BTP Thesis

Topics in Raman spectroscopic imaging

Arijit Pramanik

150020094

Supervisors: *Prof. Ajit Rajwade*

2018-19

December 23, 2019

Contents

1 Introduction

Compressive imaging has emerged as a flourishing sub-field of signal and image processing in recent times [1]. It involves acquisition of images directly in a compressed format, followed by conversion of the compressive measurements to the conventional image format typically via efficient convex optimization procedures. These procedures typically exploit the inherent sparsity or compressibility of many modalities of images in well-known orthonormal bases such as the wavelet or discrete cosine transform bases [2]. The emphasis of compressive imaging is on saving acquisition time. The time-intensive acquisition of Raman images, as well as their inherent smoothness as signals, renders Raman imaging as a potential area for applications of compressed sensing.

1.1 Raman spectroscopy

Raman spectroscopy is a non-destructive, non-invasive method that can provide important information about various materials ranging from biological tissues [3] to materials in pharmacy, chemistry or materials science [4], [5], [6], [7], to art preservation [8]. Moreover, Raman imaging does not require usage of any additional chemical agents for the acquisition, and provides very high spatial resolution. Raman spectroscopy is also employed in the field of geology [9] for determining mineral composition of rocks as well as in the field of semiconductors [10] for probing into their structure and examining the presence of traces of other materials integrated into it.

The analysis of a sample is based on the investigation of a Raman spectrum that is a recording of intensities of scattered light. Recorded spectra are usually expressed according to the wavenumber shift (generally with the unity cm^{-1}) that is the inverse of the wavelength shift between the incident laser light and the scattered light.

1.1.1 Improving the speed of acquisition

Generally, the Raman signal is very weak (one in one hundred million incident photons). The signal strength can be very low for some sample types and 'requires powerful excitation compatible with the sample damage threshold' [11]. The absence of powerful enough illumination must be compensated for by significantly higher acquisition times for Raman spectra, so as to avoid the inherent shot noise which can lower the signal to noise ratio (SNR). This is particularly the case if the CCD detectors have low efficiency, and higher efficiency detectors significantly raise the instrument cost. We aim to work towards a compressive Raman imaging architecture, which involves measuring the spectra at only a subset of the total pixels of the image. The subset can be chosen randomly or in a structured fashion. In either case, this has the potential to immediately improve acquisition time. The missing pixels can be filled in, using an application of blind compressed sensing (BCS) [12] to implement an inpainting procedure.

We also present an alternative approach for the same task using Gaussian Mixture Models under the wide umbrella of Statistical Compressed Sensing. This branch of CS [13] aims to efficiently sample an ensemble of signals having accurate reconstruction on average. SCS doesn't explicitly restrict itself to sparse models, but works on general Bayesian models. Assuming that the signals follow the distribution $f(x)$, we desire encoder-decoder pairs (Φ, Δ), which minimizes the error $\int \|x - \Delta(\Phi x)\|_p f(x)dx$ where $\|.\|_p$ denotes the p-norm. The effectiveness of GMMs in describing real signals, along with theoretical guarantees motivates us to leverage this approach. A reduced number $M = O(k)$ of measurements with the optimal decoder implemented via linear ltering is signicantly faster than the decoders applied in conventional CS, which require $O(Cklog(n/k))$ measurements for satisfying the RIP property with $A = \Phi\Psi$. The average error of Gaussian SCS is tightly upper bounded by a constant times the best k-term approximation error with overwhelming probability, the failure probability being orders of magnitude smaller than that of conventional CS.

1.1.2 Source separation

A Raman spectrum can be seen as a signal composed of peaks or bands for which the positions, intensities, widths and shapes are informative about the molecular composition and structure of the analyzed sample. Raman spectroscopy has been widely applied especially for biomedical studies such as human skin characterization [14], cancer [15] and atherosclerosis [16] diagnosis, etc. In vitro Raman analysis are realized on frozen or dewaxed paraffin embedded biopsies. Paraffin is commonly used to preserve samples from decay, but has an intense Raman signature that prevents the study of the underlying tissue. A preliminary chemical dewaxing step, which is not totally efficient and can alter the sample [17], is usually applied to remove the paraffin. Here, we propose an alternate digital solution to remove the paraffin spectra using dictionary learning techniques. This is inspired from source separation methods which assume a linear mixing of independent sources. However, due to the uorescence emission, a non-linear additive noise is superimposed. The

spectral resolution introduces a spectral shift on each recorded spectrum, while the thickness imperfections of the sample section can be expressed as width variations of the Raman peaks.

2 Literature survey

2.1 Improving upon acquisition speed

Imaging spectroscopy involves the sensing of a large amount of spatial information across a multitude of wavelengths. Conventional approaches to hyperspectral sensing scan adjacent zones of the underlying spectral scene and merge the results to construct a spectral data cube. Spectrometers based on optical bandpass filters sequentially scan the scene by tuning the bandpass filters in steps. The disadvantage of these techniques is that they require scanning a number of zones linearly in proportion to the desired spatial and spectral resolution and complicated setups. One such architecture to alleviate this problem is compressive coded aperture snapshot spectral imager (CASSI) [18]. As usual, these utilise the inherent sparsity and their consequent sparse representation in well-known orthonormal bases. Given the source image $f(x, y, \lambda)$, it is modulated by means of a coded aperture $T(x, y)$. It is then modified by a dispersive prism before impinging on a FPA detector. This induces a wavelength dependent shift in the X direction. The spectral data cube F of dimensions $N_x \times N_y \times N_\lambda$ is mapped to a 2D snapshot of dimensions $(N_x + N_\lambda - 1) \times N_y$. Several such measurements are taken, each time with a different binary code T. Any such acquired snapshot is of the form : $\tilde{f}(x, y) = \sum_{k=0}^{N_\lambda - 1} f(x - l_\lambda, y, \lambda) T(x - l_\lambda, y)$. Adhering to the incoherence principles of CS, different binary codes are used for each such measurement. Since the entire spectral data cube is not measured, we get a compression ratio of $N_\lambda : m$ where m is the number of measurements. The original image can now be reconstructed using any optimization algorithm, say by a minimization of total variation.

2.2 Source separation

Prior work in this field [19], [20] and [21] has attempted to separate the tissue component of paraffin preserved samples to a limited degree of success. This forms a typical example of Blind Source Separation problem. These earlier works have focused on using linear BSS methods and hence, require several preprocessing steps to remove the non-linearities from the Raman spectra. These steps involve elimination of saturated spectra, correction of dark current, detector and optic responses, baseline removal, correction of peak misalignment and subsequently correction of peak width heterogeneity. Post pre-processing these generally bank on methods like PCA, ICA in order to determine and filter out the components due to pure paraffin. The most successful of these approaches include identifying sample points that consist of only pure paraffin. A PCA of this subspace composed of pure paraffin points yields the signal and noise subspace. It has been empirically observed that the first three principal components are a good enough approximation to the pure paraffin spectrum. This is thereafter used in JADE, a variant of ICA on the signal subspace to obtain four principal components out of which three belong to pure paraffin sources assumed to be mutually independent and the fourth is an estimate of the global tissue spectrum for the sample. Subsequently, a Non-negatively Constrained Least Squares approach is used in order to estimate the concentrations of each of these paraffin sources, which is then subtracted from the entire spectra to yield the tissue-only spectrum. Other approaches revolve around this basic foundation to provide the tissue-only estimate accounting for local variations across pixels.

3 Compressed sensing by pixel undersampling

3.1 Dictionary Learning

Let the image of interest be denoted as H, having size $N_x \times N_y \times N_\lambda$, where N_λ is the number of wavelengths. As per our architecture, the complete spectrum (of N_λ values) will be measured at only a fraction of the $N_x N_y$ pixels. No measurements are made at other pixels. Let the acquired incomplete image be denoted as G. The missing pixels in G need to be estimated via an efficient algorithm. For this purpose, we consider dividing H (and correspondingly G) into a number of overlapping patches, each of size $p \times p$. Let h_i, g_i be the patches in H and G respectively, at location indexed as i, expressed as $p^2 \times 1$ vectors. Then we have the following:

$$g_i = \Phi_i h_i + \eta_i, \tag{1}$$

where η_i is the noise vector at location i and Φ_i is a $p^2 \times p^2$ diagonal sensing matrix such that $\Phi_{i,jj}$ (the j^{th} diagonal element in Φ_i) contains 1 if g_{ij} is measured and 0 otherwise. The 0 entries can also be dropped off to give a compact non-zero measurement vector. If we follow a random sampling pattern, then the sensing

matrices for every patch will be different. However if the sampling pattern bears more regularity, i.e. if it measures only every l^{th} pixel ($l > 0$) in both directions (often termed 'decimation'), then there will be only p^2 different sensing matrices. In this case, the reconstruction is akin to an image super-resolution problem. Let n be the total number of patches thus considered. The task now reduces to estimating $\{h_i\}_{i=1}^n$ given $\{g_i, \Phi_i\}_{i=1}^n$. We now frame this as a BCS problem. For this, we express each patch h_i as a sparse linear combination of columns of a dictionary matrix A of size $p^2 \times K$, i.e. $h_i = As_i$ where s_i is a vector of sparse coefficients. Typical choices of dictionaries would include the wavelet or discrete cosine transform, since image patches are sparse (or approximately sparse) in these bases. However, given the inherently non-negative nature of the data, we impose the constraint that both A and s_i are both element-wise non-negative. With this in mind, we now seek to minimize the following objective function:

$$\begin{aligned} J(A, \{s_i\}_{i=1}^n) = \|g_i - \Phi_i A s_i\|^2 + \lambda \|s_i\|_1, \\ \text{such that } A \succeq 0, \forall i s_i \succeq 0, \\ \forall j \in \{1, ..., K\}, \|A_{.,j}\|_2^2 = 1 \end{aligned} \tag{2}$$

where $\lambda > 0$ is a sparsity-promoting parameter, $A_{.,j}$ is the j^{th} column of A, 0 represents a zero-valued matrix or vector, and $\succeq$ represents an element-wise 'greater than' inequality. This objective function seeks to solve a compressive version of the popular non-negative sparse coding (NNSC) algorithm [22]. In this work, we choose $K \ll p^2$ since a higher K increases the number of degrees of freedom, and the size of most available Raman spectral images is usually very small.

We implement the minimization of the function in Eqn. 3 using alternating minimization on the dictionary and sparse codes, starting from a random non-negative dictionary. Each step of the minimization is performed using projected gradient descent with adaptive step-size. That is, the stepsize of the gradient descent is adaptively chosen to ensure decrease of the objective function after imposition of all the constraints in Eqn. 3. The procedure is iterated till convergence, which is guaranteed due to the biconvex nature of the objective function. The algorithm is summarized in Alg. 1. Once the individual patches $h_i = As_i$ are reconstructed, an estimate of H is assembled by sliding window averaging.

3.1.1 Dictionary Inference for Regular Sampling

This problem is equivalent to performing super-resolution on the acquired image. In this case, we do not initialize A randomly. A much better initial guess is required in this case, because the sensing matrices for applications such as super-resolution tend to have high coherence with typical dictionaries (due to the regularity of the sampling patterns). Rather, we first perform bicubic interpolation on the image G to yield $\tilde{G}$. We then infer the dictionary $A_{\tilde{G}}$ from the patches of $\tilde{G}$ by optimizing the following objective function:

$$\begin{aligned} J_2(A_{\tilde{G}}, \{\tilde{s_i}\}_{i=1}^n) = \|\tilde{g_i} - A_{\tilde{G}} \tilde{s_i}\|^2 + \lambda \|\tilde{s_i}\|_1, \\ \text{such that } A_{\tilde{G}} \succeq 0, \forall i \tilde{s_i} \succeq 0, \\ \forall j \in \{1, ..., K\}, \|A_{\tilde{G}_{.,j}}\|_2^2 = 1. \end{aligned} \tag{3}$$

This now acts as an initial guess for inferring the actual dictionary A from H via Eqn. 3.

3.2 Gaussian Mixture Models

SCS with Gaussian models have been shown to have improved performance (bounds) relative to conventional CS, the signal reconstruction calculated with an optimal decoder Δ implemented via a fast linear ltering. SCS with piecewise linear decoders have been widely investigated. Generally, only those Gaussian signals are considered that show fast eigenvalue decay for the covariance matrix. In this case, we have a linear MAP decoder which is optimal and enjoys efficient calculation via closed form linear filtering for any Φ. The linear MAP estimator is given by $\Delta(\Phi x) = argmax_x p(x|y) = \Sigma\Phi^T(\Phi\Sigma\Phi^T)^{-1}(\Phi x)$ where $x \in R^N$ is the true signal with pdf $N(0, \Sigma)$, the sensing matrix $\Phi \in R^{M \times N}$ with $M \leq N$ and $y \in R^M$ is the measured signal such that $y = \Phi x$. Following from the above formulation where G is the provided undersampled version of the original image H, we dissect it into overlapping patches of size $p \times p$. Here, we provide a different approach to the task of estimating $\{h_i\}_{i=1}^n$ given $\{g_i, \Phi_i\}_{i=1}^n$. Each patch h_i is approximated by a Gaussian component belonging to the GMM $\{\mu_i, \Sigma_i\}_{i=1}^{N_c}$ where μ_i is the mean vector of size $p^2 \times 1$ and Σ_i is the covariance matrix of size $p^2 \times p^2$ for the i^{th} component and N_c is the number of components, which is fixed beforehand. The optimization is now carried out by the means of a MAP-EM algorithm.

Consider a general setting without any additive noise. In order to decode the measured signal $y = \Phi x$, the GMM-based SCS decoder estimates the signal $\tilde{x}$ and selects the Gaussian model $\tilde{j}$ by maximizing the

Algorithm 1 `INFER` A,$S = \{s_i\}_{i=1}^n$ `from G`

Require: $\mu, \lambda, K, p, n, \{\Phi_i\}_{i=1}^N, \epsilon, \alpha$

if G obtained from H by random sampling **then**
  $A \leftarrow random(p^2 N_\lambda, K)$
else
  Initialize $A = A_{\tilde{G}}$ inferred from $\tilde{G}$
end if
$S \leftarrow random(K, n)$

Ensure: $A \succeq 0$ **and** $\|A_j\|_2^2 = 1 \forall j \in \{1, .., K\}$ **and** $S \succeq 0$

$A^{old} \leftarrow A$, $S^{old} \leftarrow S$
$J \leftarrow \sum_{i=1}^n \|g_i - \Phi_i A s_i\|^2 + \lambda \sum_{i=1}^K \sum_{j=1}^n |s_{ij}|$
while $\Delta J > \epsilon$ **do**
  $A \leftarrow A - \mu(\sum_{i=1}^n \Phi_i^t(\Phi_i A s_i - g_i) s_i^t$
  Set all negative entries of A to 0

Ensure: $\|A_j\|_2^2 = 1 \; \forall j \in \{1, ..., K\}$ **and** $A \succeq 0$

  $S \leftarrow ((\Phi_i A)^t g_i)./((\Phi_i A)^t (\Phi_i A) s_i + \lambda)$
  $J' \leftarrow \sum_{i=1}^n \|g_i - \Phi_i A s_i\|^2 + \lambda \sum_{i=1}^K \sum_{j=1}^n |s_{ij}|$
  if $J^{'} > J$ **then**
    $\mu = \mu \times \alpha$ $(\alpha < 1$ is a reduction factor)
    $A \leftarrow A^{old}$, $S \leftarrow S^{old}$
  end if
end while

log a-posteriori probability, i.e. $(\tilde{x}, \tilde{j}) = argmax_{(x,j)} f(x|y, \mu_j, \Sigma_j)$. In order to realise this, we first compute the piecewise linear MAP estimates $\tilde{x_j}$ for each of the components and then select the best (MAP) model $\tilde{j}$.

$$\tilde{x}_j = \Sigma_j \Phi^T (\Phi \Sigma_j \Phi^T + \Sigma_\eta)^{-1} (y - \Phi \mu_j) + \mu_j$$

$$\tilde{j} = argmin_j \|y - \Phi \tilde{x}_j\|_{\Sigma_\eta}^2 + \|\tilde{x}_j - \mu_j\|_{\Sigma_j}^2 + log|\Sigma_j|$$

where $\|z\|_A$ denotes $z^T A^{-1} z$. We provide both flavours of SCS with GMMs on the fly as well as learning offline on a representative set of images. The latter is achieved by means of training on one-third of the image, and then reconstructing the rest from an undersampled version of the rest two-thirds of the image. A much better initial guess is required in the former case, because the initialized mean and covariance matrices are evolved down the line to yield the signal estimates and hence demand reasonable starting estimates for $\{\mu_j, \Sigma_j\}_{j=1}^{N_c}$ where N_c is the number of components. Rather, we first perform bicubic interpolation on the image G to yield $\tilde{G}$. We then infer the component $\{\mu_j, \Sigma_j\}_{j=1}^{N_c}$ from the patches of $\tilde{G}$ optimizing the following objective function:

$$J_3(\{h_i\}_{i=1}^n) = \|\tilde{g}_i - \Phi \tilde{h_{i\tilde{j}}}\|_{\Sigma_\eta}^2 + \|\tilde{h_{i\tilde{j}}} - \mu_{\tilde{j}}\|_{\Sigma_{\tilde{j}}}^2 + log|\Sigma_{\tilde{j}}|$$

where $\tilde{h_{i\tilde{j}}} = \Delta(\tilde{g}_i)$ for each patch using piecewise linear estimation as above. We now provide the brief steps for Maximum a-posteriori Expectation-Maximization algorithm which are carried out in an alternating fashion until convergence in terms of the objective function or the likelihood of the data given the GMM parameters. The core steps are :

- E-step : This carries out the piecewise linear estimation for $\{\tilde{h}\}_{i i=1}^n$ using $\{\mu_j, \Sigma_j\}_{j=1c}^N$ utilizing $\tilde{h_{i\tilde{j}}} = argmin_j \tilde{h_{ij}} = argmax_{(h_i,j)} f(h_i|\tilde{g_i}, \mu_j, \Sigma_j)$.
- M-step : In this step, the mean and covariance matrices are updated, given the signal estimates $(\tilde{h_i}, \tilde{j})$. We utilize the probability of each estimate belonging to a particular Gaussian component by means of a weight matrix W of size $n \times N_c$. Each entry w_{ij} corresponds to the probability of the i^{th} signal, (i.e. h_i) belonging to the j^{th} Gaussian component(characterised by (μ_j, Σ_j). Hence $w_{ij} = \frac{1}{(2\pi)^{\frac{N}{2}} |\Sigma_j|^{\frac{1}{2}}} exp(-\frac{1}{2} h_i^T \Sigma_j^{-1} h_i)$. Correspondingly, for the j^{th} component, we have

$$\mu_j = \frac{1}{S_j} \sum_i w_{ij} \tilde{h}_i, \qquad \Sigma_j = \frac{1}{S_j} \sum_i w_{ij} (\tilde{h}_i - \mu_j)(\tilde{h}_i - \mu_j)^T$$

  where $S_j = \sum_{i=1}^n w_{ij}$ is the sum of probabilities of all signals belonging to the j^{th} Gaussian component.

3.3 Results

3.4 Dictionary Learning

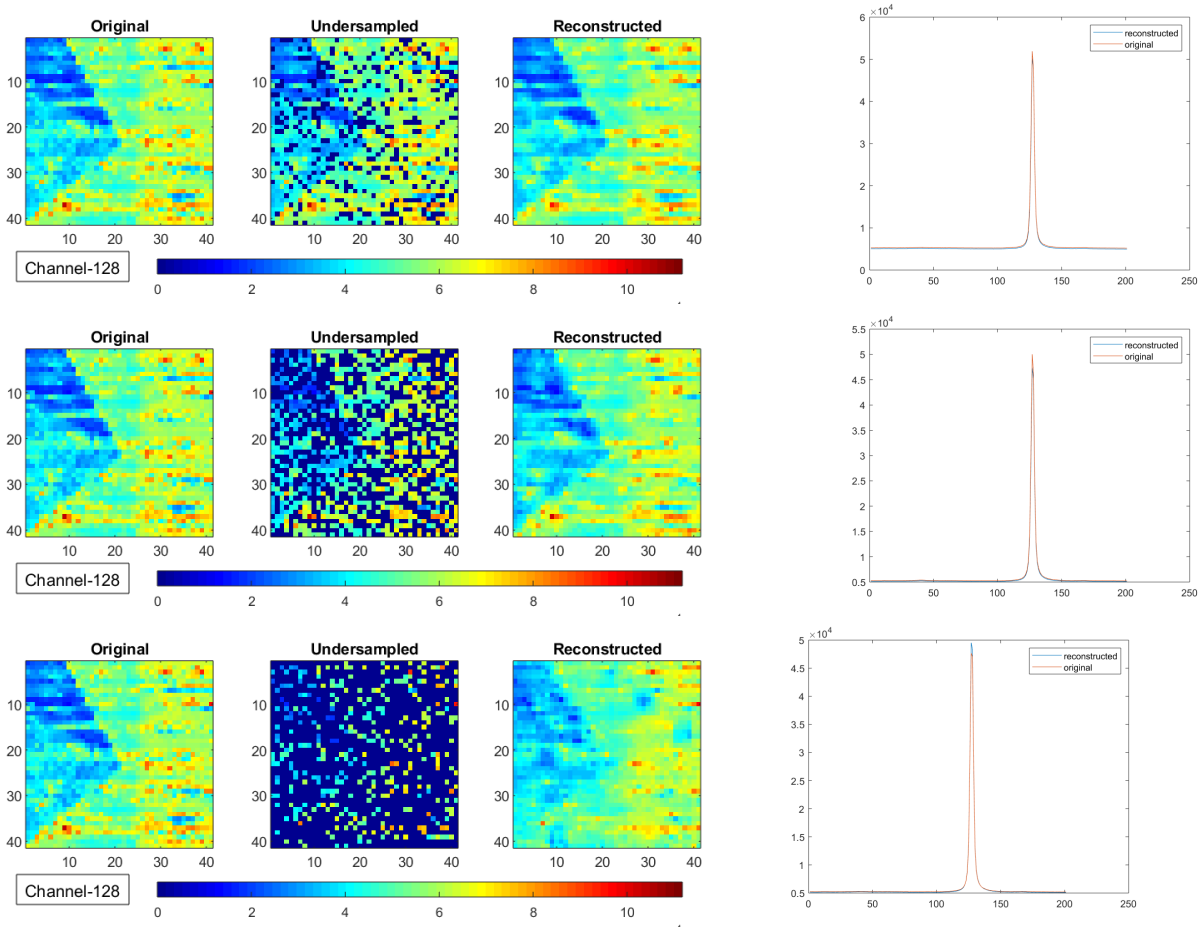
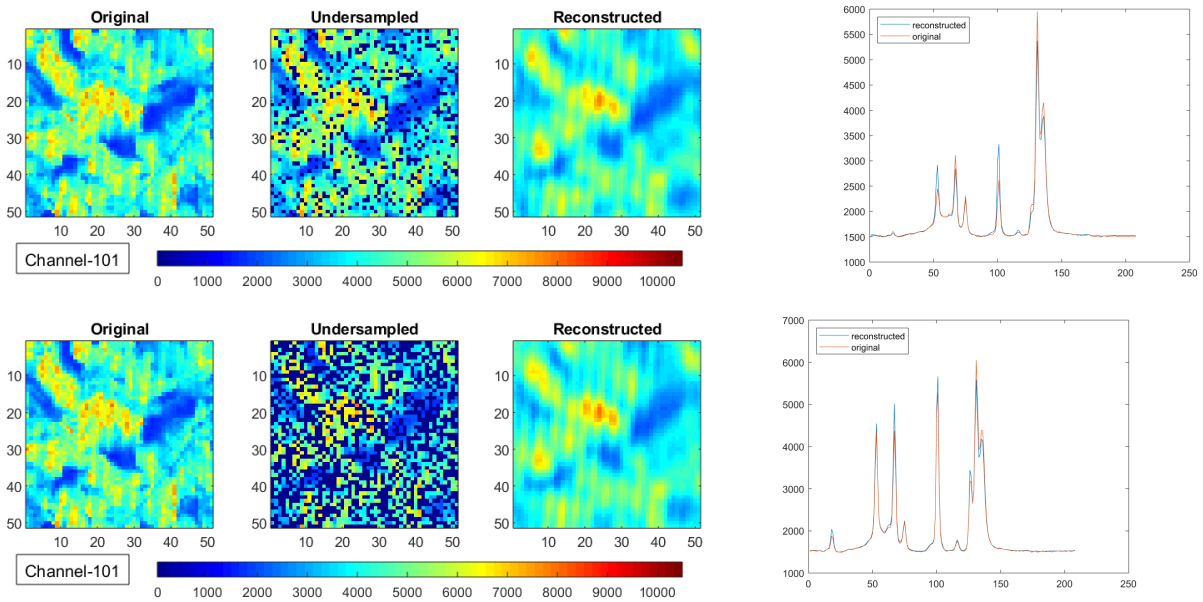


Figure 1: Top to bottom : 80%, 50%, 20% sampling with **RMSE : 0.07405**, **0.08940** and **0.14299** respectively. Left to right per row: ground truth, **randomly** sampled, reconstructed Raman spectral image and spectral plot of reconstructed pixel of pure silicon of size $41 \times 41 \times 226$ at 128^{th} spectral band. See supp. mat. for video results.



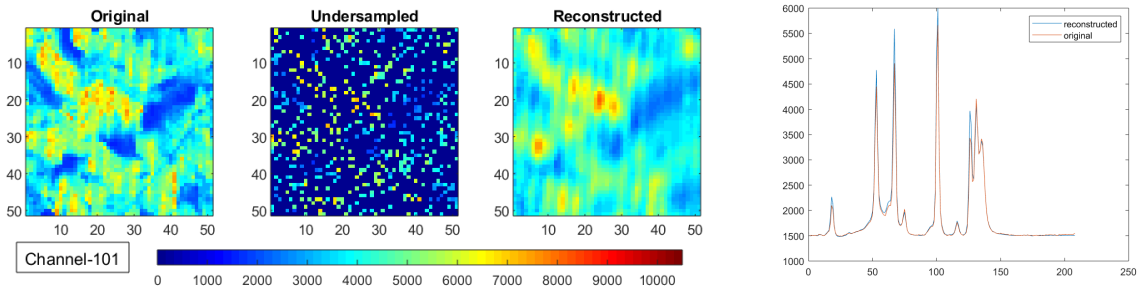


Figure 2: Top to bottom: 80%, 50%, 20% sampling with **RMSE : 0.08269**, **0.08467** and **0.09388** respectively. Left to right per row: ground truth, **randomly** sampled, reconstructed image and spectral plot of reconstructed pixel of pure paraffin of size $51 \times 51 \times 208$ at 101^{st} spectral band. See supp. mat. for video results.

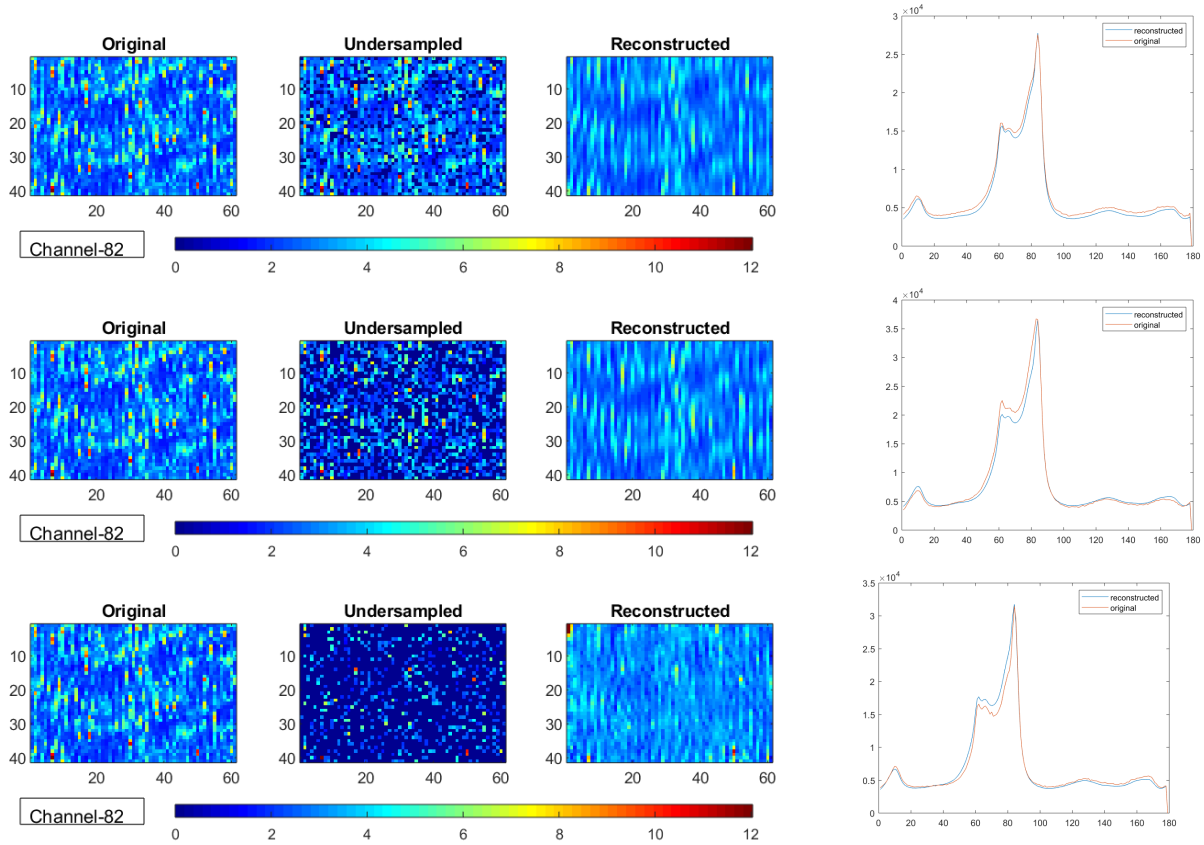
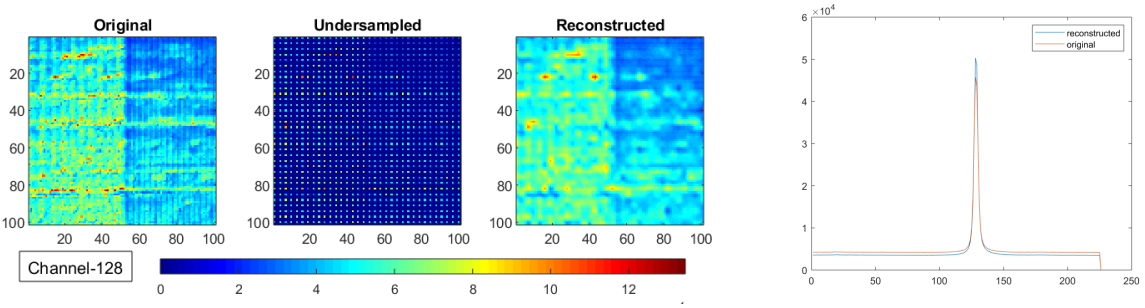


Figure 3: Top to bottom: 80%, 50%, 20% sampling with **RMSE : 0.2547**, **0.2686** and **0.3359** respectively. Left to right per row: ground truth, **randomly** sampled, reconstructed image and spectral plot of reconstructed pixel of Si + GaN of size $41 \times 61 \times 179$ at 82^{th} spectral band. See supp. mat. for video results.



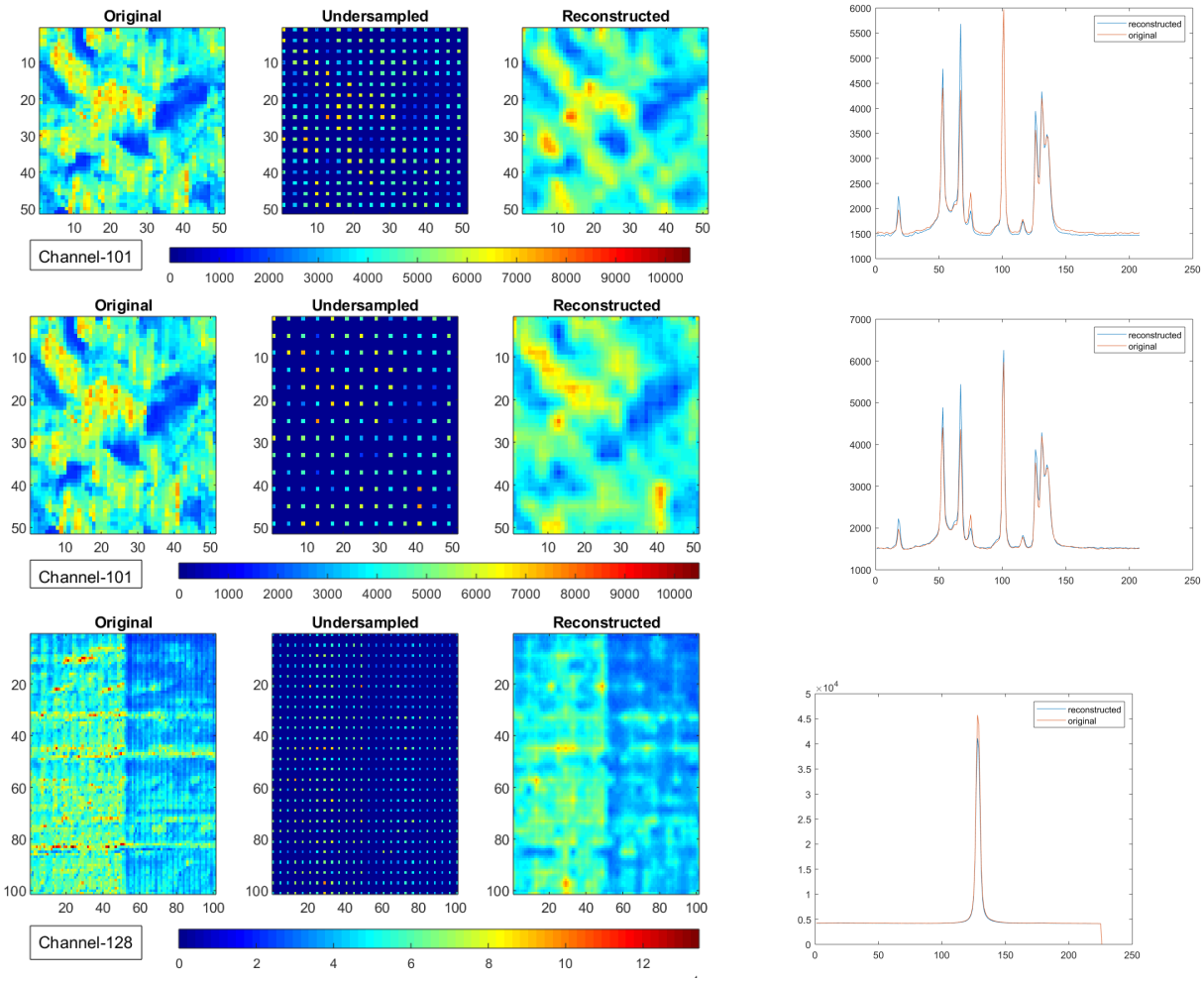
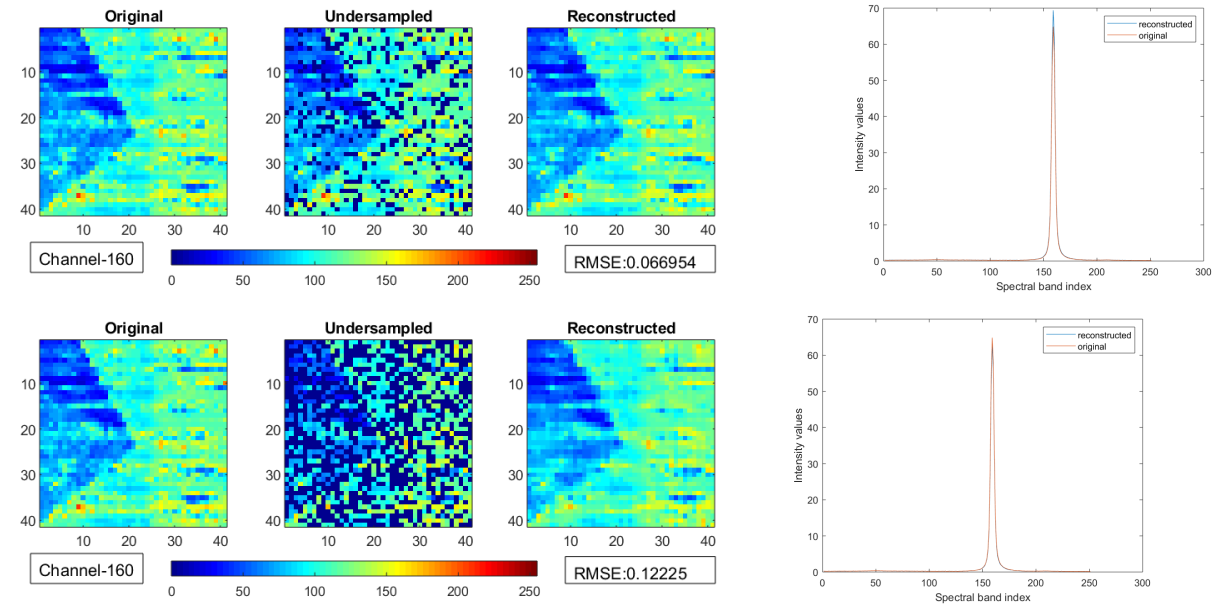


Figure 4: Top to bottom : (a) pure silicon of size $101 \times 101 \times 226$ at 128^{th} spectral band with **RMSE : 0.152** (b) pure paraffin of size $51 \times 51 \times 208$ at 101^{st} spectral band with **RMSE : 0.103**. Left to right per row: ground truth, **structurally** sampled with 4^{th}, , 5^{th} pixel in both directions, reconstructed image and corresponding spectral plots at s missing pixel. See supp. mat. for video results.

3.5 Gaussian Mixture Models



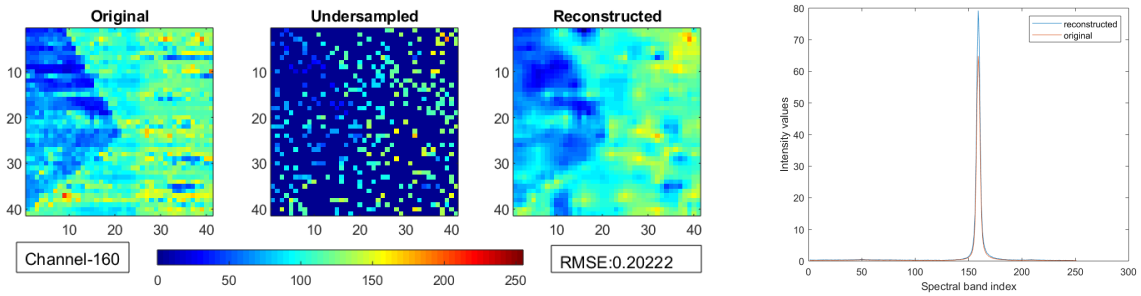


Figure 5: Top to bottom : 80%, 50%, 20% sampling with **RMSE : 0.067**, **0.122** and **0.202** respectively. Left to right per row: ground truth, **randomly** sampled, reconstructed Raman spectral image and spectral plot of reconstructed pixel of pure silicon of size $41 \times 41 \times 251$ at 128^{th} spectral band. See supp. mat. for video results.

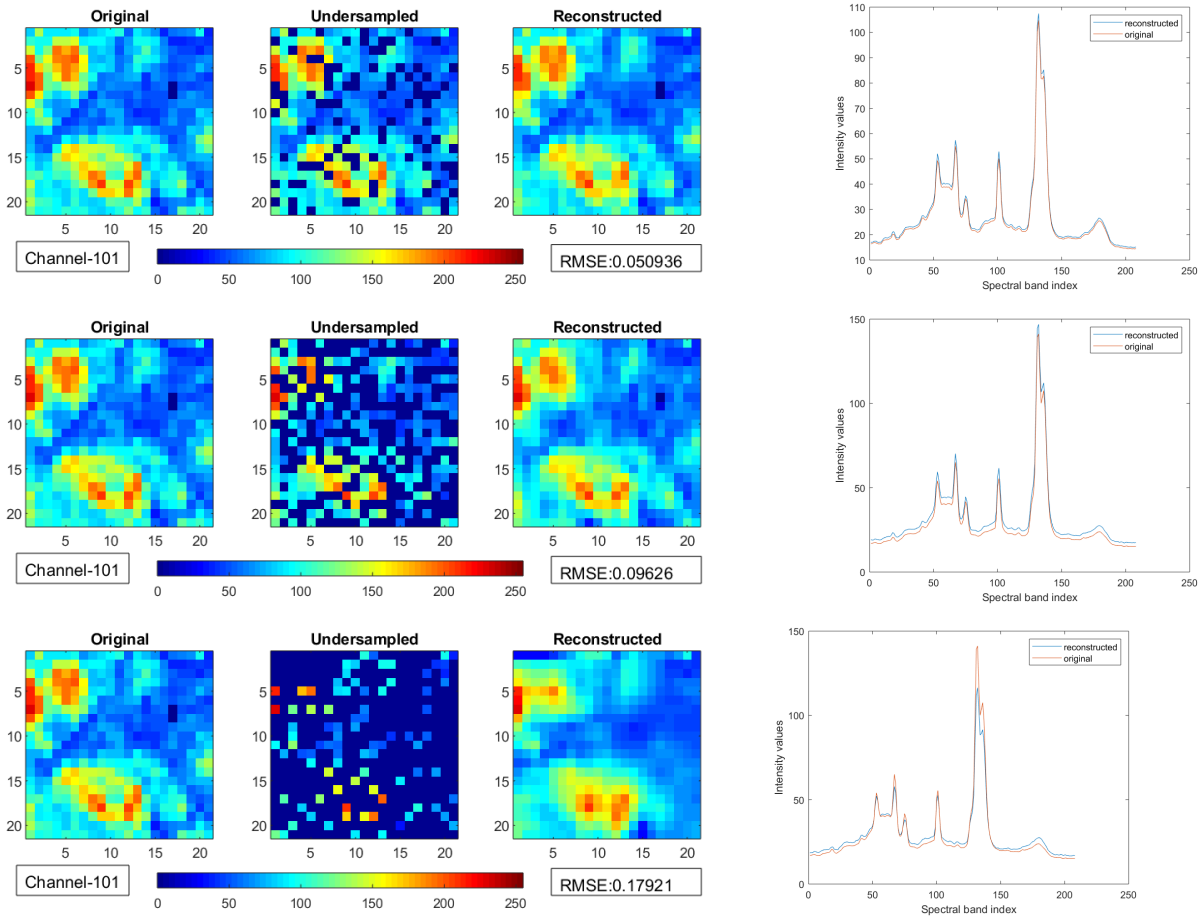


Figure 6: Top to bottom: 80%, 50%, 20% sampling with **RMSE : 0.051**, **0.096** and **0.179** respectively. Left to right per row: ground truth, **randomly** sampled, reconstructed image and spectral plot of reconstructed pixel of pure paraffin of size $21 \times 21 \times 208$ at 101^{st} spectral band. See supp. mat. for video results.

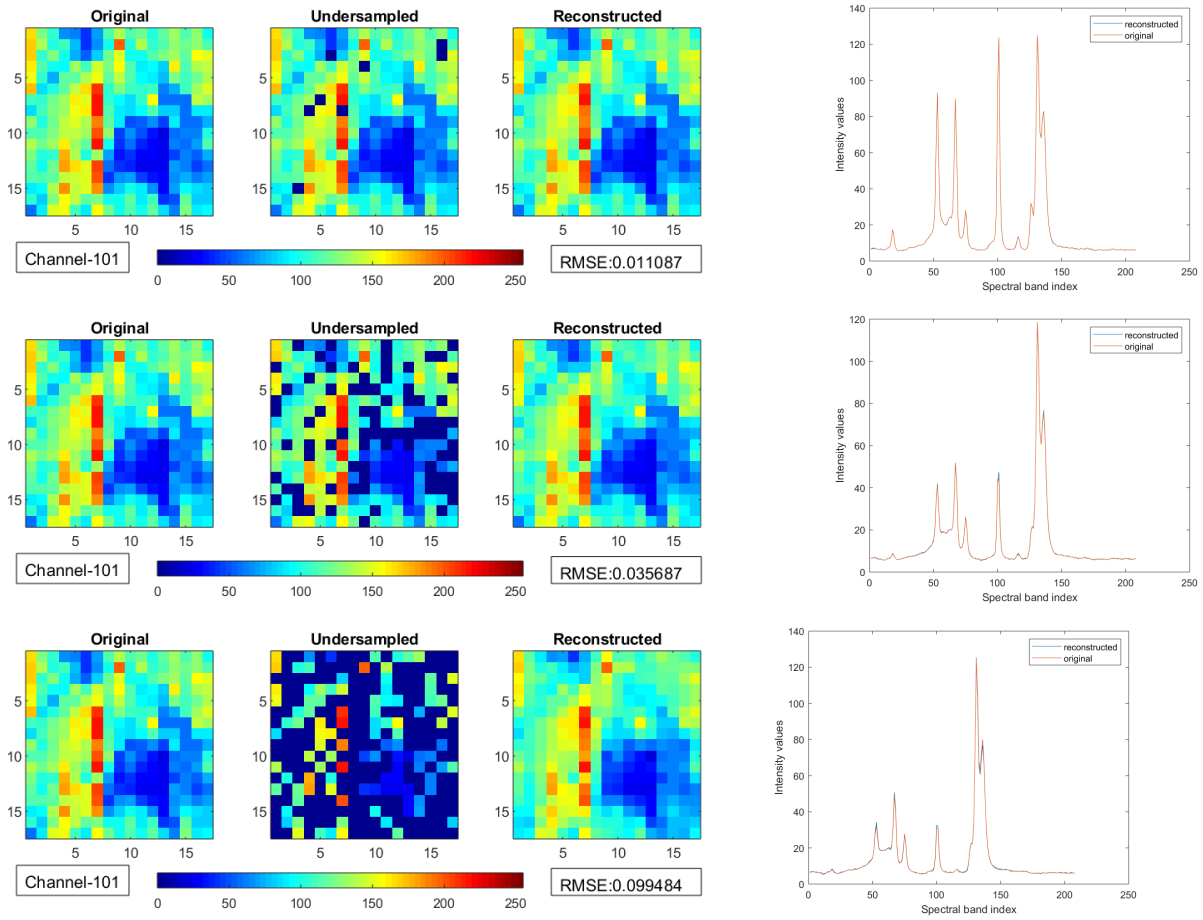
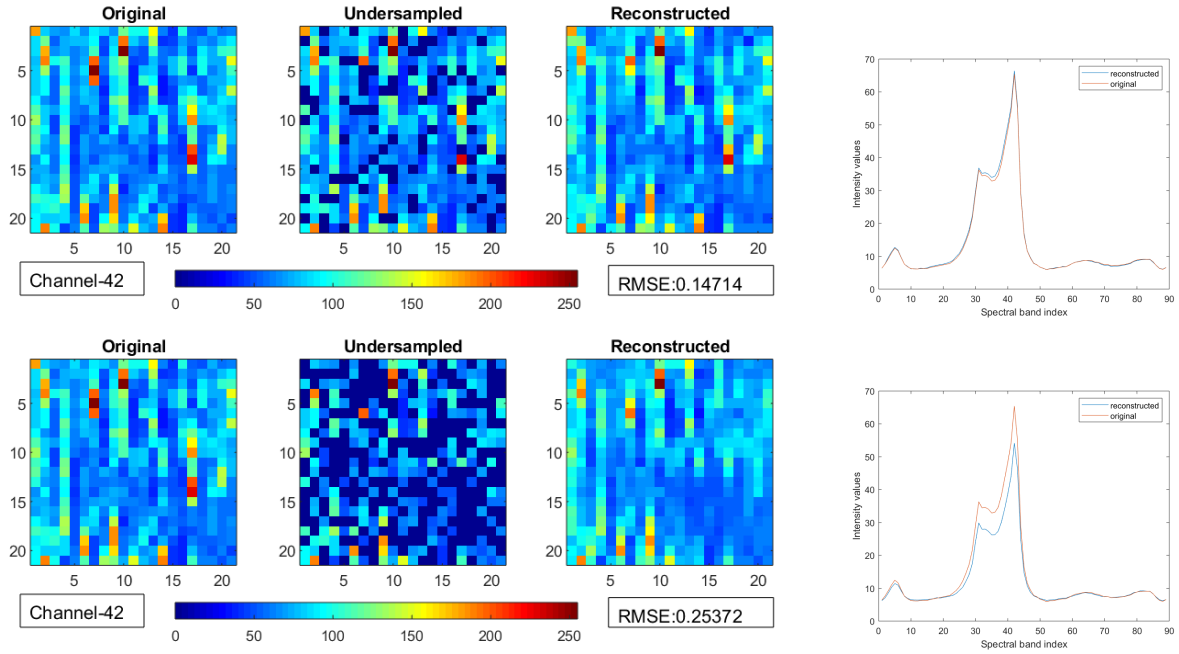


Figure 7: Top to bottom: 80%, 50%, 20% sampling with **RMSE : 0.011**, **0.036** and **0.099** respectively. Left to right per row: ground truth, **randomly** sampled, reconstructed image and spectral plot of reconstructed pixel of pure paraffin of size $21 \times 21 \times 208$ at 101^{st} spectral band **learnt offline**. See supp. mat. for video results.



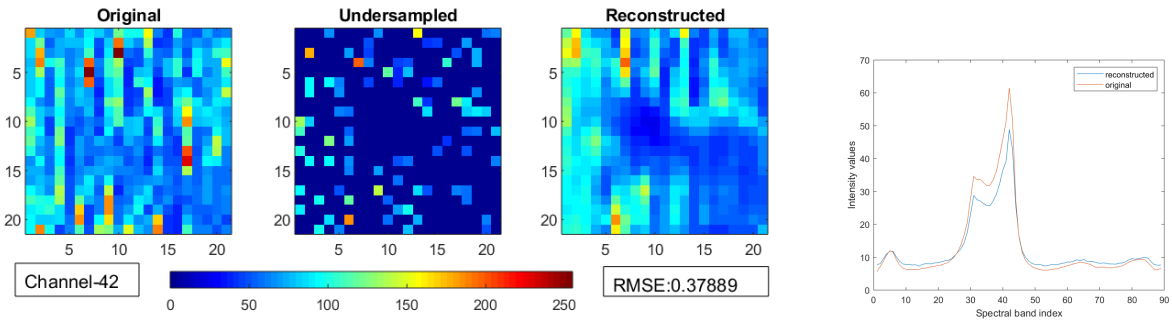


Figure 8: Top to bottom: 80%, 50%, 20% sampling with **RMSE : 0.147**, **0.253** and **0.378** respectively. Left to right per row: ground truth, **randomly** sampled, reconstructed image and spectral plot of reconstructed pixel of Si + GaN of size $21 \times 21 \times 89$ at 41^{st} spectral band. See supp. mat. for video results.

3.6 Comments

The above method can also be incorporated for pixel undersampling in the spectral dimension, so that the entire range of wavelengths for a pixel need not be measured. This will further bring down acquisition times than shown above, with almost no loss of reconstruction quality.

We focus here on approaches related to blind compressed sensing since a set of representative images for training is difficult to find. We also see above that the GMMs work better than conventional dictionary learning methods only for higher sampling rates, whereas for lower sampling rates, the conventional approach generally works better. We know that a good eigenvalue decay for the covariance matrices are required to aid in better reconstruction. We desire Gaussian components in the GMM to be such that they are orthogonal to one another, i.e, the principal components should be such that the first principal component of the first Gaussian is aligned with the last principal component of the last Gaussian and so on. This leads to a better model selection which is at the heart of this problem. Accurate model selection can be achieved with very low sampling rates, given that the energy of the signals is concentrated in the first few principal dimensions. But we are more interested in signal reconstruction, and that works well if the signals are in higher dimensions and the energy is concentrated in the first few dimensions. GMMs work better in the offline setting due to more data being available and better tuning of the mean and the variance matrices. Here, as usual, we see that with lower sampling, the MSE increases for both dictionary learning and GMM approach.

For the problem of super-resolution, we achieve a superior performance of the reconstructed image over the initial interpolation of the undersampled image, showing that the dictionary inference improves upon the image significantly.

The training strategy for GMMs is the same as super-resolution, using a bicubic interplolated image for initial inference phase. In case of lower sampling rates in the data starved situations, the signals might be assigned to erroneous Gaussian components, which leads to a bad initialization. Both this and super-resoultion face the same problems in invertibility of a matrix which is almost always not full rank.

4 Spectral Separation

4.1 Dictionary Learning

We can consider both the variants of the complete image of the sample measured as H or the incomplete undersampled image G is measured. The aim is to recover the underlying spectra os the tissue given the spectra of the paraffin-preserved problem, which might be coupled with the problem of reconstructing the spectra at missing points as well. Here again, we divide the image into overlapping patches of size $p \times p$. Let $\{h_i\}_{i=1}^n$ and $\{g_i\}_{i=1}^n$ of size $p^2 \times a$ and $m \times 1$ respectively be the patches of G and H vectorised into a column. Φ_i is a $m \times p^2$ matrix where each row is a one-hot vector consisting of a 1 at the pixel index which has been measured, i.e, if the $(i,j)^{th}$ pixel corresponds to the k^{th} measurement, then $\Phi_{k,i(p-1)+j} = 1$. We adhere to the random sampling pattern making each Φ_i a row-subsampled version of the canonical basis.

Now, each h_i can be expressed as a sparse linear combination of two dictionary matrices, namely A_p and A_s of size $p^2 \times K_p$ and $p^2 \times K_s$ corresponding to pure paraffin spectra and tissue spectra respectively, viz. $h_i = A_p s_p + A_s s_s$ where $[s_p \; s_s]$ is a concatenated sparse vector of coefficients. In this case, we obtain a pure paraffin sample by making a complete measurement and leverage the strategy described in 3.1 to get a good estimate of A_p. Again, we choose $K_p \ll p^2$ as outlined to minimize the number of degrees of freedom in the problem. As before, keeping in mind the non-negative nature of the data, we impose constraints for s_p, A_s

and s being element-wise non-negative. We now seek to minimize the following objective function :

$$J(A_s, \{s_{pi}\}_{i=1}^n, \{s_{si}\}_{i=1}^n) = \|g_i - \Phi_i A_p s_{pi} - \Phi_i A_s s_{si}\|^2 + \lambda_1 \|s_{pi}\|_1 + \lambda_2 \|s_{si}\|_1, \quad (4)$$
$$\text{such that } A_s \succeq 0, \forall i (s_{pi} \succeq 0, s_{si} \succeq 0)$$
$$\forall j \in \{1, ..., K_s\}, \|A_{s_{.,j}}\|_2^2 = 1$$

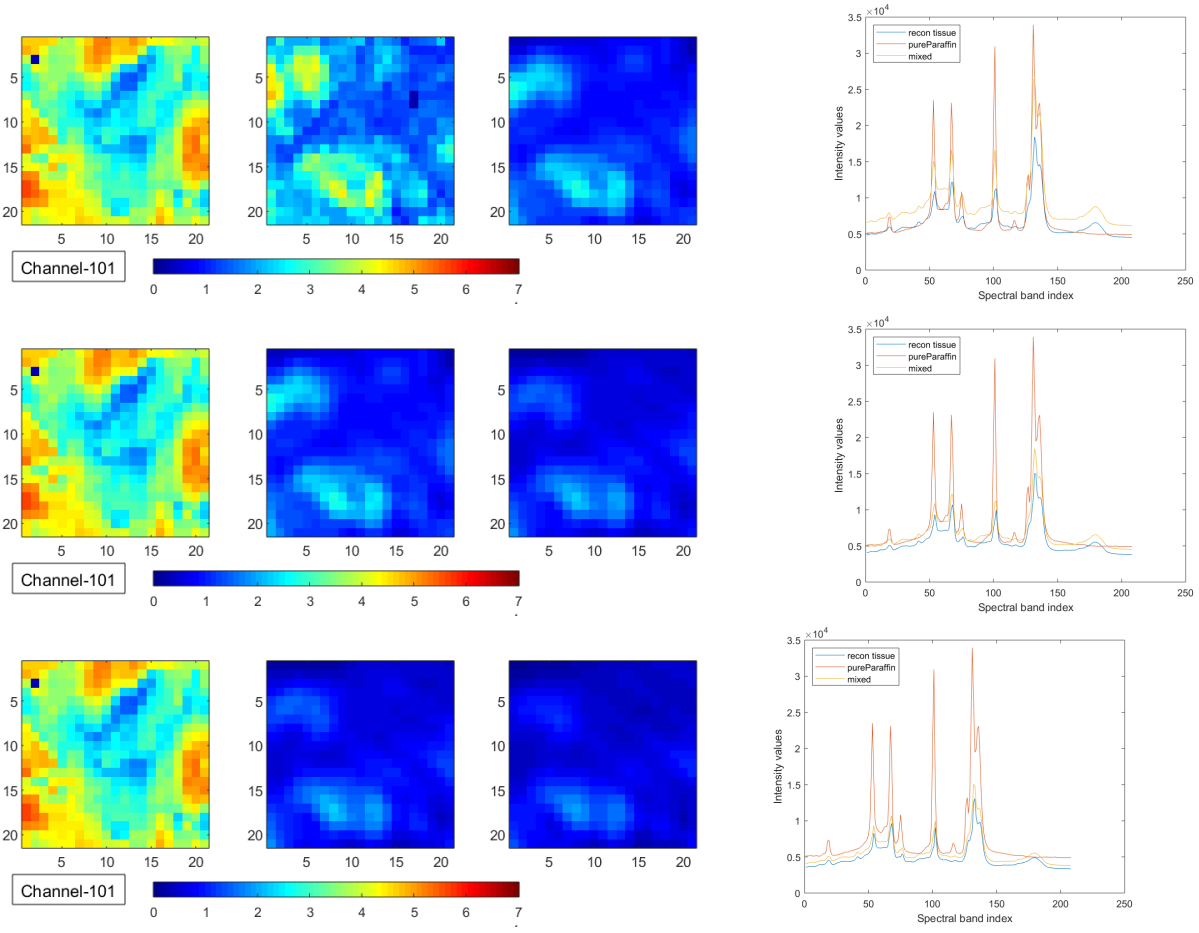
where $\lambda_1 > 0, \lambda_2 > 0$ are sparsity-promoting parameters, $A_{s_{.,j}}$ is the j^{th} column of A_s, 0 represents a zero-valued matrix or vector, and $\succeq$ represents an element-wise 'greater than' inequality. This objective function seeks to solve a different flavour of the problem stated in 3.1. In this work, we choose $K_s \ll p^2$ since a higher K_s increases the number of degrees of freedom, and the size of given Raman spectral sample images is very small.

We implement the minimization of the function in Eqn. 5 using alternating minimization on the dictionary and sparse codes, starting from a random non-negative dictionary. Each step of the minimization is performed using projected gradient descent with adaptive step-size. That is, the stepsize of the gradient descent is adaptively chosen to ensure decrease of the objective function after imposition of all the constraints in Eqn. 5. The procedure is iterated till convergence, which is guaranteed due to the biconvex nature of the objective function. The algorithm is summarized in Alg. 1. Once the dictionary A_s and coefficients $\{s_{si}\}_{i=1}^n$ are inferred, an estimate of the tissue spectrum is obtained by sliding window averaging across patches. Once an initial estimate of G_s is obtained which comprises of pure tissue spectra, we attempt to iteratively keep subtracting any paraffin component that might be present in the tissue spectra estimate. Here again, let g_{si} represent a vectorised version of corresponding patch of G_s of the same size $p \times p$, We again run the same optimization problem with the only change in the observed data utilising the same paraffin dictionary A_p :

$$J(A_s, \{s_{pi}\}_{i=1}^n, \{s_{si}\}_{i=1}^n) = \|g_{si} - \Phi_i A_p s_{pi} - \Phi_i A_s s_{si}\|^2 + \lambda_1 \|s_{pi}\|_1 + \lambda_2 \|s_{si}\|_1, \quad (5)$$
$$\text{such that } A_s \succeq 0, \forall i (s_{pi} \succeq 0, s_{si} \succeq 0)$$
$$\forall j \in \{1, ..., K_s\}, \|A_{s_{.,j}}\|_2^2 = 1$$

with the same constraints that had been used to solve 5.

4.2 Results



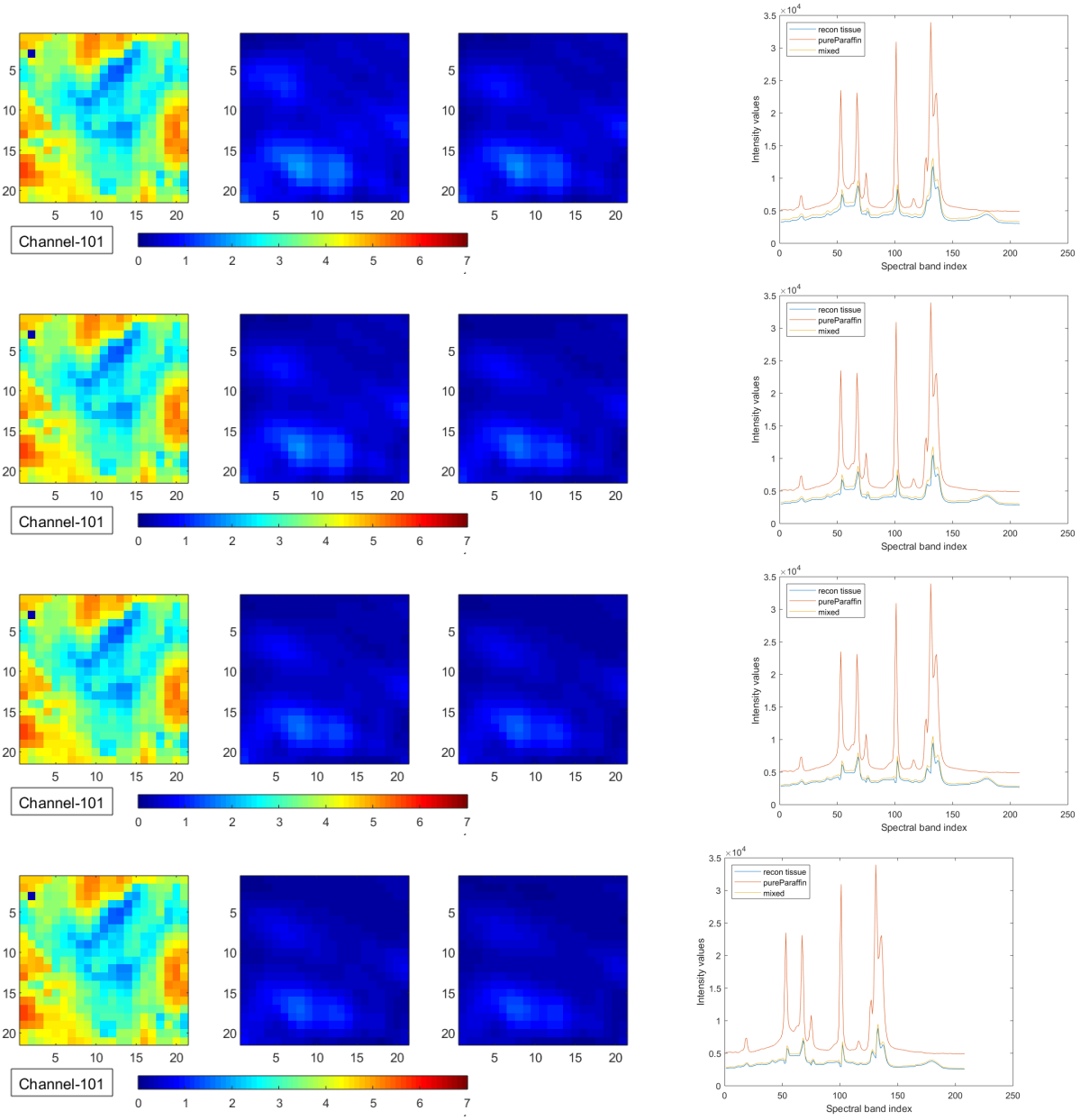


Figure 9: Top to bottom: Each of the 7 estimates of the iterative paraffin dewaxing of given tissue sample of size $21 \times 21 \times 208$ at 101^{st} spectral band along with their spectral plots. Left to Right : pure paraffin, estimate of the previous iteration and current reconstructed tissue spectral estimate.

4.3 Comments

The above results do not look too promising given the nature of the sample spectra. The paraffin spectra overwhelmingly dominates the tissue spectra. In the above case, the iterative process of inferring the tissue dictionary and corresponding coefficients whiles trying to weed out the pure paraffin component also leads to some part of the tissue spectra getting absorbed into the pure paraffin component, which can't be recovered later. Hence, subsequent iterates have lesser and lesser tissue component, albeit more 'separate' from the paraffin component. These are generally obserbed to weaken in magnitude and merge with the baeline after just a few iterations. Also, some highly intense 'stubborn' pure paraffin peaks can never be completely removed. In such a case, increasing the number of paraffin dictionary columns to subtract out more of the paraffin component also doesn't seem to help.

We have also tried ad-hoc approaches for the same because the case being that generally, the tissue signal is weaker. In such a case, when one signal dominates the other, it is generally difficult to perform source separation. Hence, in this case, we tried to identify the peaks and correspondingly tried to bring the peaks corresponding to pure paraffin in the paraffin & tissue sample to the same scale as the paraffin peaks in the pure paraffin sample. This is again achieved by a least squares minimization approach, thereafter which we

subtract the pure paraffin spectrum from the sample spectra with non-negativity constraints imposed in an effort to obtain the pure tissue spectra.

5 Future Work

One important line of future work will involve development of such a hardware prototype. The reconstruction algorithm can also be further improved using deep learning techniques [23] or tensor factorizations [24]. Such a deep learning framework can also be utilized for the above spectral separation problem.

References

[1] E. Candes and M. Wakin, "An introduction to compressive sampling," *IEEE Sig. Proc. Mag.*, 2008.

[2] J. Romberg, "Imaging via compressive sampling," *IEEE Sig. Proc. Mag.*, 2008.

[3] M. Schaeberle et al, "Raman chemical imaging: histopathology of inclusions in human breast tissue," *Anal. Chem*, 1996.

[4] D. Wetzel and S. LeVine, "Imaging molecular chemistry with infrared microscopy," *Science*, 1999.

[5] Y. Zhang Y, H. Hong, and W. Cai, "Imaging with raman spectroscopy," *Current Pharmaceutical Biotechnology*, 2010.

[6] S. Kawata et al, "Near-field optics and spectroscopy for molecular nano-imaging," *Sci. Prog*, 2004.

[7] J. Ling, S. Weitman, V. Miller, R. Moore, and A. Bovik, "Direct raman imaging techniques for study of the subcellular distribution of a drug," *Applied Optics*, 2002.

[8] D. Lauwers, Ph. Brondeel, L. Moens, and P. Vandenabeele, "In situ raman mapping of art objects," *Philosophical Transactions of the Royal Society A*, 2016.

[9] L. Haskin et al, "Raman spectroscopy for mineral identification and quantificafionor in sire planetary surface analysis: A point count method," *Journal of Geophysical Research*, 1997.

[10] A. Tiberj and J. Camassel, *Raman Imaging in Semiconductor Physics: Applications to Microelectronic Materials and Devices*, 2012.

[11] Chandrasekhara Venkata Raman and Kariamanikkam Srinivasa Krishnan, "A new type of secondary radiation," *Nature*, vol. 121, no. 3048, pp. 501, 1928.

[12] A. Rajwade, D. Kittle, T.-H. Tsai, D. Brady, and L. Carin, "Coded hyperspectral imaging and blind compressive sensing," *SIAM J. Imaging Sciences*, 2013.

[13] Guoshen Yu and Guillermo Sapiro, "Statistical compressed sensing of gaussian mixture models," *IEEE Transactions on Signal Processing*, vol. 59, no. 12, pp. 5842–5858, 2011.

[14] Peter PJ Caspers, *In vivo skin characterization by confocal Raman microspectroscopy*, 2003.

[15] Monika Gniadecka, HC Wulf, N Nymark Mortensen, O Faurskov Nielsen, and Daniel H Christensen, "Diagnosis of basal cell carcinoma by raman spectroscopy," *Journal of Raman spectroscopy*, vol. 28, no. 2-3, pp. 125–129, 1997.

[16] Hendrik P Buschman, Jason T Motz, Geurt Deinum, Tjeerd J Römer, Maryann Fitzmaurice, John R Kramer, Arnoud van der Laarse, Albert V Bruschke, and Michael S Feld, "Diagnosis of human coronary atherosclerosis by morphology-based raman spectroscopy," *Cardiovascular Pathology*, vol. 10, no. 2, pp. 59–68, 2001.

[17] Eoghan Ó Faoláin, Mary B Hunter, Joe M Byrne, Peter Kelehan, Helen A Lambkin, Hugh J Byrne, and Fiona M Lyng, "Raman spectroscopic evaluation of efficacy of current paraffin wax section dewaxing agents," *Journal of Histochemistry & Cytochemistry*, vol. 53, no. 1, pp. 121–129, 2005.

[18] D. Kittle, K. Choi, A. Wagadarikar, and D. J. Brady, "Multiframe image estimation for coded aperture snapshot spectral imagers," *Appl. Opt.*, 2010.

[19] Cyril Gobinet, Valeriu Vrabie, Michel Manfait, and Olivier Piot, "Preprocessing methods of raman spectra for source extraction on biomedical samples: application on paraffin-embedded skin biopsies," *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 5, pp. 1371–1382, 2009.

[20] Cyril Gobinet, Valeriu Vrabie, Ali Tfayli, Olivier Piot, Regis Huez, and Michel Manfait, "Pre-processing and source separation methods for raman spectra analysis of biomedical samples," in *2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 2007, pp. 6207–6210.

[21] Cyril Gobinet, David Sebiskveradze, Valeriu Vrabie, Ali Tfayli, Olivier Piot, and Michel Manfait, "Digital dewaxing of raman spectral images of paraffin-embedded human skin biopsies based on ica and ncls," in *2008 16th European Signal Processing Conference*. IEEE, 2008, pp. 1–5.

[22] P. O. Hoyer, "Non-negative sparse coding," in *IEEE Workshop on Neural Networks for Signal Processing*, 2002.

[23] A. Lucas, M. Iliadis, R. Molina, and A. Katsaggelos, "Using deep neural networks for inverse problems in imaging," *IEEE Signal Processing Magazine*, 2018.

[24] P. Zhou, C. Lu, Z. Lin, and C. Zhang, "Tensor factorization for low-rank tensor completion," *IEEE Transactions on Image Processing*, 2018.