MULTILEVEL ILLUMINATION CODING FOR FOURIER TRANSFORM INTERFEROMETRY IN FLUORESCENCE SPECTROSCOPY

A. Moshtaghpour[1], L. Jacques[1]

ABSTRACT

Fourier Transform Interferometry (FTI) is an interferometric procedure for acquiring HyperSpectral (HS) data. Recently, it has been observed that the light source highlighting a (biologic) sample can be coded before the FTI acquisition in a procedure called Coded Illumination-FTI (CI-FTI). This turns HS data reconstruction into a Compressive Sensing (CS) problem regularized by the sparsity of the HS data. CI-FTI combines the high spectral resolution of FTI with the advantages of reduced-light-exposure imaging in biology.

In this paper, we leverage multilevel sampling scheme recently developed in CS theory to adapt the coding strategy of CI-FTI to the spectral sparsity structure of HS data in Fluorescence Spectroscopy (FS). This structure is actually extracted from the spectral signatures of actual fluorescent dyes used in FS. Accordingly, the optimum illumination coding as well as the theoretical recovery guarantee are derived. We conduct numerous numerical experiments on synthetic and experimental data that show the faithfulness of the proposed theory to experimental observations.

***Index Terms*—** Hyperspectral, Fourier transform interferometry, Fluorescence spectroscopy, Compressive sensing.

1. INTRODUCTION

Fourier Transform Interferometry (FTI) has received renewed interests in biomedical Fluorescence Spectroscopy (FS) where acquiring HyperSpectral (HS) data with high spectral resolution is crucial to distinguish constituents with slightly different spectral signatures [1–3].

When it is designed from a Michelson interferometer [4], the working principles of FTI for microscopic HS imaging of a biologic sample is explained as follows (see also Fig. 1). A spatially magnified HS light beam, denoted by $\boldsymbol{X}_c$, originating from the highlighted sample (*e.g.*, as obtained in confocal microscopy) is first divided by a Beam Splitter (BS). The resulting beam copies are then reflected back to the BS either by a fixed or by a moving mirror, this last element controlling the Optical Path Difference (OPD) of the two beams. After their recombination by the BS, the interferometric intensity of the resulting beam is finally recorded by a 2D imaging sensor.

Physical optics shows that FTI observations collected in each imager pixel for multiple values of the OPD $\xi \in \mathbb{R}$ sample the Fourier transform of the HS volume along the wavenumber domain parameterized by $\nu \in \mathbb{R}$, *i.e.*, ξ and ν are (Fourier) dual parameters. As an advantage, the spectral resolution of the HS volume can be increased by enlarging the range of recorded OPD values. However, this increase of resolution is limited by the durability of the fluorescent dyes when exposed to illumination. In fact, over-exposed fluorochromes lose their ability to fluoresce, *i.e.*, a phenomenon known as *photo-bleaching* [5].

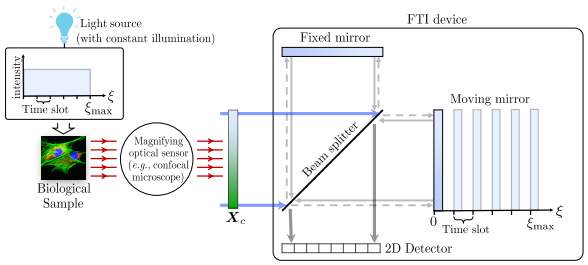

Fig. 1: Operating principle of FTI and CI-FTI.

Coded Illumination-FTI (CI-FTI) has been introduced in [6, 7] to mitigate the limitation of conventional FTI. By resorting to the theory of Compressive Sensing (CS) [8, 9] and temporally coding the illumination of the light source, CI-FTI succeeds in reconstructing the target HS volume while minimizing the global light exposure of the observed sample. The authors in [10] exploited a type of Variable Density Sampling (VDS) strategy developed in [11] to recover any HS volume whose spectra are sparsely representable in a wavelet basis.

This paper proposes to optimize further the illumination coding strategy for FS by leveraging the notion of MultiLevel Sampling (MLS) [12] (see Sec. 2), a promising VDS extension already applied in Magnetic Resonance Imaging (MRI) and in other FS experiments [13]. An optimum multilevel illumination coding scheme is thus established in Sec. 3 by exploiting the typical structure underlying all spectral sparsity patterns in FS experiments. This entails studying fluorochrome dataset [14] and the spectra therein. Conversely to [13] where the spectral dimension is scanned sequentially and MLS is applied on the spatial dimension, our approach applies MLS on the Fourier transform of the spectra.

The rest of the paper is structured as follows. We first summarize the recovery guarantee associated to MLS theory in Sec. 2. The new multilevel illumination coding is presented in Sec. 3. Sec. 4 demonstrates numerically the power of this approach before concluding the paper.

Notations: Vectors and matrices are associated with bold symbols. For a matrix $\boldsymbol{U} = (\boldsymbol{u}_1, \cdots, \boldsymbol{u}_{N_2}) \in \mathbb{C}^{N_1 \times N_2}$, $\boldsymbol{u} = \mathrm{vec}(\boldsymbol{U}) := (\boldsymbol{u}_1^\top, \cdots, \boldsymbol{u}_{N_2}^\top)^\top \in \mathbb{C}^{N_1 N_2}$ denotes the vectorization of $\boldsymbol{U}$. The probability of an event $\mathcal{U}$ reads $\mathbb{P}(\mathcal{U})$. 3D HS volumes will be either represented by their vector or matrix representations. The ℓ_2-norm is denoted by $\|\cdot\|$, while other ℓ_p norms read $\|\cdot\|_p$ with $1 \leq p \leq \infty$. The identity matrix of size $N \times N$ is represented as $\boldsymbol{I}_N$, the index set is

[1] ISPGroup, ICTEAM, UCLouvain, Belgium. The authors thank P. Antoine and M. Roblin (Lambda-X SA, Nivelles, Belgium) for their help in the acquisition of the FTI measurements. AM is funded by the FRIA/FNRS. LJ is funded by the F.R.S.-FNRS.

$[N] := \{1, \cdots, N\}$, and $|\mathcal{S}|$ is the cardinality of a set $\mathcal{S}$. The symbols $\mathbf{\Phi}$ and $\mathbf{\Psi}$ are reserved for unitary bases, with $\mathbf{\Phi}_{\rm DFT}$ and $\mathbf{\Psi}_{\rm DHW}$ denoting the 1D discrete Fourier (DFT) and Haar wavelet (DHW) basis , respectively. For two functions f and g, we write $f \lesssim g$ if $f \leq c\, g$ for some universal constant $c > 0$, and $f \gtrsim g$ if $g \lesssim f$.

2. COMPRESSED SENSING FOR SIGNALS WITH STRUCTURED SPARSITY

One branch of CS theory studies the recovery of a signal $\boldsymbol{x} \in \mathbb{C}^N$ from a vector of measurements $\boldsymbol{y} = \boldsymbol{P}_\Omega \mathbf{\Phi}^* \boldsymbol{x} + \boldsymbol{\eta}$ [15], where $\Omega = \{\Omega_1, \cdots, \Omega_M\} \subset [N]$ is a set of indices of cardinality M, $\boldsymbol{P}_\Omega \in \mathbb{C}^{N\times N}$ is a projection operator with $(\boldsymbol{P}_\Omega \boldsymbol{x})_j = x_j$ if $j \in \Omega$ (and zero otherwise), and $\boldsymbol{\eta}$ is an additive noise on the measurement with bounded power $\|\boldsymbol{\eta}\| \leq \varepsilon$. If $\boldsymbol{x}$ is assumed sparse (or well approximated by a sparse representation) in some basis $\mathbf{\Psi}$, *i.e.*, $\boldsymbol{x} = \mathbf{\Psi}\boldsymbol{s}$, with $K := |\mathrm{supp}\,(\boldsymbol{s})| \ll N$, then this signal can be recovered by

$$\hat{\boldsymbol{x}} = \arg\min_{\boldsymbol{v}\in\mathbb{C}^N} \|\mathbf{\Psi}^* \boldsymbol{v}\|_1 \text{ s.t. } \|\boldsymbol{y} - \boldsymbol{P}_\Omega \mathbf{\Phi}^* \boldsymbol{v}\| \leq \varepsilon, \quad (1)$$

provided $\mathbf{\Phi}$ and $\mathbf{\Psi}$ respect some incoherent condition [15].

In this paper, we are concerned in optimizing the sensing procedure above for signals displaying a structured form of sparsity. Following [12], this structure is best captured by splitting the signal sparsity patterns "in levels", which in turn will also split similarly the sensing procedure itself. Let us describe this in details.

For a fixed $r \in \mathbb{N}$ we first decompose $[N]$ into r disjoint *sparsity levels* $\mathcal{T} := \{\mathcal{T}_1, \cdots, \mathcal{T}_r\}$ such that $\bigcup_{\ell=1}^r \mathcal{T}_\ell = [N]$. Similarly, r disjoint *sampling levels* are defined as $\mathcal{W} := \{\mathcal{W}_1, \cdots, \mathcal{W}_r\}$ with $\bigcup_{\ell=1}^r \mathcal{W}_\ell = [N]$.

Given the parameters $\boldsymbol{k} = (k_1, \cdots, k_r)^\top \in \mathbb{N}^r$, a vector $\boldsymbol{s} \in \mathbb{R}^N$ is called $(\boldsymbol{k}, \mathcal{T})$-sparse in levels, and we write $\boldsymbol{s} \in \Sigma_{\boldsymbol{k},\mathcal{T}}$, if $|\mathrm{supp}\, \boldsymbol{P}_{\mathcal{T}_\ell} \boldsymbol{s}| \leq k_\ell$ for all $\ell \in [r]$. For an arbitrary vector $\boldsymbol{s}$, its $(\boldsymbol{k}, \mathcal{T})$-approximation error is $\sigma_{\boldsymbol{k},\mathcal{T}}(\boldsymbol{s}) := \min\{\|\boldsymbol{s} - \boldsymbol{z}\|_1 : \boldsymbol{z} \in \Sigma_{\boldsymbol{k},\mathcal{T}}\}$. Moreover, given an isometry $\boldsymbol{U} \in \mathbb{C}^{N\times N}$, the $t^{\rm th}$ *relative sparsity* is defined by

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

Given $\boldsymbol{m} = (m_1, \cdots, m_r)^\top \in \mathbb{N}^r$, the set $\Omega_{\mathcal{W},\boldsymbol{m}} := \bigcup_{t=1}^r \Omega_t$ provides a multilevel sampling scheme, or $(\mathcal{W}, \boldsymbol{m})$-MSS, 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 in $\mathcal{W}_t$. Furthermore, the $(t,\ell)^{\rm th}$ *local coherence* of $\mathbf{\Phi}$ with respect to $\mathbf{\Psi}$ is

$$\mu_{t,\ell}^{\mathcal{W},\mathcal{T}}(\mathbf{\Phi}, \mathbf{\Psi}) := \sqrt{\mu(\boldsymbol{P}_{\mathcal{W}_t} \mathbf{\Phi}^* \mathbf{\Psi})\, \mu(\boldsymbol{P}_{\mathcal{W}_t} \mathbf{\Phi}^* \mathbf{\Psi} \boldsymbol{P}_{\mathcal{T}_\ell})}, \quad (3)$$

where $\mu(\boldsymbol{U}) := \max_{i,j} |U_{i,j}|^2 \in \left[N^{-1}, 1\right]$. Within this context, (1) is guaranteed to find a good signal estimate in the following sense [12].

Proposition 1 ([12]). *Let $\Omega = \Omega_{\mathcal{W},\boldsymbol{m}}$ be a $(\mathcal{W}, \boldsymbol{m})$-MSS and $(\boldsymbol{k}, \mathcal{T})$ be any pair such that the following holds: for $0 < \epsilon \leq \exp(-1)$, $K = k_1 + \cdots + k_r$, and $1 \leq t \leq r$,*

$$m_t \gtrsim |\mathcal{W}_t|\, (\textstyle\sum_{\ell=1}^r \mu_{t,\ell}^{\mathcal{W},\mathcal{T}}(\mathbf{\Phi}, \mathbf{\Psi})\, k_\ell)\, \log(K\epsilon^{-1})\, \log(N), \quad (4)$$

$\mathbf{\Psi}_{\rm DHW}$	$\mathcal{T}_\ell = \{l_{\ell-1}+1, \cdots, l_\ell\}$, $1 \leq \ell \leq r$. $l_0 = 0$, $l_\ell = 2^\ell$, for $\ell \in [r]$; and $r = \log_2(N)$.
	$\mathcal{W}_{t+1} := \{-n_t+1, \cdots, n_t\} \backslash \mathcal{W}_t$, $1 \leq t \leq r-1$ and $\mathcal{W}_1 = \{0, 1\}$. $n_0 = 0$, $n_t = 2^t$, for $t \in [r]$.
$\mathbf{\Psi}_{\rm DFT}$	$\mathcal{T}_\ell = \{-l_\ell+1, \cdots, -l_{\ell-1}\} \bigcup \{l_{\ell-1}, \cdots, l_\ell\}$, $1 \leq \ell \leq r$. $l_0 = 0$, $l_\ell = \ell\, N\, 2^{-q+1}$, for $\ell \in [r]$; and $r = 2^q \ll N$.
	$\mathcal{W}_t = \mathcal{T}_t$, $1 \leq t \leq r$. $n_0 = 0$, $n_t = t\, N\, 2^{-q+1}$, for $t \in [r]$.

Table 1: Definitions of sparsity and sampling levels where $\mathbf{\Phi} = \mathbf{\Phi}_{\rm DFT}$ and $\mathbf{\Psi}$ is indicated in the first column. See Cor. 1.

where $m_t \gtrsim \hat{m}_t\, \log(K\epsilon^{-1})\, \log(N)$, and $\hat{m}_t$ is such that

$$1 \gtrsim \textstyle\sum_{t=1}^r (\frac{|\mathcal{W}_t|}{\hat{m}_t} - 1)\, \mu_{t,\ell}^{\mathcal{W},\mathcal{T}}(\mathbf{\Phi}, \mathbf{\Psi})\, K_t(\mathcal{W}, \mathcal{T}, \boldsymbol{k}), \ \ \ell \in [r]. \quad (5)$$

Suppose that $\hat{\boldsymbol{x}} \in \mathbb{C}^N$ is a minimizer of (1)*. Then, with probability exceeding $1 - \epsilon$, we have*

$$\|\boldsymbol{x} - \hat{\boldsymbol{x}}\| \leq \beta_1\, \sigma_{\boldsymbol{k},\mathcal{T}}(\mathbf{\Psi}^* \boldsymbol{x}) + \beta_2\, \varepsilon, \quad (6)$$

for some constant $\beta_1, \beta_2 > 0$.

Special sparsity cases: The acquisition physics of many applications, *e.g.*, MRI, tomography, electron microscopy, radio interferometry, and FS using FTI (the purpose of this paper), imposes using the Fourier operator in the sensing model. While [12] has considered the first four applications above in the context of the Haar wavelet sparsity basis [16], this paper proposes studying (in Prop. 1) the benefit of two possible sparsity bases for FTI in FS, *i.e.*, $\mathbf{\Psi} = \mathbf{\Psi}_{\rm DHW}$ and $\mathbf{\Psi} = \mathbf{\Psi}_{\rm DFT}$. For the DHW, we follow the conventions of [13] by assigning the sparsity levels to the natural dyadic wavelet levels, also associated with dyadic bands for the sampling levels. For the DFT choice, we impose $\mathcal{W} = \mathcal{T}$ and symmetric levels (around the DC frequency) with identical cardinality. Table 1 gathers all our level definitions for the two systems.

Corollary 1. *In the context of the sparsity and sampling levels defined in Table 1, let us set $\mathbf{\Phi} = \mathbf{\Phi}_{\rm DFT}$. Each of the two requirements below implies* (4) *and* (5) *in Prop. 1:*

$$\text{(DHW)}\ m_t \gtrsim (\textstyle\sum_{\ell=1}^r 2^{-|t-\ell|/2}\, k_\ell)\, \log(K\epsilon^{-1})\, \log(N) \quad (7)$$

$$\text{(DFT)}\ \ m_t \gtrsim \min\left\{|\mathcal{W}_t|, |\mathcal{W}_t|\, k_t\, \log(K\epsilon^{-1})\, \log(N)\right\}. \quad (8)$$

Proof. See [17] for the case of DHW. In the second case, given $\mathcal{W} = \mathcal{T}$ we can show that $K_t(\mathcal{W}, \mathcal{T}, \boldsymbol{k}) \leq k_t$ and $\mu_{t,\ell}^{\mathcal{W},\mathcal{T}}(\mathbf{\Phi}_{\rm DFT}, \mathbf{\Psi}_{\rm DFT}) = \delta_{t,\ell}$, where δ_{ij} is the Kronecker symbol. Applying these in Prop. 1 completes the proof. □

Relation (8) enforces full sampling for the levels where $k_t > 0$. However, in the case that the majority of k_t are zero, the use of DFT sparsity basis will require very few number of measurements (see Sec. 3 and Sec. 4).

3. MAIN RESULTS: CI-FTI IN FS

We here focus on the framework of CI-FTI proposed in [6]. We refer the reader to [10] for the details on the acquisition principles of FTI and its variations. In short, a vectorized form of the CI-FTI acquisition is modeled as $\boldsymbol{Y} = \boldsymbol{P}_\Omega \mathbf{\Phi}^*_{\rm DFT} \boldsymbol{X} + \boldsymbol{W}_{\rm CI}$, where $\boldsymbol{W}_{\rm CI} := \boldsymbol{P}_\Omega \boldsymbol{W}_{\rm Nyq}$, $\boldsymbol{X} \in \mathbb{R}^{N_\xi \times N_p}$ is an HS volume with N_ξ spectral bands and N_p spatial pixels, and $\boldsymbol{W}_{\rm Nyq} = (\boldsymbol{w}_1, \cdots, \boldsymbol{w}_{N_p})$ is a noise matrix corrupting the Nyquist measurements and satisfying $\|\boldsymbol{w}_j\| \leq \varepsilon_{\rm Nyq}/\sqrt{N_p}$

Approach	α	β_1	β_2
[10]	$\sqrt{M_\xi/N_p}$	$2/\sqrt{K}$	$\sqrt{N_p}$
This work	$\sqrt{M_\xi/N_\xi N_p}$	c	$c\cdot\sqrt{M_\xi N_p/N_\xi}$

Table 2: The value of the variables in (9) and (10) with respect to different approaches, where c is a constant and $M_\xi = m_1 + \cdots + m_r$.

for all $j \in [N_p]$. Similarly to [10], in this paper we would like to recover the spectra at all the pixels by solving the following convex optimization problem for all $j \in [N_p]$:

$$\hat{\boldsymbol{x}}_j = \underset{\boldsymbol{u}\in\mathbb{R}^{N_\xi}}{\arg\min} \|\boldsymbol{\Psi}^*\boldsymbol{u}\|_1 \text{ s.t. } \|\boldsymbol{D}(\boldsymbol{y}_j - \boldsymbol{P}_\Omega \boldsymbol{\Phi}^*_{\mathrm{DFT}}\boldsymbol{u}_j)\| \le \alpha\varepsilon_{\mathrm{Nyq}}, \quad (9)$$

where $\boldsymbol{D}$ and α are defined below according to the selected sensing scheme. As will be clear later, we can show that the error on $\boldsymbol{x}$ is bounded as

$$\|\boldsymbol{x} - \hat{\boldsymbol{x}}\| \le \beta_1 \cdot \textstyle\sum_{j=1}^{N_p} \sigma_{\boldsymbol{k},\mathcal{T}}(\boldsymbol{\Psi}^*\boldsymbol{x}_j) + \beta_2 \cdot \varepsilon_{\mathrm{Nyq}}, \quad (10)$$

for some $\beta_1, \beta_2 > 0$ only depending on $|\Omega|$, N_ξ, and N_p.

Initial VDS scheme: Moshtaghpour et al. [10] propose a VDS scheme for temporal coding of light illumination inspired by [11]. They subsample the rows of $\boldsymbol{\Phi}^*_{\mathrm{DFT}}$ by selecting $M_\xi \gtrsim K\log^3(K)\log^2(N_\xi)$ OPD indices iid according to a *pmf* inversely proportional to the OPD magnitude, *i.e.*, $p(i) \propto \min\{1, |i - N_\xi/2|^{-1}\}$ for all $i \in [N_\xi]$. The reconstruction procedure is then ensured by (9) by assigning the diagonal matrix $\boldsymbol{D}$ as $D_{ii} = p(\Omega_i)^{-1/2}$, and the reconstruction error is then bounded, with high probability (w.h.p.), as in (10), by setting there $\boldsymbol{k} = (K) \in \mathbb{N}$, $\mathcal{T} = [N_\xi]$, and the other variables being given in Table 2.

FS-driven MLS: Although this VDS scheme in [10] can be applied on any compressive FTI system (provided the HS data are sparse), boosted reconstruction quality can be reached by leveraging the sparsity structure in levels [12]. Practically, the sparsity pattern of a target signal is unknown. However, a common approach is to consider a class of similar signals and to estimate a sparsity pattern, *i.e.*, inclusive for that class of signals. For some examples of this approach see [12, 18].

Given a sparsity basis $\boldsymbol{\Psi}$ and the sparsity levels $\mathcal{T}$, we estimate the spectral sparsity pattern of the HS volumes in FS as follows. *(i)* We form a dictionary $\boldsymbol{H} \in \mathbb{R}^{N_\nu \times N_f}$ by collecting the spectra of N_f fluorochromes commonly used in FS [1], see Fig. 2-(top). *(ii)* The columns of this dictionary are then represented in $\boldsymbol{\Psi}$ domain, *i.e.*, $\boldsymbol{H} = \boldsymbol{\Psi}\tilde{\boldsymbol{H}}$. *(iii)* For every $\tilde{\boldsymbol{h}}_i$, *i.e.*, the i^{th} column of $\tilde{\boldsymbol{H}}$, we define $\pi_{i,j}$ as the index set of the j-largest (in absolute value) coefficients of $\tilde{\boldsymbol{h}}_i$. Given $\rho \in [0,1]$, we define the local sparsity of $\tilde{\boldsymbol{h}}_i$ at level ℓ, as

$$k_{i,\ell}(\rho) := |\pi_{i,k(\rho)} \cap \mathcal{T}_\ell|,$$

where $k(\rho) := \min\{n : \|\tilde{\boldsymbol{h}}_{i,\pi_{i,n}}\| \ge \rho\|\tilde{\boldsymbol{h}}_i\|\}$. (iv) Finally, in order to obtain an estimation applicable in FS experiments, we consider the worst local sparsity value among all the fluorochromes included in dictionary $\boldsymbol{H}$, *i.e.*,

$$k^0_\ell(\rho) := \max\{k_{i,\ell}(\rho) : 1 \le i \le N_f\}.$$

For a proper choice of sparsity basis and sparsity levels we should observe that the normalized ratios $k^0_\ell(\rho)/|\mathcal{T}_\ell|$ decay rapidly for a fixed ρ.

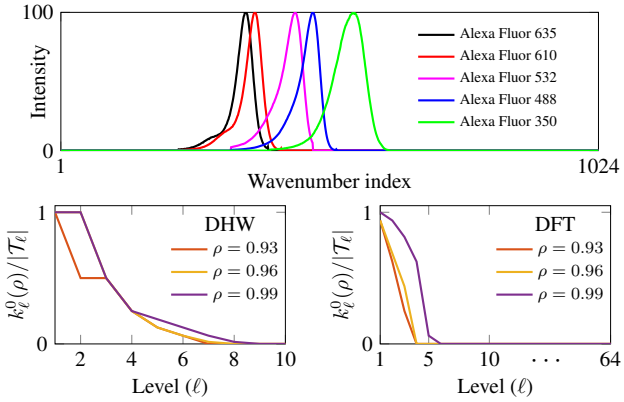

Fig. 2: The spectra of five fluorochromes (top); The estimated local sparsity ratio for a collection of 38 fluorochrome spectra (bottom).

As a proof of concept, we have applied our approach on a collection of $N_f = 38$ spectra of common fluorochromes [14], frequently used as cell and tissue labels in FS [1]. This includes the spectra of Alexa Fluors (Fig. 2-top). We conducted our test with $\boldsymbol{\Psi} = \boldsymbol{\Psi}_{\mathrm{DHW}}$ and $\boldsymbol{\Psi} = \boldsymbol{\Psi}_{\mathrm{DFT}}$, and the corresponding sparsity and sensing levels defined in Table 1 (with $q = 6$ for the DFT case). We observe in Fig. 2-bottom that fluorochrome spectra do display structured sparsity pattern in both DFT and DHW bases. In the DHW basis, local sparsity ratio decreases when the wavelet level increases. Moreover, in the DFT basis, even for a severe $\rho = 0.99$, 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 DFT basis is the main reason for superior HS reconstruction that will be followed in Sec. 4.

Validity of this approach: Our recovery guarantee is valid under the Linear Mixing Model (LMM), which is a common assumption (see *e.g.*, [19]). In LMM, any HS volume $\boldsymbol{X} \in \mathbb{R}^{N_\xi \times N_p}$ explained by the spectra matrix $\boldsymbol{H}$ can be modeled as the product $\boldsymbol{X} = \boldsymbol{H}\boldsymbol{G}$, where $\boldsymbol{G} \in \mathbb{R}_+^{N_f \times N_p}$ is a mixing matrix with nonnegative entries representing the spatial concentration of each fluorochromes. If $\boldsymbol{S} := \boldsymbol{\Psi}^*\boldsymbol{X}$ is the transformation of (the columns of) $\boldsymbol{X}$ in the spectral basis $\boldsymbol{\Psi}$, then, $\boldsymbol{S} = \tilde{\boldsymbol{H}}\boldsymbol{G}$ and, for each pixel $1 \le j \le N_p$, the coefficients $\boldsymbol{s}_j$ are mixed as $\boldsymbol{s}_j = \tilde{\boldsymbol{H}}\boldsymbol{g}_j$. Therefore, $\operatorname{supp}\boldsymbol{s}_j \subseteq \bigcup_{i=1}^{N_f} \operatorname{supp}\tilde{\boldsymbol{h}}_i$. Consequently, any sparsity structure shared among the fluorochrome spectra in $\boldsymbol{H}$ is preserved by every spectrum of the HS volume.

HS Recovery guarantees: Let us consider a CI-FTI system developed based on one of the two schemes of Table 1 and adjusted to the sparsity structure of $\boldsymbol{H}$ as described above. Prop. 1 can be invoked to characterize the error bound (10) on the reconstruction of HS data from (9) with $\boldsymbol{D} = \boldsymbol{I}_{N_\xi}$. We select two possible strategies. First, we can set $\boldsymbol{\Psi} = \boldsymbol{\Psi}_{\mathrm{DHW}}$, and, for each level $t \in [r]$, m_t OPD indices are picked uniformly at random in each sensing level with

$$m_t \gtrsim \left(\textstyle\sum_{\ell=1}^{r} 2^{-|t-\ell|/2}\, k_\ell\right) \log(KN_p\epsilon^{-1})\, \log(N_\xi). \quad (11)$$

Second, we set $\boldsymbol{\Psi} = \boldsymbol{\Psi}_{\mathrm{DFT}}$ and suppose that there exists an integer $r_0 \ll r$ such that $k_{t>r_0} = 0$. In this case, (8) imposes

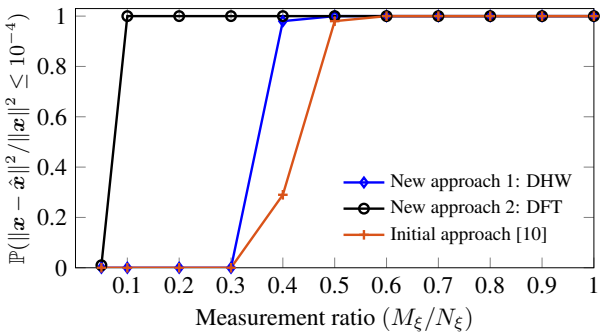

Fig. 3: Comparison of the successful recovery rate with the proposed approaches and the initial approach in [10].

to fully sample the frequencies in the levels $t \in [r_0]$ for any fixed value of $0 < \epsilon \leq \exp(-1)$.

Then, from Cor. 1, we can easily deduce that, by union bound over all the N_p pixels/spectra, the HS recovery error in (10) holds with probability exceeding $1 - N_p(\epsilon/N_p) = 1 - \epsilon$ and with the parameters of Table 2. Note that when the number of measurements increases, β_2 in (10) increases too. Thus, taking more measurements does not necessarily imply superior reconstruction. We will see this effect in Sec. 4.

4. NUMERICAL RESULTS

Let us now evaluate the performance of CI-FTI using the two FS-driven approaches. All HS data were reconstructed by solving (9) with the SPGL1 toolbox [20].

The first experiment traces the phase transition curves of successful recovery of a single spectrum with respect to the measurement ratio M_ξ/N_ξ in the noiseless case. The measurements are formed as $\boldsymbol{y} = \boldsymbol{P}_\Omega \boldsymbol{\Phi}^*_{\mathrm{DFT}} \boldsymbol{x}$, where $\boldsymbol{x} := \boldsymbol{\Psi H g}$ is a synthetic spectrum resulting from the linear mixing of N_f spectra, as realized by $\boldsymbol{g} \in [0,1]^{N_f}$ with $g_i \sim_{\mathrm{iid}} \mathcal{U}([0,1])$. In our approaches the set Ω is generated depending on the local sparsity values (see Sec. 3). Fig. 3 shows the probability of successful reconstruction over 100 independent realizations of Ω and $\boldsymbol{g}$. The improvement of the two approaches is due to the structured sparsity exploited for the optimum subsampling strategy. In addition, since the fluorescent spectra are highly compressible in the DFT basis, the successful recovery rate is boosted for the DFT-based approach.

The second experiment consists in simulating CI-FTI measurements from actual FTI measurements recorded at Nyquist regime. The full description of this experiment is explained in [10]. In short, we observed a thin layer of a transparent cell, *i.e.*, Convallaria, lily of the valley. The Nyquist-FTI measurements of size $(N_\xi, N_x, N_y) = (1024, 128, 128)$ was obtained at current level 700 mA. We formed CI-FTI measurement by subsampling 10% of the Nyquist measurements using the projection $\boldsymbol{P}_\Omega$. The reference HS volume, in this test, is the one reconstructed from the Nyquist sensing.

The reconstructed HS volumes are illustrated in Fig. 4. Recall that in CI-FTI the spatial dimension of the measurements is not subsampled; hence, the spatial configuration of

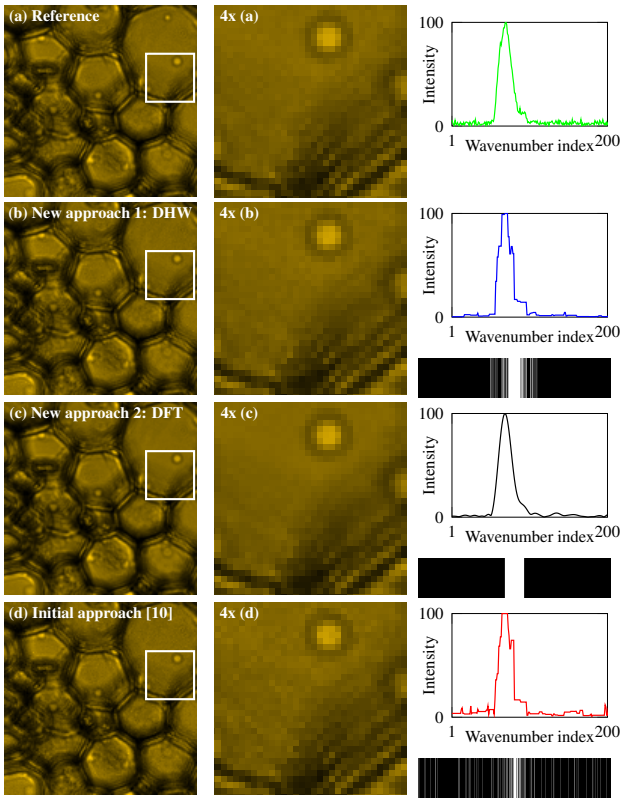

Fig. 4: The reconstructed HS volumes. (left) The spatial maps at 594 nm. (right) The spectra at the center pixel, here restricted to the first 200 indices of wavenumber axis. The coding pattern is shown in the last column. The DFT-based approach results in smoother and less noisy reconstruction.

the specimen is preserved. We have thus to assess the existence of the noise elements in the reconstructed spatial maps as well as the spectra. The former is clear in the second column. The reconstructed spectra witness the superiority of the two proposed approaches. As mentioned before, increasing the number of measurements results in more noise terms in the reconstructed spectra (recall parameter $\beta_2 = O(M_\xi^{1/2})$ in Table 2). This can be seen in the reference spectra. However, omitting the noise terms, we can see the shape of the reference spectrum. This shape is accurately preserved by the DFT-based approach, thanks to the proper choice of sparsity basis. Furthermore, as a result of subsampling few coefficients, the noise terms are significantly reduced.

5. CONCLUSION

We have presented new coding designs for CI-FTI that are adapted to FS. These schemes are derived from the notions of sparsity in levels, multilevel sampling, and local coherence in levels [12]. For our application, the common sparsity pattern among different fluorochrome spectra has been extracted. Our conclusion is that adapting CI-FTI with the corresponding multilevel illumination coding improves the quality of the reconstructed spectra. Extending this work to coded aperture-FTI, *i.e.*, spatio-temporal FTI coding [7], will be the scope of future research.

References

[1] 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.

[2] 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.

[3] G. Lu and B. Fei, "Medical hyperspectral imaging: a review," *Journal of biomedical optics*, vol. 19, no. 1, pp. 010 901–010 901, 2014.

[4] R. Bell, *Introductory Fourier transform spectroscopy*. Elsevier, 2012.

[5] R. Ghauharali and G. Brakenhoff, "Fluorescence photobleaching-based image standardization for fluorescence microscopy," *Journal of microscopy*, vol. 198, no. 2, pp. 88–100, 2000.

[6] 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.

[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] D. L. Donoho, "Compressed sensing," *IEEE Transactions on information theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[9] E. J. Candès and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?" *IEEE transactions on information theory*, vol. 52, no. 12, pp. 5406–5425, 2006.

[10] A. Moshtaghpour, V. Cambareri, P. Antoine, M. Roblin, and L. Jacques, "A variable density sampling scheme for compressive Fourier transform interferometry," 2018, arXiv:1801.10432v1.

[11] 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.

[12] B. Adcock, A. C. Hansen, C. Poon, and B. Roman, "Breaking the coherence barrier: A new theory for compressed sensing," *arXiv preprint arXiv:1302.0561*, 2013.

[13] B. Roman, A. C. Hansen, and B. Adcock, "On asymptotic structure in compressed sensing," *arXiv preprint arXiv:1406.4178*, 2014.

[14] "Fluorophores.org - Database of fluorescent dyes, properties and applications," retrieved on Jan. 5th, 2018. [Online]. Available: http://www.fluorophores.tugraz.at

[15] E. Candès and J. Romberg, "Sparsity and incoherence in compressive sampling," *Inverse problems*, vol. 23, no. 3, p. 969, 2007.

[16] S. Mallat, *A wavelet tour of signal processing: the sparse way*. Academic press, 2008.

[17] B. Adcock, A. C. Hansen, and B. Roman, "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.

[18] A. Bastounis and A. C. Hansen, "On the absence of the RIP in real-world applications of compressed sensing and the RIP in levels," *arXiv preprint arXiv:1411.4449*, 2014.

[19] N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE signal processing magazine*, vol. 19, no. 1, pp. 44–57, 2002.

[20] E. Vandenberg and M. P. Friedlander, "SPGL1: A solver for large-scale sparse reconstruction," June 2007, http://www.cs.ubc.ca/labs/scl/spgl1.