTI, $N_{\rm p}$ Hadamard patterns are programmed at each major OPD point; resulting in $N_{\rm hs}$ total Hadamard patterns. Compressive SP-FTI, on the other hand, allows for the selection of $M_{l_\xi} \leq N_{\rm p}$ different Hadamard patterns at each $l_\xi^{\rm th}$ major OPD point with $l_\xi \in [\![N_\xi]\!]$; resulting in $M := \sum_{l_\xi} M_\xi \leq N_{\rm hs}$ total Hadamard patterns. This is also different from SI-FTI where at each OPD point different *spatial locations* are selected.

To understand the flexibility of compressive SP-FTI with Hadamard patterns, observe that the Nyquist FTI scheme can be obtained from such a system. It can be simply done by programming all the possible Hadamard patterns on the SLM for all major OPD samples followed by an application of the inverse Hadamard transform to the rows of the SP-FTI data matrix $\boldsymbol{Y}^{\rm sp}$, *i.e.*, $\boldsymbol{Y}^{\rm fti} = \boldsymbol{Y}^{\rm sp}\boldsymbol{\Phi}_{\rm 2had} + \boldsymbol{N}$, where $\boldsymbol{N}$ is still a zero-mean random iid Gaussian noise; it is a consequence of the orthonormality of the Hadamard matrix.

Similar to the FTI model, we assume, as defined in (4.4), that the light source delivers a constant intensity $I^{\rm src}$ per second and per unit area on any location of the specimen. Therefore, in the Nyquist SP-FTI, the total light exposure at each major OPD point on the whole biological specimen is constant and equal[2] to $2^{-1}(N_{\rm p}+1)N_{\rm p}\delta_S\tau_\xi^{\rm min}I^{\rm src}$, where $\tau_\xi^{\rm min}$ corresponds to the duration of each minor time (or OPD) sample, and δ_S is the SLM pixel area. Therefore, the total light exposure on the whole biological specimen during the whole Nyquist SP-FTI acquisition is equal to $2^{-1}(N_{\rm p}+1)N_{\rm hs}\delta_S\tau_\xi^{\rm min}I^{\rm src}$.

By modifying the programming of the SLM in compressive SP-FTI, if only M_{l_ξ} Hadamard patterns (out of $N_{\rm p}$ possible patterns) are programmed on the SLM at the $l_\xi^{\rm th}$ major OPD sample[3] with $l_\xi \in [\![N_\xi]\!]$, the total light exposure at each major OPD point (during the whole compressive SP-FTI acquisition) on the whole biological specimen is decreased to $2^{-1}(N_{\rm p}+1)M_{l_\xi}\delta_S\tau_\xi^{\rm min}I^{\rm src}$ (respectively, $2^{-1}(N_{\rm p}+1)M\delta_S\tau_\xi^{\rm min}I^{\rm src}$) by a factor of $M_{l_\xi}/N_{\rm p}$ (respectively, $M/N_{\rm hs}$).

We now turn our attention to compressive SP-FTI sensing model. Let us follow the general random VDS scheme of § 2.3, with special care to integrate the geometry of SP-FTI. We consider the set $\Omega = \{\omega_1, \cdots, \omega_M\} \subset [\![N_{\rm hs}]\!]$ generated from M (scalar) r.v.s $\omega_l \sim_{\rm iid} \beta$ (with $l \in [\![M]\!]$), where the r.v. $\beta \in N_{\rm hs}$ is defined from the pmf

$$p(l_{\rm hs}) := \mathbb{P}[\beta = l_{\rm hs}], \quad \forall l_{\rm hs} \in [\![N_{\rm hs}]\!]. \tag{5.25}$$

[2]This is a direct consequence of the definition of the Hadamard matrix $\boldsymbol{\Psi}_{\rm 2had} \in \mathbb{R}^{N_{\rm p}\times N_{\rm p}}$, *i.e.*, the first row of $\boldsymbol{\Psi}_{\rm 2had}$ contains $N_{\rm p}$ entries equal to $N_{\rm p}^{-1/2}$ and the rest of the rows contain $N_{\rm p}/2$ entries equal to $N_{\rm p}^{-1/2}$. Combining this with the scaling (5.18) implies the mentioned light exposure.

[3]We assume that the SLM blocks out the light illumination during the other $N_{\rm p} - M_{l_\xi}$ minor time slots.

From (5.24), compressive SP-FTI observation then amounts to

$$\boldsymbol{y}^{\mathrm{csp}} = \boldsymbol{P}_\Omega \boldsymbol{\Phi}_{\mathrm{sp}}^* \boldsymbol{x} + \boldsymbol{n} \in \mathbb{R}^M. \tag{5.26}$$

Remark 5.7. *Similar to the practical challenges of SI-FTI noted in Remark 4.14, the SLM module in compressive SP-FTI needs to be aware of the position of the moving mirror inside the MI device and respect the discretized time domain characterized by the fps rate of the single-pixel sensor. However, since the spatial content of the scene is here integrated into a single-beam, non-identical pixel size of the SLM with single-pixel sensor, as well as their alignment, are not of practical issues in compressive SP-FTI.*

5.3.2 Reconstruction Method and Guarantee

The goal of this section is to provide a stable and robust scheme to recover every HS volume from compressive SP-FTI measurements (5.26) by following the VDS scheme of § 2.3.

Given the noisy compressive SP-FTI measurements as in (5.26), an HS volume $\boldsymbol{x}$ with N_{hs} voxels can be reconstructed using the convex optimization problem

$$\hat{\boldsymbol{x}} = \underset{\boldsymbol{u}\in\mathbb{C}^{N_{\mathrm{hs}}}}{\arg\min} \|\boldsymbol{\Psi}_{\mathrm{sp}}^\top \boldsymbol{u}\|_1 \ \text{ s. t. } \ \|\boldsymbol{D}(\boldsymbol{y}^{\mathrm{csp}} - \boldsymbol{P}_\Omega \boldsymbol{\Phi}_{\mathrm{sp}}^* \boldsymbol{u})\| \le \varepsilon_{\mathrm{sp}}\sqrt{M}, \tag{5.27}$$

where $\Omega = \{\omega_1, \cdots, \omega_M\} \subset [\![N_{\mathrm{hs}}]\!]$ is randomly generated according to the pmf of (5.25), $\boldsymbol{\Psi}_{\mathrm{sp}}$ is the sparsity basis, $\varepsilon_{\mathrm{sp}}$ must satisfy $\|\boldsymbol{D}\boldsymbol{n}\| \le \varepsilon_{\mathrm{sp}}\sqrt{M}$ with high probability, and $\boldsymbol{D} = \mathrm{diag}(\boldsymbol{d}) \in \mathbb{R}^{M\times M}$ with $d_l = 1/(p(\omega_l))^{1/2}$ for $l \in [\![M]\!]$. Notice the similarities between (5.27) and the recovery problem of SI-FTI in (4.27). We also regularize (5.27) similar to (4.27) by considering a joint spatiospectral HS sparsity model described in § 4.3, *i.e.*, $\boldsymbol{\Psi}_{\mathrm{sp}} = \boldsymbol{\Psi}_{\mathrm{si}} = \boldsymbol{\Psi}_{\mathrm{idhw}} \otimes \boldsymbol{\Psi}_{\mathrm{dhw}}$.

We now focus on adjusting the pmf (5.25) determining Ω. Following Thm. 2.4, the preconditioned matrix $\frac{1}{\sqrt{M}}\boldsymbol{D}\boldsymbol{P}_\Omega\boldsymbol{\Phi}_{\mathrm{sp}}^*\boldsymbol{\Psi}_{\mathrm{sp}}$ respects the RIP of order K with probability exceeding $1-\epsilon$ if

$$M \gtrsim \delta^{-2}\|\boldsymbol{\kappa}\|^2 K \log(\epsilon^{-1})$$

SP-FTI measurements are recorded with respect to the pmf $p(l_{\rm hs}) := \kappa_{l_{\rm hs}}^2/\|\boldsymbol{\kappa}\|^2$ for $l_{\rm hs} \in [\![N_{\rm hs}]\!]$, where the vector $\boldsymbol{\kappa} \in \mathbb{R}_+^{N_{\rm hs}}$ is a bound for the local coherence $\mu_{l_{\rm hs}}^{\rm loc}(\boldsymbol{\Phi}_{\rm sp}^*\boldsymbol{\Psi}_{\rm sp})$ with $l_{\rm hs} \in [\![N_{\rm hs}]\!]$.

The next proposition (proved in § 5.5.1) bounds this local coherence, and thus determines the pmf p.

Proposition 5.8. *For the context of SP-FTI explained above,* $\mu_{l_{\rm hs}}^{\rm loc}(\boldsymbol{\Phi}_{\rm sp}^*\boldsymbol{\Psi}_{\rm sp}) \leq \kappa_{l_{\rm hs}}$*, where, for* $l_{\rm hs} \xleftrightarrow{N_\xi,\bar{N}_{\rm p},\bar{N}_{\rm p}} (l_\xi, l_1, l_2)$*,*

$$\kappa_{l_{\rm hs}} := \sqrt{2}\,\min\Big(1, \frac{1}{\sqrt{|l_\xi - N_\xi/2|}}\Big)\cdot \min\Big(1, 2^{-\lfloor \log_2(\max(l_1,l_2)-1)\rfloor}\Big), \tag{5.28}$$

with $\|\boldsymbol{\kappa}\|^2 \lesssim \log(N_\xi)\log_2(\bar{N}_{\rm p})$*. In this case,* (2.8) *reads*

$$p(l_{\rm hs}) = \tfrac{C_{N_\xi}}{3\log_2(\bar{N}_{\rm p})+1}\min\Big(1, |l_\xi - \tfrac{N_\xi}{2}|^{-1}\Big)\cdot\min\big(1, 2^{-\lfloor\log_2(\max(l_1,l_2)-1)\rfloor}\big), \tag{5.29}$$

where the normalization constant C_{N_ξ} *respects* $2\log(N_\xi/2) < C_{N_\xi}^{-1} < 4 + 2\log(N_\xi/2)$*.*

The indices $l_1, l_2 \in [\![\bar{N}_{\rm p}]\!]$ in Prop. (5.8) are here referred to as *spatial Hadamard frequencies*. According to Thm. 2.4, if $\frac{1}{\sqrt{M}}\boldsymbol{D}\boldsymbol{P}_\Omega\boldsymbol{\Phi}_{\rm sp}^*\boldsymbol{\Psi}_{\rm sp}$ respects the RIP with the pmf specified in the previous proposition, the estimation error of (5.27) can be bounded as

$$\|\boldsymbol{x} - \hat{\boldsymbol{x}}\| \leq c_1\frac{\sigma_K(\boldsymbol{\Psi}_{\rm sp}^\top\boldsymbol{x})_1}{\sqrt{K}} + c_2\varepsilon_{\rm sp}. \tag{5.30}$$

Moreover, following Prop. 5.8, Cor. 2.9, Remark 2.8, and Cor. 5.5, the level $\varepsilon_{\rm sp}$ of a Gaussian noise in the SP-FTI model (5.26) can be estimated according to the next proposition.

Proposition 5.9. *In the context of compressive SP-FTI described above, the following upper bound for the weighted noise power in* (5.27), i.e.,

$$\frac{1}{M}\|\boldsymbol{D}\boldsymbol{n}\|^2 \le \varepsilon_{\rm sp}^2(s) := \varepsilon_{\bar\sigma,s}^2(N_{\rm hs}, M, \bar\rho) + \frac{N_\xi}{M}\varepsilon_{\hat\sigma,s}^2(N_\xi, N_\xi, \hat\rho), \quad (5.31)$$

holds with probability exceeding $1 - 6e^{-s/2}$, *where* $\bar\sigma^2 := 5N_{\rm p}^{-1}\sigma_{\rm fti}^2$, $\hat\sigma^2 := 6N_{\rm p}^{-1}\sigma_{\rm fti}^2$, *and* $\varepsilon_{\sigma,s}$ *is defined in* (2.21).

Proof. See § 5.5.3. □

Let us now summarize the complete analysis of compressive SP-FTI in the following theorem.

Theorem 5.10 (Compressive SP-FTI)**.** *Given* $s > 0$, *fix integers* K, $N_{\rm hs} = N_\xi N_{\rm p}$ *such that* $K \gtrsim \log(N_{\rm hs})$ *and*

$$M \gtrsim K \log(N_\xi)\log_2(\bar N_{\rm p})\log(\epsilon^{-1}). \quad (5.32)$$

Generate M *random (non-unique) indices associated with a (1-D) index set* $\Omega = \{\omega_1, \cdots, \omega_M\}$ *such that* $\omega_l \sim_{\rm iid} \beta$ *for* $l \in [\![M]\!]$, *with* β *a r.v. with the pmf* (5.29). *Then, given the noisy SP-FTI measurements* $\boldsymbol{y}^{\rm csp}$ *in* (5.26), *the HS volume* $\boldsymbol{x}$ *can be approximated by solving* (5.27) *with the bound* $\varepsilon_{\rm sp}(s)$ *in Prop. 5.9, up to an error*

$$\|\boldsymbol{x} - \hat{\boldsymbol{x}}\| \le c_1 \frac{\sigma_K(\boldsymbol{\Psi}_{\rm sp}^\top \boldsymbol{x})_1}{\sqrt{K}} + c_2\varepsilon_{\rm sp}(s), \quad (5.33)$$

where the constants c_1 *and* c_2 *are specified in Thm. 2.4, with probability exceeding* $1 - \epsilon - 6e^{-s/2}$.

Proof. A combination of Thm. 2.4 and Prop. 5.8 with Prop. 5.9, (5.30) completes the proof. □

We observe in Thm. 5.10 that the 1-D pmf of selecting the rows of the matrix $\mathbf{\Phi}_{\rm sp}^*$ is linked to the 3-D geometry of the sensing domain. In (5.29), the probability of programming a Hadamard pattern on the SLM (*i.e.*, $p(l_{\rm hs})$ for fixed l_1 and l_2) decreases inversely proportional to the magnitude of the OPD point (its distance from the zero-OPD). In addition, the probability of selecting a Hadamard pattern at a given major OPD point (*i.e.*, $p(l_{\rm hs})$ for fixed l_ξ) is inversely proportional to the maximum magnitude of spatial Hadamard "frequencies". For the sake of comparison, a UDS strategy, *i.e.*, $p(l_{\rm hs}) = N_{\rm hs}^{-1}$ for $l_{\rm hs} \in [\![N_{\rm hs}]\!]$, would result in $M \gtrsim N_{\rm hs} K \log(\epsilon^{-1})$ measurements, which is equivalent to overexposing the observed object.

Remark 5.11. *The authors of [59] proposed a compressive SP-FTI model that reads $\boldsymbol{Y} = \boldsymbol{F}^* \boldsymbol{X} \mathbf{\Phi}_{\rm 2had}^\top \boldsymbol{P}_\Omega^\top + \boldsymbol{N}$, where the Hadamard patterns are subsampled according to the UDS scheme. Moreover, HS volumes are recovered via a TV-norm minimization. Analysis of this model is lacking in [59] and we do not cover it in this thesis. Moreover, during the preparation of this thesis we became aware of another implementation of compressive SP-FTI, called Neospectra [3], which is a single-chip MI implemented using micro electro mechanical systems. However, investigating the applicability of our results on such a device is left to a future study.*

Remark 5.12. *Neglecting delicate issues of photon efficiency and photon noise, we provide two perspectives on the comparison of the total light exposure among CI-FTI, SI-FTI, and compressive SP-FTI. We first recall from § 4.4.1 and § 4.5.1 that the total light exposure on the whole biological specimen during the full acquisition equals $M \tau_\xi \delta_S I^{\rm src}$, with $M = M_\xi N_{\rm p}$ for CI-FTI, where τ_ξ is the duration of time slots associated with the (major) OPD points. Then, depending on the setting of $\tau_\xi^{\rm min}$ and $\tau_\xi^{\rm maj} = \tau_\xi^{\rm min} N_{\rm p}$ (in SP-FTI) with respect to τ_ξ (in FTI), we find the comparison reported in Table 5.1. An interesting observation from this table is that for $\tau_\xi^{\rm maj} = \tau_\xi$,*

Model	Total light exposure		Total measurements
	if $\tau_\xi^{\rm min} = \tau_\xi$	if $\tau_\xi^{\rm maj} = \tau_\xi$	
Nyquist FTI	$N_{\rm hs}\, C_{\rm e}$		$N_{\rm hs}$
CI-FTI (4.22)	$M\, C_{\rm e}$		$M \gtrsim N_{\rm p}\, K_\xi\, C_{\rm m}$
SI-FTI (4.32)	$M\, C_{\rm e}$		$M \gtrsim N_{\rm p}\, K\, C_{\rm m}$
Nyquist SP-FTI	$2^{-1}(N_{\rm p}+1)\, N_{\rm hs}\, C_{\rm e}$	$2^{-1}(N_{\rm p}+1)\, N_\xi\, C_{\rm e}$	$N_{\rm hs}$
Compressive SP-FTI (5.32)	$2^{-1}(N_{\rm p}+1)\, M\, C_{\rm e}$	$2^{-1}(N_{\rm p}^{-1}+1)\, M\, C_{\rm e}$	$M \gtrsim K\, \log_2(\bar{N}_{\rm p})\, C_{\rm m}$

Table 5.1: Comparison of the total light exposure in different HS imaging models covered in this thesis. In this table, $C_{\rm e} := \tau_\xi \delta_S I^{\rm src}$ and $C_{\rm m} := \log(N_\xi)\log(\epsilon^{-1})$. Note that for $\tau_\xi^{\rm maj} = \tau_\xi$, the total acquisition time for all the systems is equal to $N_\xi \tau_\xi$ seconds. Recall also that, according to our design in this thesis, the total acquisition time for compressive FTI systems (and compressive SP-FTI) is equal to the one of Nyquist FTI (respectively, Nyquist SP-FTI).

since $(1 + N_{\rm p}^{-1})2^{-1} \leq 1$, the total light exposure in compressive SP-FTI models is always less than in both CI- and SI-FTI systems.

5.4 Numerical Results

We conduct several simulations to verify the performance of the proposed compressive SP-FTI on a simulated biological HS volume of size $(N_\xi, N_{\rm p}) = (512, 64^2)$ used in the experiments of § 4.8.2; see Fig. 4.6. Compressive SP-FTI observations are formed according to (5.26), where the subset Ω is randomly generated from the pmf (5.29) and the variance of the associated Gaussian noise $\sigma_{\rm sp}^2$ is fixed with respect to the desired SNR values reported in Fig. 5.2. This sensing procedure is repeated over 10 random realizations of the noise and the subset Ω. Fig. 5.2 depicts the average SRE (3.18) in dB as a function of measurement ratio. The value of $\varepsilon_{\rm sp}$ is computed via the empirical $95^{\rm th}$ percentile curve of the weighted noise power $\|\boldsymbol{D}\boldsymbol{n}\|$ over 100 Monte-Carlo realizations of the Gaussian noise and the index set Ω. We recover HS volumes through (5.27) us-

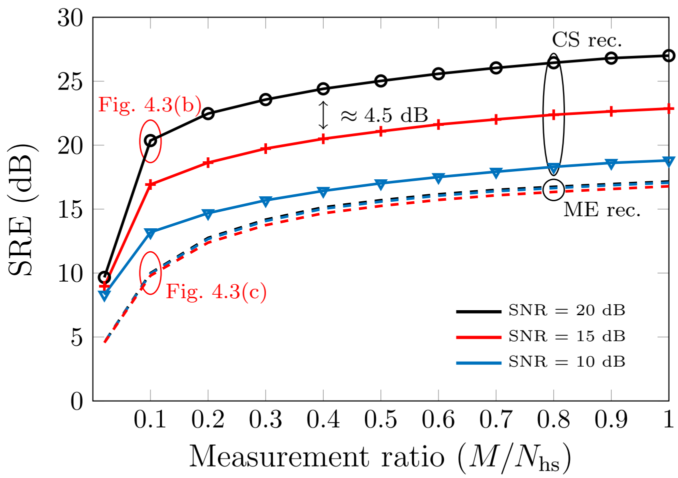

Figure 5.2: Performance of the proposed compressive SP-FTI system.

ing SPGL1 toolbox [118], referred to as CS reconstruction, and the ME problem, *i.e.*, $\hat{\boldsymbol{x}}_{\text{me}} := (\boldsymbol{P}_\Omega \boldsymbol{\Phi}_{\text{sp}}^*)^\dagger \boldsymbol{y}^{\text{csp}}$, where † denotes the pseudo-inverse operator [22].

Remark 5.13. *Since the VDS scheme proposed in Thm. 5.10 allows the selection of repeated indices, the matrix $\boldsymbol{P}_\Omega \boldsymbol{\Phi}_{\text{sp}}^*$ is not necessarily full rank. Therefore, the algebraic formula mentioned in Notations may not hold.*

Following the observations in Fig. 5.2, poor performance of the ME solution (see dashed lines) highlights the necessity of leveraging sparsity prior of the HS data. Since ME reconstruction does not consider the noise power, its performance does not change with respect to different SNR values. On the contrary, CS reconstruction is robust to the noisy measurements and is stable for compressible HS volumes. An example of the reconstructed HS volume for $M/N_{\text{hs}} = 0.1$ is visualized in Fig. 5.3. A comparison between these results and the ground truth values confirms that the spatial and spectral content of the HS volume is successfully preserved, even with 10% of the SP-FTI measurements.

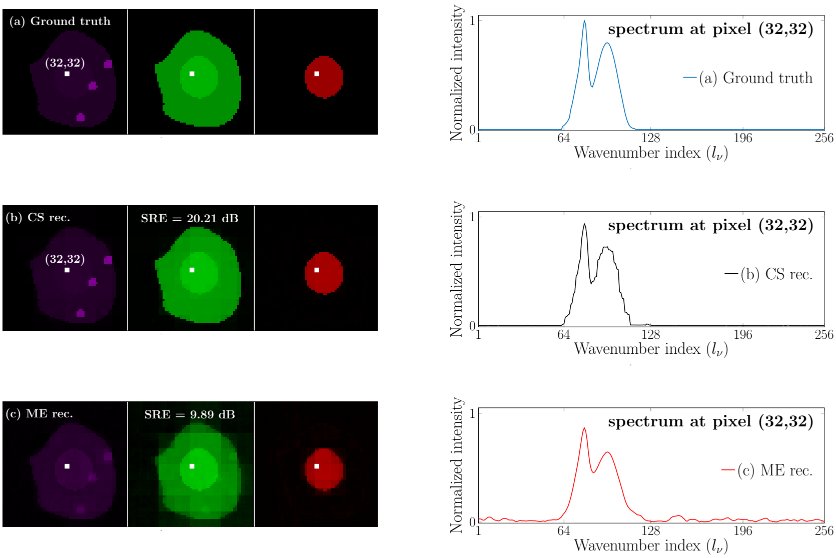

Figure 5.3: An example of reconstructed HS volumes from 10% of the measurements: (left) spatial maps corresponding to $l_\nu \in \{72, 79, 96\}$; (right) the spectral content at the centered pixel.

5.5 Proofs

5.5.1 Proof of Prop. 5.8

We follow similar steps as we did in § 4.9.2 for the proof of Prop. 4.15. To prove the proposition we need to compute a bound on the local coherence $\mu^{\rm loc}_{l_{\rm hs}}(\mathbf{\Phi}^*_{\rm sp}\mathbf{\Psi}_{\rm sp})$ for $l_{\rm hs} \in [\![N_{\rm hs}]\!]$.

We first consider the relation $l_{\rm hs} \xrightleftharpoons{N_\xi, N_{\rm p}} (l_\xi, l_{\rm h})$ between the 1-D "$l_{\rm hs}$" and the 2-D "$(l_\xi, l_{\rm h})$" index representations. From the definition of local coherence (2.6), noting $\mathbf{\Phi}^*_{\rm sp}\mathbf{\Psi}_{\rm sp} = (\mathbf{\Phi}^\top_{\rm 2had}\mathbf{\Psi}_{\rm idhw}) \otimes (\boldsymbol{F}^*\mathbf{\Psi}_{\rm dhw})$, and using (3.20b) we find

$$\mu^{\rm loc}_{l_{\rm hs}}(\mathbf{\Phi}^*_{\rm sp}\mathbf{\Psi}_{\rm sp}) = \underbrace{\max_{l'_\xi}|(\boldsymbol{F}^*\mathbf{\Psi}_{\rm dhw})_{l_\xi, l'_\xi}|}_{=:\ \mu^{\rm loc}_{l_\xi}(\boldsymbol{F}^*\mathbf{\Psi}_{\rm dhw})} \cdot \underbrace{\max_{l'_{\rm p}}|(\mathbf{\Phi}^\top_{\rm 2had}\mathbf{\Psi}_{\rm idhw})_{l_{\rm h}, l'_{\rm p}}|}_{=:\ \mu^{\rm loc}_{l_{\rm h}}(\mathbf{\Phi}^\top_{\rm 2had}\mathbf{\Psi}_{\rm idhw})}. \qquad (5.34)$$

The proof requires now to upper bound the two terms on the right-hand side of (4.55). For the first one, (4.53) provides

$$\begin{cases} \mu^{\mathrm{loc}}_{l_\xi}(\boldsymbol{F}^*\boldsymbol{\Psi}_{\mathrm{dhw}}) \leq \kappa^{\xi}_{l_\xi} := \sqrt{2}\,\min\left\{1, \frac{1}{\sqrt{|l_\xi - N_\xi/2|}}\right\}, \\ \|\boldsymbol{\kappa}^{\xi}\|^2 \leq 8 + 4\log(N_\xi/2) \lesssim \log N_\xi. \end{cases} \tag{5.35}$$

For the second term, Prop. 3.7 provides

$$\begin{cases} \mu^{\mathrm{loc}}_{l_{\mathrm{h}}}(\boldsymbol{\Phi}^{\top}_{\mathrm{2had}}\boldsymbol{\Psi}_{\mathrm{idhw}}) = \kappa^{\mathrm{h}}_{l_{\mathrm{h}}} := \min\left(1, 2^{-\lfloor \log_2(\max(l_1,l_2)-1)\rfloor}\right), \\ \|\boldsymbol{\kappa}^{\mathrm{h}}\|^2 = 3\log_2(\bar{N}_{\mathrm{p}}) + 1, \end{cases} \tag{5.36}$$

where $l_{\mathrm{h}} \xrightleftharpoons{\bar{N}_{\mathrm{p}}} (l_1, l_2)$. Inserting (5.35) and (5.36) in (5.34) results in $\mu^{\mathrm{loc}}_{l_{\mathrm{hs}}}(\boldsymbol{\Phi}^*_{\mathrm{sp}}\boldsymbol{\Psi}_{\mathrm{sp}}) \leq \kappa^{\mathrm{hs}}_{l_{\mathrm{hs}}}$, where

$$\kappa^{\mathrm{hs}}_{l_{\mathrm{hs}}} := \sqrt{2}\,\min\left\{1, \frac{1}{\sqrt{|l_\xi - N_\xi/2|}}\right\} \cdot \min\left(1, 2^{-\lfloor \log_2(\max(l_1,l_2)-1)\rfloor}\right), \tag{5.37}$$

and thus,

$$\|\boldsymbol{\kappa}^{\mathrm{hs}}\|^2 \leq (8 + 4\log(N_\xi/2))(3\log_2(\bar{N}_{\mathrm{p}}) + 1) \lesssim \log(N_\xi)\log_2(\bar{N}_{\mathrm{p}}), \tag{5.38}$$

which provides (5.28).

The pmf associated with SP-FTI can be then formulated from (2.8) and (5.37) as

$$p(l_{\mathrm{hs}}) = \tfrac{C_{N_\xi}}{3\log_2(\bar{N}_{\mathrm{p}})+1}\min\left\{1, |l_\xi - N_\xi/2|^{-1}\right\} \cdot \min\left(1, 2^{-\lfloor \log_2(\max(l_1,l_2)-1)\rfloor}\right),$$

where the normalizing constant C_{N_ξ}, *i.e.*, such that $\sum_{l_{\mathrm{hs}}} p(l_{\mathrm{hs}}) = 1$, is formulated in § 4.9.1.

5.5.2 Proof of Cor. 5.5

To get simpler notation, we write $\boldsymbol{N}, \bar{\boldsymbol{N}}, \sigma, \bar{\sigma}, l, j, M$, and N instead of $\boldsymbol{N}^{\mathrm{fti}}, \boldsymbol{N}^{\mathrm{sp}}, \sigma_{\mathrm{fti}}, \sigma_{\mathrm{sp}}, l_\xi, l_{\mathrm{p}}, N_\xi$, and N_{p}, respectively. From (5.22), we have, for $l \in [\![M]\!]$ and $j \in [\![N]\!]$,

$$\bar{n}_{l,j} = \frac{1}{\sqrt{N}}(2\,n_{l,j} - n_{l,1}) = \begin{cases} \frac{n_{l,1}}{\sqrt{N}}, & \text{if } l \in [\![M]\!], j = 1, \\ \frac{2\,n_{l,j} - n_{l,1}}{\sqrt{N}}, & \text{if } l \in [\![M]\!], j \in [\![N]\!]\backslash\{1\}. \end{cases} \tag{5.39}$$

Since, for $l \in [\![M]\!], j \in [\![N]\!]\backslash\{1\}$, $n_{l,j}$ and $n_{l,1}$ are two independent Gaussian r.v.s, $2\,n_{l,j} - n_{l,1}$ will be a Gaussian r.v. and

$$\mathbb{E}[\bar{n}_{l,j}] = N^{-1/2}(2\mathbb{E}[n_{l,j}] - \mathbb{E}[n_{l,1}]) = 0, \tag{5.40}$$

and

$$\begin{aligned} \mathbb{E}[\bar{n}_{l,j}^2] &= N^{-1}\mathbb{E}[(2\,n_{l,j} - n_{l,1})^2] \\ &= N^{-1}(\mathbb{E}[4n_{l,j}^2] + \mathbb{E}[n_{l,1}^2] - 4\,\mathbb{E}[n_{l,j}n_{l,1}]) \\ &= \frac{5\sigma^2}{N}. \end{aligned} \tag{5.41}$$

Inserting (5.40) and (5.41) in (5.39) gives

$$\begin{cases} \bar{n}_{l,j} \sim_{\text{iid}} \mathcal{N}(0, N^{-1}\sigma^2), & \text{if } l \in [\![M]\!], j = 1, \\ \bar{n}_{l,j} \sim_{\text{iid}} \mathcal{N}(0, 5N^{-1}\sigma^2), & \text{if } l \in [\![M]\!], j \in [\![N]\!]\backslash\{1\}. \end{cases} \tag{5.42}$$

The proof is complete by showing that $E(l,j) := \mathbb{E}[\bar{n}_{l,1}\bar{n}_{l,j}] \neq 0$, for $l \in [\![M]\!], j \in [\![N]\!]\backslash\{1\}$: from (5.39), we get

$$\begin{aligned} E(l,j) &= \frac{1}{N}\mathbb{E}[n_{l,1}(2\,n_{l,j} - n_{l,1})] = \frac{1}{N}\big(2\mathbb{E}[n_{l,1}n_{l,j}] - \mathbb{E}[n_{l,1}^2]\big) \\ &= -\frac{\sigma^2}{N}. \end{aligned}$$

5.5.3 Proof of Prop. 5.9

Recall our assumption in Remark 5.6 that the first Hadamard pattern is always subsampled for all the OPD points. Define a subset of indices in $[\![N_{\text{hs}}]\!]$ that corresponds to the first Hadamard frequency, *i.e.*, $\Omega^{(1)} := \{l \in$

$[\![N_{\rm hs}]\!] : l \xrightleftharpoons{N_\xi, N_{\rm p}} (l_\xi, 1), l_\xi \in [\![N_\xi]\!]\}$, with cardinality $|\Omega^{(1)}| = N_\xi$. In the following, this subset which will be used to isolate the first column of the Hadamard matrix. Define also $\Omega^{(2)} := \Omega \backslash \Omega^{(1)}$ with $|\Omega^{(2)}| = M - N_\xi$. Observe that the noise vector in (5.26) reads $\boldsymbol{n} = [(\boldsymbol{n}^{(1)})^\top, (\boldsymbol{n}^{(2)})^\top]^\top \in \mathbb{R}^M$, where

$$\boldsymbol{n}^{(1)} := \boldsymbol{P}_{\Omega^{(1)}} \boldsymbol{n} \in \mathbb{R}^{N_\xi}, \quad \text{and} \quad \boldsymbol{n}^{(2)} := \boldsymbol{P}_{\Omega^{(2)}} \boldsymbol{n} \in \mathbb{R}^{M - N_\xi}. \tag{5.43}$$

Observe from Cor. 5.5 that

$$\begin{aligned} n_l^{(1)} &\sim_{\rm iid} \mathcal{N}(0, N_{\rm p}^{-1} \sigma_{\rm fti}^2), \quad \text{for} \quad l \in [\![N_\xi]\!] \\ n_l^{(2)} &\sim_{\rm iid} \mathcal{N}(0, 5N_{\rm p}^{-1} \sigma_{\rm fti}^2), \quad \text{for} \quad l \in [\![M - N_\xi]\!]. \end{aligned}$$

We define a Gaussian random vector $\boldsymbol{v} \in \mathbb{R}^{N_\xi}$ with $v_l \sim_{\rm iid} \mathcal{N}(0, 5N_{\rm p}^{-1} \sigma_{\rm fti}^2)$ such that its entries are independent from the entries of both $\boldsymbol{n}^{(1)}$ and $\boldsymbol{n}^{(2)}$. Using this random vector, we rewrite the noise vector as $\boldsymbol{n} = \bar{\boldsymbol{n}} + [\hat{\boldsymbol{n}}^\top, \boldsymbol{0}^\top]^\top$, where

$$\bar{\boldsymbol{n}} := [\boldsymbol{v}^\top, (\boldsymbol{n}^{(2)})^\top]^\top \in \mathbb{R}^M, \quad \text{with} \quad \bar{n}_l \sim_{\rm iid} \mathcal{N}(0, 5N_{\rm p}^{-1} \sigma_{\rm fti}^2), \tag{5.44}$$

$$\hat{\boldsymbol{n}} := \boldsymbol{n}^{(1)} - \boldsymbol{v} \in \mathbb{R}^{N_\xi}, \quad \text{with} \quad \hat{n}_l \sim_{\rm iid} \mathcal{N}(0, 6N_{\rm p}^{-1} \sigma_{\rm fti}^2). \tag{5.45}$$

In the second line above, we used the independence of the elements of $\boldsymbol{v}$ and $\boldsymbol{n}^{(1)}$: $\mathbb{E}[\hat{n}_l^2] = \mathbb{E}[(n_l^{(1)})^2] + \mathbb{E}[v_l^2]$.

Moreover, the diagonal weighting matrix in (5.27) can be partitioned as

$$\boldsymbol{D} = \begin{bmatrix} \boldsymbol{D}^{(1)} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{D}^{(2)} \end{bmatrix},$$

where

$$\begin{aligned} \boldsymbol{D}^{(1)} &:= \boldsymbol{P}_{\Omega^{(1)}} \boldsymbol{D} \boldsymbol{P}_{\Omega^{(1)}}^\top \in \mathbb{R}^{N_\xi \times N_\xi}, \\ \boldsymbol{D}^{(2)} &:= \boldsymbol{P}_{\Omega^{(2)}} \boldsymbol{D} \boldsymbol{P}_{\Omega^{(2)}}^\top \in \mathbb{R}^{(M - N_\xi) \times (M - N_\xi)}. \end{aligned}$$

Recall that, for $\Omega = \{\omega_l\}_{l=1}^{M}$, the entries of the diagonal matrix reads $d_{l,l} = p(\omega_l)^{-1/2}$, for $l \in [\![M]\!]$, with the sampling pmf p proposed in (5.29). Following the construction of the subset $\Omega^{(1)} = \{\omega_l^{(1)}\}_{l=1}^{N_\xi}$ we find that $d_{l,l}^{(1)} = p^{(1)}(\omega_l^{(1)})^{-1/2}$ for $l \in [\![N_\xi]\!]$, where

$$p^{(1)}(l_\xi) := p(l_{\rm hs}|l_1 = l_2 = 1) = \tfrac{C_{N_\xi}}{3\log_2(\bar{N}_{\rm p})+1} \min\Big(1, |l_\xi - \tfrac{N_\xi}{2}|^{-1}\Big). \quad (5.46)$$

With this setup in hand, we now aim for computing the weighted noise bound $\frac{1}{M}\|\boldsymbol{D}\boldsymbol{n}\|^2$ which can be now decomposed into two terms, *i.e.*,

$$\frac{1}{M}\|\boldsymbol{D}\boldsymbol{n}\|^2 = \frac{1}{M}\|\boldsymbol{D}\bar{\boldsymbol{n}}\|^2 + \frac{1}{M}\|\boldsymbol{D}^{(1)}\hat{\boldsymbol{n}}\|^2. \quad (5.47)$$

To compute the first term in (5.43), from Remark 2.8, we need to compute $\bar{\rho}$, *i.e.*, a bound on $N_{\rm hs}^{-1}\|p(l_{\rm hs})^{-1}\|_{\psi_1} \leq \bar{\rho}$. Recalling from Prop. 5.8 that $C_{N_\xi}^{-1} < 2(2 + \log_2(N_\xi))$,

$$\begin{aligned} N_{\rm hs}^{-1}\|p(l_{\rm hs})^{-1}\|_{\psi_1} &\leq N_{\rm hs}^{-1}C_{N_\xi}^{-1}(3\log_2(\bar{N}_{\rm p}) + 1)(N_\xi - N_\xi/2)(2^{\lfloor \log_2(\bar{N}_{\rm p}-1)\rfloor}) \\ &\leq (3\log_2(\bar{N}_{\rm p}) + 1)(2 + \log_2(N_\xi))\bar{N}_{\rm p}^{-1} =: \bar{\rho}. \end{aligned}$$

Therefore, Cor. 2.9 gives

$$\frac{1}{M}\|\boldsymbol{D}\bar{\boldsymbol{n}}\|^2 \leq \bar{\varepsilon}^2(s) := \varepsilon_{\bar{\sigma},s}^2(N_{\rm hs}, M, \bar{\rho}), \quad (5.48)$$

where $\bar{\sigma}^2 := 5N_{\rm p}^{-1}\sigma_{\rm fti}^2$, with probability exceeding $1 - 3e^{-s/2}$ and $\varepsilon_{\sigma,s}$ defined in (2.21).

Similarly, to compute the second term in (5.43), from Remark 2.8, we need to compute $\hat{\rho}$, *i.e.*, a bound on $N_\xi^{-1}\|p^{(1)}(l_\xi)^{-1}\|_{\psi_1} \leq \hat{\rho}$.

$$\begin{aligned} N_\xi^{-1}\|p(l_\xi)^{-1}\|_{\psi_1} &\leq N_\xi^{-1}C_{N_\xi}^{-1}(3\log_2(\bar{N}_{\rm p}) + 1)(N_\xi - N_\xi/2) \\ &\leq (3\log_2(\bar{N}_{\rm p}) + 1)(2 + \log_2(N_\xi))N_{\rm p}^{-1} =: \hat{\rho} = \bar{\rho}\bar{N}_{\rm p}^{-1}. \end{aligned}$$

Therefore, Cor. 2.9 gives

$$\frac{1}{N_\xi}\|\boldsymbol{D}^{(1)}\hat{\boldsymbol{n}}\|^2 \leq \hat{\varepsilon}^2(s) := \varepsilon^2_{\hat{\sigma},s}(N_\xi, N_\xi, \hat{\rho}), \tag{5.49}$$

where $\hat{\sigma}^2 := 6N_{\rm p}^{-1}\sigma^2_{\rm fti}$, with probability exceeding $1 - 3e^{-s/2}$.

Inserting (5.48) and (5.49) in (5.47) and using a union bound gives

$$\frac{1}{M}\|\boldsymbol{D}\boldsymbol{n}\|^2 \leq \bar{\varepsilon}^2(s) + \frac{N_\xi}{M}\hat{\varepsilon}^2(s), \tag{5.50}$$

with probability exceeding $1 - 6e^{-s/2}$ which completes the proof.

5.6 Discussion

We have proposed a compressive SP-FTI where the light illumination can be efficiently structured (using Hadamard patterns) in order to minimize the light exposure imposed on the observed object and the number of measurements. Similar to the compressive FTI systems proposed in Chapter 4, our method is practically plausible without any hardware modification to the initial SP-FTI implemented in [59]. The theoretical recovery guarantee of this chapter supports any application of SP-FTI.

Similar to the constrained-exposure compressive FTI proposed in § 4.10, a scope of future work would be to extend the contribution of this chapter to a biologically friendly SP-FTI, where the total amount of light exposure is assumed fixed, *i.e.*, a constrained-exposure compressive SP-FTI. It is worth mentioning that we have studied in [81] a proof of concept for a compressive SP-FTI, where the structured light illumination is designed based on an MDS strategy. An interesting line of research is to pursue that work in order to develop a provable MDS-based compressive SP-FTI.

Chapter 6

Conclusions and Perspectives

"From small beginnings, big things arise![1]*"*

In this dissertation, as summarized in Fig. 6.1, we have presented different computational interferometry techniques for HS imaging with elaborate analysis of their efficient light coding, signal recovery performance, and feasibility in biomedical imaging applications.

In Chapter 2, we presented elements of subsampling strategies for orthonormal sensing and sparsity bases in the framework of CS. We provided a formal definition of global, local, and multilevel coherence parameters determining the design and efficiency of UDS, VDS, and MDS strategies, respectively. We then presented two types of signal recovery guarantees, *i.e.*, uniform and non-uniform, for VDS and MDS schemes, respectively. For the VDS scheme, we explained how the reconstruction quality depends on the power of a weighted noise; and as our contribution in this chapter, we provided a controllable estimation of that weighted noise power. Essentially, the obtained bound depends on the unweighted ℓ_2- and ℓ_∞-norms of the noise, and on a quantity fixed by the pmf of the VDS.

Some perspectives can be outlined as follows.

[1]"The Law of Non-Contradiction." Fargo: The Third Season, written by Matt Wolpert and Ben Nedivi, directed by John Cameron, 2017.

- The VDS scheme guided in Thm. 2.4 requires us to select the indices of Ω iid with respect to a pmf p. Therefore, the subsampling set Ω contains possible repetition of the same indices. As we discussed in § 4.8.3, this may not be feasible in practical experiments, *e.g.*, in compressive FTI and SP-FTI. A general perspective would be to extend the (iid and with replacement) VDS supported in Thm. 2.4 to a (non-iid and without replacement) VDS scheme where the indices of subsampling set Ω are unique. Alternatively, one can consider a sampling procedure with respect to a Bernoulli selector; see *e.g.*, [100] and [42, § 12.6]. However, given a number of measurements M, performing a Bernoulli selector results in a subsampling set Ω whose cardinality is a random number. For a UDS combined with Bernoulli selector, $\mathbb{E}[|\Omega|] = M$. For a VDS combined with Bernoulli selector, this argument is an open question. Another approach would be, as discussed in § 4.8.3, to construct $\bar{\Omega}$, a replica of Ω containing only unique indices, and possibly to question whether $\mathbb{E}_{\Omega}[\|\boldsymbol{P}_{\bar{\Omega}}\boldsymbol{u}\|^2] = M^{-1}\|\boldsymbol{D}\boldsymbol{P}_{\Omega}\boldsymbol{u}\|^2$ for an arbitrary vector $\boldsymbol{u}$.

- The second term on the right-hand side of the recovery error (2.17) grows asymptotically to $\sqrt{q}$. From the definition of q, if there exists a level t where $m_t = 0$, then $q \to \infty$. In this case, (2.17) reports the failure of the signal recovery, which is contrary to our observations in numerical tests. A scope of research would be to elaborate this gap between the theory and experiments. Following the proof of Prop. 7.1 in [7], we suspect that the presence of $\sqrt{q}$ in (2.17) is an artifact of the developments in Equation 7.6 in [7].

- As addressed in § 4.8, we observed numerically that the estimation (2.19) is not tight enough for small M, as the second term on the right-hand side is inversely proportional to M. As it might be an artifact of the proof in § 2.6.2, a future work would be to improve

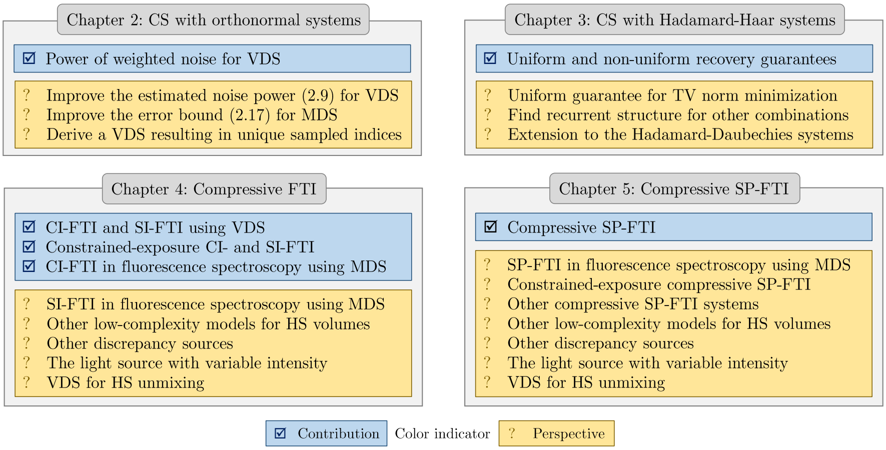

Figure 6.1: Summary of the contributions and potential perspectives of this thesis.

our developments to possibly suppress the term $\max(s/M, \sqrt{s/M})$ in (2.19).

In Chapter 3, we addressed the problem of recovering signals from subsampled Hadamard measurements using the Haar wavelet sparsity basis. We explained how the traditional sampling strategy (UDS) fails in Hadamard-Haar systems, how stable and robust signal recovery provokes VDS and MDS schemes, and how those schemes can be designed based on the local and multilevel coherence values, respectively, between the Hadamard and Haar bases. In particular, we found that the Hadamard-Haar matrix, *i.e.*, obtained by multiplication of the Hadamard and Haar matrices, contains a block structure where each block was associated with the Hadamard matrix of lower order. This enabled us to compute the exact values of local and multilevel coherence and thus achieved tight sample-complexity bounds for both uniform and non-uniform recovery guarantees. In two-dimensions, we covered two constructions of the Haar wavelet basis, *i.e.*, using either tensor product of two 1-D Haar bases or the isotropic construction of a multi-resolution analysis; we observed that either construction is associated with a different efficient sampling strategy.

Some perspectives can be outlined as follows.

- The authors in [61] proposed a VDS of Fourier measurements in two dimensions with two signal prior model: *(i)* signal sparsity in Haar wavelet basis (see [61, Thm. 3.2]), which amounts to the ℓ_1 minimization problem (2.9), and *(ii)* signal sparsity in the gradient domain (see [61, Thm. 3.1]), which is equivalent to replacing the ℓ_1-norm in (2.9) with a TV norm. The former case shares the same sparsity prior as in Thm. 3.8. A future work will thus be to extend the signal recovery program in Thm. 3.8 to a TV minimization, (possibly) by leveraging the proof procedure of Thm. 3.2 in [61].

- Our analysis in this chapter involved the recurrence relations provided by the Kronecker factorization in (3.2) and (3.7) for the Haar and Hadamard matrices, respectively. We acknowledge the application of Kronecker product (and extended operations) for describing a range of other unitary matrices, *e.g.*, the discrete Fourier transform and the related Sine, Cosine, and Hartley transforms [48, 69, 96] and the Daubechies wavelets [39]. An interesting scope of future research would be to investigate the combinations of different sensing and sparsity bases and to find other scaling structures similar to the ones supported in Prop. 3.4 and Prop. 3.6.

- Finally, several lines of research can be extracted following the unchecked cells in Table. 3.1, *e.g.*, deriving uniform and non-uniform recovery guarantees for 2-D Hadamard-Daubechies systems.

In Chapter 4, two versions of compressive FTI, *i.e.*, CI-FTI and SI-FTI, were proposed where the light exposure was coded temporally and spatially with the goal of minimizing the light exposure on the observed object and (at the same time) maximizing the spectral resolution of the resulting HS volumes.

In CI-FTI, illuminating light undergoes a binary modulation before exposing the biological specimen; while SI-FTI involves a spatial light-modulation that allows spatiotemporal light coding. By resorting to a VDS strategy guided by [61], the two proposed frameworks are proved to allow stable and robust recovery of HS volumes. We observed that SI-FTI allows us to further reduce the light exposure, compared to CI-FTI, at the cost of increasing the system complexity. Our analysis was extended to a biologically-friendly scenario where the life-time of biological elements were assumed as a constraint. A striking result in this context was that the proposed compressive FTI systems (where the illumination coding or structuring must be performed as instructed in Thm. 4.12 and Thm. 4.16) yielded higher quality reconstructed HS volumes compared to the Nyquist FTI system.

Next, we presented two improved coding designs for CI-FTI that are adapted to fluorescence spectroscopy experiment. We observed that by using the MDS scheme and extracting the sparse structure shared among different biological dyes and adjusting the sampling strategy accordingly, a superior signal recovery (compared to the VDS scheme) can be achieved. All the methods in this chapter are practically plausible without any modification of the MI.

This chapter has also some interesting perspectives that we decide to merge with those of its brother chapter, *i.e.*, Chapter 5 (see below).

In Chapter 5, as a continuation of the methods presented in Chapter 4, we proposed a variant of compressive FTI combined with a single-pixel detector, referred to as compressive SP-FTI. We have elaborated that the light illumination can be efficiently structured (using Hadamard patterns) with the goal of minimizing the light exposure received by the observed object while preserving the quality of recovered HS volume. Leveraging the results of § 3, we developed a VDS scheme enabling efficient selection of Hadamard patterns. Similar to the compressive FTI systems proposed in Chapter 4, our method is practically plausible without any hardware modification to the initial SP-FTI implemented in [59].

Some perspectives to Chapter 4 and Chapter 5 can be outlined as follows.

- In § 4.2.3 we addressed several discrepancy sources, *e.g.*, Poisson noise, quantization, instrumental response, and explained how their effect could be modeled by an additive Gaussian noise. However, this approach may not hold in practice: *e.g.*, for low photon-counting rates, a Poisson noise cannot be well approximated by a Gaussian noise. A future study in this sense would entail new signal priors and data fidelity models, followed by a new recovery program for HS volumes in CI-FTI, SI-FTI, and compressive SP-FTI systems developed in this thesis.

- The low-complexity prior models for HS volumes goes beyond the two proposed sparsity models in § 4.3. A direction for future study (for CI-FTI, SI-FTI, and compressive SP-FTI systems) would consider other low-complexity models, *e.g.*, low-rankness [38], group sparsity, dictionary-based sparsity, sparsity in shearlet [62] or gradient domains, and generative models.

- Similar to our analysis in § 4.7, two other future works would be to design improved structured illumination schemes for SI-FTI and compressive SP-FTI systems, based on the MDS scheme guided in Thm. 2.6, to boost the quality of recovered HS volumes in fluorescence spectroscopy experiments.

- Similar to the constrained-exposure CI- and SI-FTI systems developed in § 4.10, a line of research would involve designing a constrained-exposure compressive SP-FTI. Recall that since the light illumination is randomly structured in compressive SP-FTI as instructed by Thm. 5.10, the total light exposure received on each spatial location of the observed object varies. In Table 5.1, we reported the total light exposure received by the whole observed object; the main challenge to study constrained-exposure compressive SP-FTI will be to analyze the amount of light exposure received by each spatial location of the object.

- As mentioned in Remark 5.11, the acquisition model of the compressive SP-FTI proposed in [59] is different from the general model introduced in § 5.3. While we believe that our model (5.26) benefits from a greater diversity of the 3-D sampling strategy, it is yet worth analyzing the compressive SP-FTI system of Jin *et al.* [59].

- In the FTI and SP-FTI systems, we assumed that the intensity of the light source is constant in the time and spatial domains (see Assumption 4.2). An interesting alternative would be to consider a light source with variable intensity. In this context, the light source intensity would become a free parameter of compressive FTI and SP-FTI systems, which can be efficiently designed to reduce the effect of noise and in turn boost the quality of the reconstructed HS volumes. To analyze this problem, one may need to verify, for example, whether the Nyquist FTI model (4.15) will extend to $\boldsymbol{y}^{\rm fti} = \boldsymbol{\Lambda}\boldsymbol{\Phi}^*_{\rm fti}\boldsymbol{x} + \boldsymbol{n}^{\rm fti}$ with some diagonal weighting matrix $\boldsymbol{\Lambda}$ collecting on its diagonal the intensity of the light source over the time and spatial domains.

- We have proposed, as a proof of concept, another acquisition scheme for compressive SP-FTI in [81] (not covered in this thesis). Note that in the model explained in Remark 5.11, the same set of subsampled Hadamard patterns are programmed during *every* major OPD points. In [81], on the other hand, we introduced a compressive SP-FTI where the same set of subsampled Hadamard patterns are programmed during *a subset of* the major OPD points.

This also differs from the model (5.26) where a different subset of Hadamard patterns are programmed during every major OPD point. Therefore, profound analysis of the work in [81] would be also an interesting line of research.

- Finally, in this thesis we have addressed different HS imaging problems, while HS interpretation and unmixing problems are left out the present thesis. A line of research would be to unmix the HS volumes recovered from CI-FTI, SI-FTI, and the proposed compressive SP-FTI measurements. A general and challenging future work would be, however, to unmix HS volumes directly from the (subsampled) CI-, SI-, and SP-FTI measurements. The main challenge in that study would be to question how to integrate the concepts of VDS and MDS in the realm of HS unmixing and what sample-complexity bound will assert stable and robust HS unmixing. One can call this idea *variable density sampling for HS unmixing*.

References

[1] "Fluorophores.org - Database of fluorescent dyes, properties and applications," http://www.fluorophores.tugraz.at, retrieved on Aug. 5th, 2019.

[2] "LUXEON A DS100 specification sheet."

[3] "Neospectra.com - Neospectra: a Single-Pixel Fourier Transform Interferometry using Micro Electro Mechanical Systems," https://www.neospectra.com/how-it-works/mems-ft-ir/, retrieved on Oct. 15th, 2019.

[4] "Spot –- A Linear-Operator Toolbox," http://www.cs.ubc.ca/labs/scl/spot/index.html, retrieved on Aug. 25th, 2018.

[5] B. Adcock, V. Antun, and A. C. Hansen, "Uniform recovery in infinite-dimensional compressed sensing and applications to structured binary sampling," *arXiv preprint arXiv:1905.00126*, 2019.

[6] B. Adcock, C. Boyer, and S. Brugiapaglia, "On oracle-type local recovery guarantees in compressed sensing," *arXiv preprint arXiv:1806.03789*, 2018.

[7] B. Adcock, A. C. Hansen, C. Poon, and B. Roman, "Breaking the coherence barrier: A new theory for compressed sensing," in *Forum of Mathematics, Sigma*, vol. 5. Cambridge University Press, 2017.

[8] B. Adcock, A. C. Hansen, and B. Roman, "The quest for optimal sampling: Computationally efficient, structure-exploiting measurements for compressed sensing," in *Compressed Sensing and its Applications*. Springer, 2015, pp. 143–167.

[9] ——, "A note on compressed sensing of structured sparse wavelet coefficients from subsampled Fourier measurements," *IEEE signal processing letters*, vol. 23, no. 5, pp. 732–736, 2016.

[10] V. Antun, "Coherence estimates between Hadamard matrices and Daubechies wavelets," Master's thesis, 2016.

[11] H. Arguello and G. R. Arce, "Code aperture optimization for spectrally agile compressive imaging," *JOSA A*, vol. 28, no. 11, pp. 2400–2413, 2011.

[12] L. Baldassarre, C. Aprile, M. Shoaran, Y. Leblebici, and V. Cevher, "Structured sampling and recovery of IEEG signals," in *IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing*, no. EPFL-CONF-214740, 2015.

[13] R. Bell, *Introductory Fourier transform spectroscopy*. Elsevier, 2012.

[14] G. Beylkin, R. Coifman, and V. Rokhlin, "Fast wavelet transforms and numerical algorithms i," *Communications on pure and applied mathematics*, vol. 44, no. 2, pp. 141–183, 1991.

[15] J. Bigot, C. Boyer, and P. Weiss, "An analysis of block sampling strategies in compressed sensing," *IEEE transactions on information theory*, vol. 62, no. 4, pp. 2125–2139, 2016.

[16] T. Blumensath and M. E. Davies, "Iterative hard thresholding for compressed sensing," *Applied and computational harmonic analysis*, vol. 27, no. 3, pp. 265–274, 2009.

[17] C. Boyer, J. Bigot, and P. Weiss, "Compressed sensing with structured sparsity and structured acquisition," *arXiv preprint arXiv:1505.01619*, 2015.

[18] D. J. Brady, *Optical imaging and spectroscopy*. John Wiley & Sons, 2009.

[19] T. T. Cai and A. Zhang, "Sparse representation of a polytope and recovery of sparse signals and low-rank matrices." *IEEE transactions on information theory*, vol. 60, no. 1, pp. 122–132, 2014.

[20] R. Calderbank, A. Hansen, B. Roman, and L. Thesing, "On reconstructions from measurements with binary functions," *Springer, to appear*.

[21] E. Candès and J. Romberg, "Sparsity and incoherence in compressive sampling," *Inverse problems*, vol. 23, no. 3, pp. 969–985, 2007.

[22] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE transactions on information theory*, vol. 52, no. 2, pp. 489–509, 2006.

[23] E. J. Candès and T. Tao, "Decoding by linear programming," *IEEE transactions on information theory*, vol. 51, no. 12, pp. 4203–4215, 2005.

[24] ——, "Near-optimal signal recovery from random projections: Universal encoding strategies?" *IEEE transactions on information theory*, vol. 52, no. 12, pp. 5406–5425, 2006.

[25] W. L. Chan, K. Charan, D. Takhar, K. F. Kelly, R. G. Baraniuk, and D. M. Mittleman, "A single-pixel terahertz imaging system based on compressed sensing," *Applied Physics Letters*, vol. 93, no. 12, p. 121105, 2008.

[26] H. Choi, D. Wadduwage, P. T. Matsudaira, and P. T. So, "Depth resolved hyperspectral imaging spectrometer based on structured light illumination and Fourier transform interferometry," *Biomedical optics express*, vol. 5, no. 10, pp. 3494–3507, 2014.

[27] P. Clemente, V. Durán, E. Tajahuerce, P. Andrés, V. Climent, and J. Lancis, "Compressive holography with a single-pixel detector," *Optics letters*, vol. 38, no. 14, pp. 2524–2527, 2013.

[28] G. Davis, M. Maggioni, F. Warner, and F. Geshwind, "Hyperspectral analysis of normal and malignant colon tissue microarray sections using a novel DMD system," in *NIH Optical Imaging Workshop*, 2004.

[29] R. DeVerse, F. Geshwind, R. Coifman, W. Fateley, and A. Coppi, "Application of spatial light modulators for new modalities in spectrometry and imaging," Nov. 4 2004, US Patent App. 10/764,113.

[30] R. A. DeVore, B. Jawerth, and B. J. Lucier, "Image compression through wavelet transform coding," *IEEE Transactions on information theory*, vol. 38, no. 2, pp. 719–746, 1992.

[31] A. Diaspro, G. Chirico, C. Usai, P. Ramoino, and J. Dobrucki, "Photobleaching," in *Handbook of biological confocal microscopy*. Springer, 2006, pp. 690–702.

[32] D. L. Donoho, "Compressed sensing," *IEEE transactions on information theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[33] D. L. Donoho and I. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *bBiometrika*, pp. 425–455, 1994.

[34] M. F. Duarte, M. A. Davenport, D. Takhar, J. N. Laska, T. Sun, K. E. Kelly, and R. G. Baraniuk, "Single-pixel imaging via compressive sampling," *IEEE signal processing magazine*, vol. 25, no. 2, p. 83, 2008.

[35] M. Elad, P. Milanfar, and R. Rubinstein, "Analysis versus synthesis in signal priors," *Inverse problems*, vol. 23, no. 3, p. 947, 2007.

[36] B. Falkowski and S. Rahardja, "Walsh-like functions and their relations," *IEE Proceedings-Vision, Image and Signal Processing*, vol. 143, no. 5, pp. 279–284, 1996.

[37] V. Farley, M. Chamberland, P. Lagueux, A. Vallières, A. Villemaire, and J. Giroux, "Chemical agent detection and identification with a hyperspectral imaging infrared sensor," in *Proceedings SPIE*, vol. 6661, 2007, p. 66610L.

[38] M. Fazel, E. Candes, B. Recht, and P. Parrilo, "Compressed sensing and robust recovery of low rank matrices," in *42nd Asilomar Conference on Signals, Systems and Computers*, 2008, pp. 1043–1047.

[39] A. Fijany and C. P. Williams, "Quantum wavelet transforms: Fast algorithms and complete circuits," in *NASA International Conference on Quantum Computing and Quantum Communications*. Springer, 1998, pp. 10–33.

[40] B. Fino, "Relations between Haar and Walsh/Hadamard transforms," *Proceedings of the IEEE*, vol. 60, no. 5, pp. 647–648, 1972.

[41] S. Foucart, "Flavors of compressive sensing," in *International Conference Approximation Theory*. Springer, 2016, pp. 61–104.

[42] S. Foucart and H. Rauhut, *A mathematical introduction to compressive sensing*. Birkhäuser Basel, 2013, vol. 1, no. 3.

[43] M. Gehm, R. John, D. Brady, R. Willett, and T. Schulz, "Single-shot compressive spectral imaging with a dual-disperser architecture," *Optics express*, vol. 15, no. 21, pp. 14 013–14 027, 2007.

[44] M. Golbabaee, S. Arberet, and P. Vandergheynst, "Compressive source separation: Theory and methods for hyperspectral imaging," *IEEE transactions on image processing*, vol. 22, no. 12, pp. 5096–5110, 2013.

[45] M. Golbabaee and P. Vandergheynst, "Hyperspectral image compressed sensing via low-rank and joint-sparse matrix recovery," in *IEEE International Conference On Acoustics, Speech and Signal Processing (ICASSP)*. Ieee, 2012, pp. 2741–2744.

[46] ——, "Joint trace/TV norm minimization: A new efficient approach for spectral compressive imaging," in *IEEE International Conference on Image Processing (ICIP)*, 2012, pp. 933–936.

[47] A. Gowen, C. O'Donnell, P. Cullen, G. Downey, and J. Frias, "Hyperspectral imaging -– an emerging process analytical tool for food quality and safety control," *Trends in Food Science & Technology*, vol. 18, no. 12, pp. 590 – 598, 2007.

[48] J. Granata, M. Conner, and R. Tolimieri, "Recursive fast algorithm and the role of the tensor product," *IEEE transactions on signal processing*, vol. 40, no. 12, pp. 2921–2930, 1992.

[49] R. M. Gray and D. L. Neuhoff, "Quantization," *IEEE transactions on information theory*, vol. 44, no. 6, pp. 2325–2383, 1998.

[50] S. Guérit, S. Sivankutty, C. Scotté, J. A. Lee, H. Rigneault, and L. Jacques, "Compressive sampling approach for image acquisition with lensless endoscope," *arXiv preprint arXiv:1810.12286*, 2018.

[51] J. Hadamard, "Resolution d'une question relative aux determinants," *Bulletin des Sciences Mathématiques*, vol. 2, pp. 240–246, 1893.

[52] A. Hansen and L. Thesing, "On the stable sampling rate for binary measurements and wavelet reconstruction," *Applied and Computational Harmonic Analysis*, 2018.

[53] A. Hansen and L. Terhaar, "Sampling from binary measurements-on reconstructions from walsh coefficients," in *International Conference on Sampling Theory and Applications (SampTA)*. IEEE, 2017, pp. 256–260.

[54] S. S. Haykin and B. V. Veen, *Signals and Systems*, 1st ed. New York, NY, USA: John Wiley & Sons, Inc., 1998.

[55] K. J. Horadam, *Hadamard matrices and their applications*. Princeton university press, 2012.

[56] G. Huang, H. Jiang, K. Matthews, and P. Wilford, "Lensless imaging by compressive sensing," in *Image Processing (ICIP), 2013 20th IEEE International Conference on*. IEEE, 2013, pp. 2101–2105.

[57] L. Jacques, L. Duval, C. Chaux, and G. Peyré, "A panorama on multiscale geometric representations, intertwining spatial, directional and frequency selectivity," *Signal Processing*, vol. 91, no. 12, pp. 2699–2730, 2011.

[58] L. Jacques and A. Moshtaghpour, "Structured illumination and variable density sampling for compressive Fourier transform interferometry," p. 54, 2019.

[59] S. Jin, W. Hui, Y. Wang, K. Huang, Q. Shi, C. Ying, D. Liu, Q. Ye, W. Zhou, and J. Tian, "Hyperspectral imaging using the single-pixel Fourier transform technique," *Scientific reports*, vol. 7, p. 45209, 2017.

[60] M. N. Kolounzakis and M. Matolcsi, "Complex Hadamard matrices and the spectral set conjecture," *Collectanea mathematica*, vol. 57, no. 1, pp. 281–291, 2006.

[61] F. Krahmer and R. Ward, "Stable and robust sampling strategies for compressive imaging," *IEEE transactions on image processing*, vol. 23, no. 2, pp. 612–622, 2014.

[62] D. Labate, W.-Q. Lim, G. Kutyniok, and G. Weiss, "Sparse multidimensional representation using shearlets," in *Wavelets XI*, vol. 5914, 2005.

[63] F. Lacar, M. Lewis, and I. Grierson, "Use of hyperspectral imagery for mapping grape varieties in the Barossa Valley, South Australia," in *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, vol. 6, 2001, pp. 2875–2877.

[64] B. Laurent and P. Massart, "Adaptive estimation of a quadratic functional by model selection," *Annals of Statistics*, pp. 1302–1338, 2000.

[65] C. Leonard, A. Errachid, J. Daubie, D. Beghuin, P.-J. Courtois, M.-P. Mingeot-Leclercq, and D. Tyteca, "Hyperspectral analysis of Laurdan emission spectra in red blood cells and giant unilamellar vesicles," *Biophysical Journal*, vol. 108, no. 2, p. 622a, 2015.

[66] C. Li and B. Adcock, "Compressed sensing with local structure: uniform recovery guarantees for the sparsity in levels class," *Applied and Computational Harmonic Analysis*, 2017.

[67] C. Li, T. Sun, K. F. Kelly, and Y. Zhang, "A compressive sensing and unmixing scheme for hyperspectral data processing," *IEEE transactions on image processing*, vol. 21, no. 3, pp. 1200–1210, 2012.

[68] V. Ljosa, K. L. Sokolnicki, and A. E. Carpenter, "Annotated high-throughput microscopy image sets for validation," *Nature methods*, vol. 9, no. 7, p. 637, 2012.

[69] C. F. Loan, "The ubiquitous Kronecker product," *Journal of Computational and Applied Mathematics*, vol. 123, no. 1, pp. 85 – 100, 2000.

[70] G. Lu and B. Fei, "Medical hyperspectral imaging: a review," *Journal of biomedical optics*, vol. 19, no. 1, pp. 010 901–010 901, 2014.

[71] M. Lustig, D. Donoho, and J. M. Pauly, "Sparse MRI: The application of compressed sensing for rapid MR imaging," *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, vol. 58, no. 6, pp. 1182–1195, 2007.

[72] M. Lustig, D. L. Donoho, J. M. Santos, and J. M. Pauly, "Compressed sensing MRI," *IEEE signal processing magazine*, vol. 25, no. 2, pp. 72–82, 2008.

[73] J. Ma, "Single-pixel remote sensing," *IEEE Geoscience and Remote Sensing Letters*, vol. 6, no. 2, pp. 199–203, 2009.

[74] S. Mallat, *A wavelet tour of signal processing: the sparse way*. Academic press, 2008.

[75] G. Martín, J. M. Bioucas-Dias, and A. Plaza, "HYCA: A new technique for hyperspectral compressive sensing," *IEEE transactions on geoscience and remote sensing*, vol. 53, no. 5, pp. 2819–2831, 2015.

[76] L. Martínez-León, P. Clemente, Y. Mori, V. Climent, J. Lancis, and E. Tajahuerce, "Single-pixel digital holography with phase-encoded illumination," *Optics express*, vol. 25, no. 5, pp. 4975–4984, 2017.

[77] A. A. Michelson and E. W. Morley, "On the relative motion of the earth and of the luminiferous ether," *Sidereal Messenger*, vol. 6, pp. 306–310, 1887.

[78] A. Moshtaghpour, V. Cambareri, L. Jacques, P. Antoine, and M. Roblin, "Compressive hyperspectral imaging using coded Fourier transform interferometry," in *Signal Processing with Adaptive Sparse Structured Representations workshop (SPARS)*, 2017.

[79] A. Moshtaghpour, K. Degraux, V. Cambareri, A. Gonzalez, M. Roblin, L. Jacques, and P. Antoine, “Compressive hyperspectral imaging with Fourier transform interferometry,” in *3rd International Traveling Workshop on Interactions between Sparse models and Technology*, 2016, pp. 27–29.

[80] A. Moshtaghpour, L. Jacques, V. Cambareri, P. Antoine, and M. Roblin, “A variable density sampling scheme for compressive Fourier transform interferometry,” *SIAM journal on imaging sciences*, vol. 12, no. 2, pp. 671–715, 2019.

[81] A. Moshtaghpour, J. M. Bioucas-Dias, and L. Jacques, “Compressive hyperspectral imaging: Fourier transform interferometry meets single pixel camera,” in *international Traveling Workshop on Interactions between low-complexity data models and Sensing Techniques (iTWIST)*, 2018.

[82] ——, “Compressive single-pixel Fourier transform imaging using structured illumination,” in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2019, pp. 7810–7814.

[83] A. Moshtaghpour, V. Cambareri, K. Degraux, A. C. Gonzalez Gonzalez, M. Roblin, L. Jacques, and P. Antoine, “Coded-illumination Fourier transform interferometry,” in *the Golden Jubilee Meeting of the Royal Belgian Society for Microscopy (RBSM)*, 2016, pp. 65–66.

[84] A. Moshtaghpour, J. B. Dias, and L. Jacques, “Close encounters of the binary kind: Signal reconstruction guarantees for compressive Hadamard sampling with Haar wavelet basis,” *arXiv preprint arXiv:1907.09795*, 2019.

[85] A. Moshtaghpour, J. M. B. Dias, and L. Jacques, “Single pixel hyperspectral imaging using Fourier transform interferometry,” in *the International Biomedical and Astronomical Signal Processing (BASP) Frontiers workshop*, 2019, p. 78.

[86] A. Moshtaghpour and L. Jacques, “Multilevel illumination coding for Fourier transform interferometry in fluorescence spectroscopy,” in *25th IEEE International Conference on Image Processing (ICIP)*, 2018, pp. 1433–1437.

[87] A. Moshtaghpour, L. Jacques, V. Cambareri, K. Degraux, and C. De Vleeschouwer, “Consistent basis pursuit for signal and matrix estimates in quantized compressed sensing,” *IEEE signal processing letters*, vol. 23, no. 1, pp. 25–29, 2016.

[88] H. Nagahara, C. Zhou, T. Watanabe, H. Ishiguro, and S. K. Nayar, "Programmable aperture camera using lcos," in *European Conference on Computer Vision*. Springer, 2010, pp. 337–350.

[89] M. H. Neumann, R. Von Sachs *et al.*, "Wavelet thresholding in anisotropic function classes and application to adaptive estimation of evolutionary spectra," *The Annals of Statistics*, vol. 25, no. 1, pp. 38–76, 1997.

[90] R. D. Nowak and R. G. Baraniuk, "Wavelet-based transformations for nonlinear signal processing," *IEEE Transactions on signal processing*, vol. 47, no. 7, pp. 1852–1865, 1999.

[91] A. C. Polak, M. F. Duarte, and D. L. Goeckel, "Performance bounds for grouped incoherent measurements in compressive sensing," *IEEE Transactions on Signal Processing*, vol. 63, no. 11, pp. 2877–2887, 2015.

[92] S. Popa, "Classification of subfactors: the reduction to commuting squares," *Inventiones mathematicae*, vol. 101, no. 1, pp. 19–43, 1990.

[93] G. Puy, P. Vandergheynst, and Y. Wiaux, "On variable density compressive sampling," *IEEE signal processing letters*, vol. 18, no. 10, pp. 595–598, 2011.

[94] H. M. Rafiq and M. U. Siddiqi, "The Haar-recursive transform and its consequence to the Walsh-paley spectrum and the autocorrelation function," *International Journal of Engineering Research and Application (IJERA)*, vol. 6, no. 11, pp. 46–58, 2016.

[95] H. Rauhut, "Compressive sensing and structured random matrices," *Theoretical foundations and numerical methods for sparse recovery*, vol. 9, pp. 1–92, 2010.

[96] P. A. Regalia and M. K. Sanjit, "Kronecker products, unitary matrices and signal processing applications," *SIAM review*, vol. 31, no. 4, pp. 586–613, 1989.

[97] P. Rigollet, "High-dimensional statistics," *Lecture notes for course 18S997*, 2015.

[98] R. T. Rockafellar, *Convex analysis*. Princeton university press, 2015.

[99] B. Roman, A. C. Hansen, and B. Adcock, "On asymptotic structure in compressed sensing," *arXiv preprint arXiv:1406.4178*, 2014.

[100] M. Rudelson and R. Vershynin, "On sparse reconstruction from Fourier and Gaussian measurements," *Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences*, vol. 61, no. 8, pp. 1025–1045, 2008.

[101] P. Ruusuvuori, A. Lehmussola, J. Selinummi, T. Rajala, H. Huttunen, and O. Yli-Harja, "Benchmark set of synthetic images for validating cell image analysis algorithms," in *16th European signal processing conference*, 2008, pp. 1–5.

[102] R. G. Sellar and G. D. Boreman, "Comparison of relative signal-to-noise ratios of different classes of imaging spectrometer," *Applied optics*, vol. 44, no. 9, pp. 1614–1624, Mar 2005.

[103] L. A. Shepp and B. F. Logan, "The Fourier reconstruction of a head section," *IEEE transactions on nuclear science*, vol. 21, no. 3, pp. 21–43, June 1974.

[104] S. Sivankutty, V. Tsvirkun, O. Vanvincq, G. Bouwmans, E. R. Andresen, and H. Rigneault, "Nonlinear imaging through a fermat's golden spiral multicore fiber," *Optics letters*, vol. 43, no. 15, pp. 3638–3641, 2018.

[105] E. J. Stollnitz, A. DeRose, and D. H. Salesin, "Wavelets for computer graphics: a primer. 1," *IEEE Computer Graphics and Applications*, vol. 15, no. 3, pp. 76–84, 1995.

[106] V. Studer, J. Bobin, M. Chahid, H. S. Mousavi, E. Candès, and M. Dahan, "Compressive fluorescence microscopy for biological and hyperspectral imaging," *Proceedings of the National Academy of Sciences*, vol. 109, no. 26, pp. E1679–E1687, 2012.

[107] P. Sudhakar, L. Jacques, X. Dubois, P. Antoine, and L. Joannes, "Compressive imaging and characterization of sparse light deflection maps," *SIAM Journal on Imaging Sciences*, vol. 8, no. 3, pp. 1824–1856, 2015.

[108] B. Sun, M. P. Edgar, R. Bowman, L. E. Vittert, S. Welsh, A. Bowman, and M. Padgett, "3d computational imaging with single-pixel detectors," *Science*, vol. 340, no. 6134, pp. 844–847, 2013.

[109] T. Sun, G. Woods, M. F. Duarte, K. Kelly, C. Li, and Y. Zhang, "Obic measurements without lasers or raster-scanning based on compressive sensing," in *Int. Symposium for Testing and Failure Analysis (ISTFA), San Jose, CA*, 2009, pp. 272–277.

[110] J. J. Sylvester, "Thoughts on inverse orthogonal matrices, simultaneous signsuccessions, and tessellated pavements in two or more colours, with applications to newton's rule, ornamental tile-work, and the theory of numbers," *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 34, no. 232, pp. 461–475, 1867.

[111] D. Takhar, J. N. Laska, M. B. Wakin, M. F. Duarte, D. Baron, S. Sarvotham, K. F. Kelly, and R. G. Baraniuk, "A new compressive imaging camera architecture using optical-domain compression," in *Computational Imaging IV*, vol. 6065. International Society for Optics and Photonics, 2006, p. 606509.

[112] C. Taswell, "The what, how, and why of wavelet shrinkage denoising," *Computing in science & engineering*, vol. 2, no. 3, pp. 12–19, 2000.

[113] L. Thesing and A. Hansen, "Non uniform recovery guarantees for binary measurements and wavelet reconstructions," *to appear*.

[114] L. Thesing and A. C. Hansen, "Linear reconstructions and the analysis of the stable sampling rate," vol. 17, pp. 103–126, 2018.

[115] A. Thompson and R. Calderbank, "Compressive imaging using fast transform coding," in *Emerging Imaging and Sensing Technologies*, vol. 9992. International Society for Optics and Photonics, 2016, p. 99920F.

[116] Y. Tsaig and D. L. Donoho, "Extensions of compressed sensing," *Signal processing*, vol. 86, no. 3, pp. 549–571, 2006.

[117] E. Vandenberg and M. P. Friedlander, "SPGL1: A solver for large-scale sparse reconstruction," June 2007, http://www.cs.ubc.ca/labs/scl/spgl1.

[118] ——, "Probing the pareto frontier for basis pursuit solutions," *SIAM Journal on Scientific Computing*, vol. 31, no. 2, pp. 890–912, 2008.

[119] H. Vargas and H. Arguello, "A low rank model for compressive spectral image classification," *IEEE transactions on computational imaging, to appear*.

[120] R. Vershynin, *Introduction to the non-asymptotic analysis of random matrices*. Cambridge University Press, 2012, ch. 5, pp. 210–268.

[121] ——, *High-Dimensional Probability: An Introduction with Applications in Data Science*, ser. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2018.

[122] A. Wagadarikar, R. John, R. Willett, and D. Brady, "Single disperser design for coded aperture snapshot spectral imaging," *Applied optics*, vol. 47, no. 10, pp. B44–B51, 2008.

[123] M. Wakin, J. N. Laska, M. F. Duarte, D. Baron, S. Sarvotham, D. Takhar, K. F. Kelly, and R. G. Baraniuk, "Compressive imaging for video representation and coding," in *Picture Coding Symposium*, vol. 1, 2006, p. 13.

[124] Z. Wang and G. R. Arce, "Variable density compressed image sampling," *IEEE transactions on image processing*, vol. 19, no. 1, pp. 264–270, 2010.

[125] Y. Wiaux, L. Jacques, G. Puy, A. M. Scaife, and P. Vandergheynst, "Compressed sensing imaging techniques for radio interferometry," *Monthly Notices of the Royal Astronomical Society*, vol. 395, no. 3, pp. 1733–1742, 2009.

[126] M. Woringer, X. Darzacq, C. Zimmer, and M. Mir, "Faster and less phototoxic 3D fluorescence microscopy using a versatile compressed sensing scheme," *Optics express*, vol. 25, no. 12, pp. 13 668–13 683, Jun 2017.

[127] H. Young and R. Freedman, *University Physics: With Modern Physics*. Addison-Wesley, 2007.

[128] D. Yudovsky, A. Nouvong, and L. Pilon, "Hyperspectral imaging in diabetic foot wound care," *Journal of Diabetes Science and Technology*, vol. 4, no. 5, pp. 1099–1113, 2010.

[129] Q. Zhang, R. Plemmons, D. Kittle, D. Brady, and S. Prasad, "Joint segmentation and reconstruction of hyperspectral data with compressed measurements," *Applied optics*, vol. 50, no. 22, pp. 4417–4435, 2011.

[130] T. Zhang, "Sparse recovery with orthogonal matching pursuit under RIP," *IEEE transactions on information theory*, vol. 57, no. 9, pp. 6215–6221, Sep. 2011.

[131] Y. Zhang, M. P. Edgar, B. Sun, N. Radwell, G. M. Gibson, and M. J. Padgett, "3d single-pixel video," *Journal of Optics*, vol. 18, no. 3, p. 035203, 2016.

[132] Z. Zhang, X. Wang, G. Zheng, and J. Zhong, "Hadamard single-pixel imaging versus Fourier single-pixel imaging," *Optics Express*, vol. 25, no. 16, pp. 19 619–19 639, 2017.

[133] L. Zhihua and Z. Qishan, "Ordering of walsh functions," *IEEE Transactions on Electromagnetic Compatibility*, no. 2, pp. 115–119, 1983.