Multi-resolution Reconstruction Algorithm for Phase Retrieval in X-ray Crystallography

Jhon Angarita♦, Samuel Pinilla†, Hans Garcia†, Henry Arguello⋆
♦*Dept. of Physics*, †*Dept. of Electrical Engineering*, ⋆*Dept. of Computer Science*
Universidad Industrial de Santander, Colombia

***Abstract*—Phase Retrieval (PR) in X-ray Crystallography (XC) is an inverse problem that consists on recovering an image from phaseless data. Recently, it has been shown that an image in XC can be sparsely represented in the Fourier domain. This fact implies that the number of required measurements to retrieve the phase in XC is determined by the sparsity, which is much smaller than the size of the image. However, the computational complexity to retrieve the phase still depends on the image size, implying more time to solve this problem in XC. Therefore, this work proposes a reconstruction algorithm that exploits the sparsity of the image by grouping sets of pixels of its sparse representation, called super-pixels, in order to reduce the total number of unknowns in the inverse problem. The proposed recovery methodology leads to a reduction in time of at least 80% and improves the reconstruction quality in up to 6% in terms of the Structural Similarity Index Measure (SSIM) compared to state-of-art counterparts.**

I. Introduction

Phase retrieval (PR) is an inverse problem which consists on recovering an image from phaseless measurements. This problem appears in a wide range of engineering and physics applications. To name a few, optical imaging [1] , astronomy [2], coherent diffraction imaging [3] and X-ray crystallography (XC) [4]. The latter is the object of study in this work. XC is an experimental tecnique to determine the three-dimensional (3D) structure of a crystal [5]. Specifically, the crystal can be uniquely identified by means of the phase of its diffraction patterns [6]. Further, because the phase of the diffraction patterns cannot be directly measured, different methods in the state-of-the-art have been developed to recover it from the intensity of the diffraction patterns [7], [8].

The authors of [8] have shown that the phase can be more efficiently recovered (up to a global unimodular constant) from measurements acquired with architectures including an optical element known as coded aperture, which is assumed to be a random matrix. Mathematically, the PR problem in XC from coded diffraction patterns consists on recovering an image $\mathbf{x} \in \mathbb{C}^n$ from the phaseless data $\mathbf{y} = |\langle \mathbf{a}_i, \mathbf{x} \rangle|^2$, where $\mathbf{a}_i \in \mathbb{C}^n$ for $i = 1, \cdots, \hat{m}$ are the sampling vectors.

In section II, we will show that the image in XC can be sparsely represented in the Fourier domain, that is, $\|\boldsymbol{\theta}\|_0 = k \ll n$ where $\boldsymbol{\theta} = \mathbf{F}\mathbf{x}$, and $\|\cdot\|_0$ is the ℓ_0-pseudonorm. This fact implies that the number of required coded measurements to retrieve the phase in XC is determined by the sparsity, which is much smaller than the size of the image. However, the computational complexity to solve the PR problem still depends on the image size n, implying long reconstruction times for XC. In fact, supposing a signal with size greater than n, with sparsity k, the computational cost of its reconstruction will be higher compared with that of a signal of size n.

This paper proposes a reconstruction algorithm that exploits the sparsity k of the image by grouping sets of pixels, called super-pixels, in the sparse representation to substantially reduce the total number of unknowns in the inverse problem. These super-pixels are determined from the acquired data by grouping zero coefficients of the sparse representation of the image. This fact leads to an improvement in the reconstruction quality and a reduction in the reconstruction time since the sparse representation of a high-spatial-resolution image is mapped to a lower spatial dimension reducing the number of unknowns. The super-pixel concept has been previously used to compress images as in simple linear iterative clustering [9] and compressive spectral imaging [10]. Also, it is worth to mention that these state-of-the-art works operate directly over the scene, in contrast, our methodology focuses on processing the sparse representation making this paper first in its kind. Further, a super-pixel approach has not been up to date extended to the PR problem.

Numerical simulations show that the super-pixel methodology attains a reconstruction time reduction of at least 80% compared with traditional sparse-based reconstruction methods. Also, we validate through simulations that the proposed approach improves the reconstruction quality in up to 6% in terms of the Structural Similarity Index Measure (SSIM) compared with sparse-based reconstruction methodologies.

Notation. In the remainder, we denote $(\cdot)^T$, $\overline{(\cdot)}$ and $(\cdot)^H$ as the transpose, conjugate, and conjugate transpose operations, respectively. The ceiling operation $\lfloor \cdot \rfloor$ returns the greater integer, smaller than or equal to the given number. For vectors $\|\mathbf{x}\|_p$ is the usual ℓ_p norm. Also, $\|\cdot\|_0$ denotes the ℓ_0 pseudo-norm counting the number of nonzero entries. Further, we will denote $\mathbb{E}[\cdot]$ as the expected value and operator $card(\mathcal{B})$ represents cardinality of $\mathcal{B}$.

II. Problem formulation

X-ray crystallography is a technique used in material analysis to determine the 3D structure of a crystal [6]. A recent proposed acquisition system for XC is shown in Fig. 1, where the object of interest is illuminated by an X-ray source and modulated by a coded aperture [11]. As a result, a coded diffraction pattern (CDP) is produced.

Notice that, if we change the spatial configuration of the coded aperture, this acquisition system allows to acquire multiple projections of the scene. Mathematically, the acquisition

978-9-0827-9703-9/19/$31.00 ©2019 IEEE

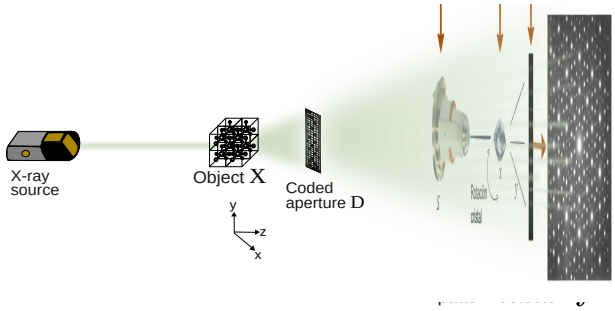

Fig. 1: Illustrative configuration to acquire coded diffraction patterns in XC.

process of the coded system illustrated in Fig. 1 can be modeled as

$$\mathbf{y}_\ell = |\mathbf{F}\mathbf{D}_\ell\mathbf{x}|^2, \ell = 1, \cdots, L, \tag{1}$$

where $\mathbf{F} \in \mathbb{C}^{n\times n}$ is the discrete Fourier transform matrix, $|\cdot|^2$ denotes the element-wise absolute-squared value, $\mathbf{x} \in \mathbb{C}^n$ is the image to recover, $\mathbf{D}_\ell$ is a diagonal matrix that models the coded aperture at the ℓ-th projection, and L is the number of projections. Further, defining $\mathbf{y} = [\mathbf{y}_1^T, \cdots, \mathbf{y}_L^T]^T \in \mathbb{R}^{m=nL}$ and the matrix $\mathbf{A} = [\mathbf{D}_1\mathbf{F}, \cdots, \mathbf{D}_L\mathbf{F}]^T$, we can rewrite the quadratic model in (1) as

$$\mathbf{y} = |\mathbf{A}\mathbf{x}|^2, \tag{2}$$

where each row of $\mathbf{A}$ is given by $\mathbf{a}_i = \overline{\mathbf{D}}_{r_i}\mathbf{f}_{u_i}$ with $r_i = \lfloor (i-1)/n \rfloor + 1$, $u_i = [(i-1) \mod n] + 1$, and $\mathbf{f}_{u_i}$ the rows of $\mathbf{F}$, for $i = 1 \cdots nL$.

Fig. 2 illustrates typical observed $|\mathbf{F}\mathbf{x}|^2$ in XC. In particular, the magnitude of the Fourier transform of a crystal of *cytidine* is presented. The white and black pixels represent low and high-intensity elements, respectively. Thus, it shows that the image of interest $\mathbf{x}$ in XC can be sparsely represented in the Fourier domain, that is, $\mathbf{F}\mathbf{x} = \boldsymbol{\theta}$ where $\|\boldsymbol{\theta}\|_0 = k \ll n$.

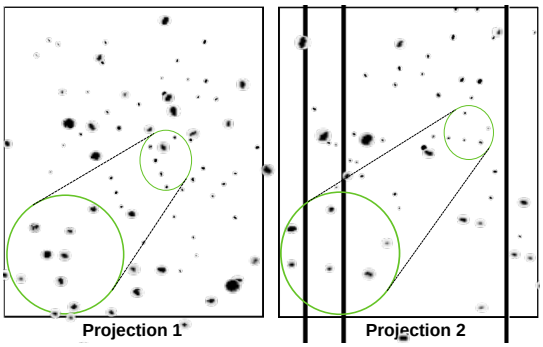

Fig. 2: Real $|\mathbf{F}\mathbf{x}|^2$ in XC, where white and black pixels represent low and high intensity elements, respectively.

Considering this previous observation, and the coded diffraction model in (2), the image $\mathbf{x}$ in XC can be reconstructed as

$$\begin{aligned} \min_{\boldsymbol{\theta}\in\mathbb{C}^n} \quad & f(\boldsymbol{\theta}) = \|\sqrt{\mathbf{y}} - |\mathbf{A}\mathbf{F}^H\boldsymbol{\theta}|\|_2^2 \\ \text{subject to} \quad & \|\boldsymbol{\theta}\|_0 \le k, \end{aligned} \tag{3}$$

where $\sqrt{\cdot}$ denotes the element-wise squared root of a vector. Observe that, the image of interest can then be estimated as $\mathbf{x} = \mathbf{F}^H\boldsymbol{\theta}$. However, the computational complexity to solve (3) still depends on the image size n, implying long reconstruction times.

To alleviate the time complexity of PR in XC we propose to design a decimation matrix $\mathbf{M} \in \mathbb{R}^{r\times n}$ with $r < n$, to build super-pixels which are groups of zero coefficients of $\boldsymbol{\theta}$. In fact, using the matrix $\mathbf{M}$, the sparse representation of a high-spatial-resolution image is mapped to one of lower resolution, reducing the total number of unknowns and therefore the reconstruction time. Mathematically, introducing the effect of the decimation matrix $\mathbf{M}$, we can reformulate (3) as

$$\begin{aligned} \min_{\boldsymbol{\theta}\in\mathbb{C}^n} \quad & f(\boldsymbol{\theta}) = \|\sqrt{\mathbf{y}} - |\mathbf{A}\mathbf{F}^H\mathbf{M}^T\mathbf{M}\boldsymbol{\theta}|\|_2^2 \\ \text{subject to} \quad & \|\boldsymbol{\theta}\|_0 \le k. \end{aligned} \tag{4}$$

Notice that, the surrogate optimization problem in (4) is equivalent to (3), because the vector $\boldsymbol{\theta}$ is assumed to be sparse, and $\mathbf{M}$ is forming groups of zero coefficients which implies that $\mathbf{F}^H\mathbf{M}^T\mathbf{M}\boldsymbol{\theta} = \mathbf{F}^H\boldsymbol{\theta}$, preserving the sparsity level. Thus, defining $\mathbf{M}\boldsymbol{\theta} = \hat{\boldsymbol{\theta}} \in \mathbb{C}^r$ and $\hat{\mathbf{A}} = \mathbf{A}\mathbf{F}^H\mathbf{M}^T$, we obtain that (4) can be rewritten as

$$\begin{aligned} \min_{\hat{\boldsymbol{\theta}}\in\mathbb{C}^r} \quad & f(\hat{\boldsymbol{\theta}}) = \|\sqrt{\mathbf{y}} - |\hat{\mathbf{A}}\hat{\boldsymbol{\theta}}|\|_2^2 \\ \text{subject to} \quad & \|\hat{\boldsymbol{\theta}}\|_0 \le k. \end{aligned} \tag{5}$$

Considering (5), the following section develops a strategy to design the matrix $\mathbf{M}$. We call our decimation strategy multi-resolution because the non-zero coefficients are not decimated and the zero coefficients are decimated in an irregular fashion, leading to super-pixels of different sizes, as it will be explained in Section III.

III. Multi-resolution strategy

This section describes the design principles of the decimation matrix $\mathbf{M}$ in two steps as follows:

(a) Non-zero coefficient estimation: we statistically estimate the non-zero coefficients of $\boldsymbol{\theta}$, from the coded measurements in (2). In fact, the coded apertures can be optimized to attain the best estimation of the non-zero coefficients of $\boldsymbol{\theta}$.

(b) Building super-pixels: From the outcome of the first step, rectangular super-pixels of different sizes can be formed by grouping neighboring zero coefficients of $\boldsymbol{\theta}$ into super-pixels.

A. Non-zero coefficients estimation

In order to estimate the non-zero coefficients of $\boldsymbol{\theta}$ we extend the strategy developed in [12] to CDP. But first, we specify that the entries of each matrix $\mathbf{D}_\ell$ are $i.i.d.$ copies of an admissible random variable d defined as follows.

Definition 1 (*Admissible random variable*)**.** *A discrete random variable obeying $|d| \le 1$ almost sure, is said to be admissible.*

Considering that $\mathbf{D}_\ell$ for $\ell = 1, \cdots, L$, are random matrices, define $Z_{r,s} := y_r|b_{r,s}|^2, 1 \le r \le m, 1 \le s \le n$, where $b_{r,s}$ is the element at row r and column s of a matrix $\mathbf{B} = \mathbf{A}\mathbf{F}^H$. Then, we have that

$$\mathbb{E}[Z_{q,p}] = c_1(n-1)\|\mathbf{x}\|_2^2 + c_2|\boldsymbol{\theta}_p|^2, \tag{6}$$

where $c_1, c_2 > 0$ are constants depending on d. Also, given the fact that $\boldsymbol{\theta}$ is sparse, $\boldsymbol{\theta}_p \neq 0$ or $\boldsymbol{\theta}_p = 0$. Then it is clear from (6), as long as the constant c_2 is sufficiently large, the non-zero coefficients of $\boldsymbol{\theta}$ can be identified in this way. Thus, the k non-zero coefficients of $\boldsymbol{\theta}$ are the k largest indices of $Z_{q,p}$. In fact, we are interested in optimizing the set of coded apertures in order to obtain the best estimation of the non-zero coefficients of $\boldsymbol{\theta}$. Specifically, the following lemma theoretically proves (6) and establishes that an admissible random variable d satisfying $\mathbb{E}[d], \mathbb{E}[\overline{d}] \neq 0$ attains the best performance.

Lemma 1. *Considering the coded measurements* $\mathbf{y}$ *in* (2) *and Definition 1, an admissible random variable* d *satisfying* $\mathbb{E}[d], \mathbb{E}[\overline{d}] \neq 0$ *attains the best performance for estimating the non-zero coefficients of* $\boldsymbol{\theta}$.

Proof. See Appendix A. □

B. Building super-pixels

In this section, we describe how to group sets of zero coefficients of $\boldsymbol{\theta}$ into super-pixels of different sizes. More precisely, once the non-zero coefficients are estimated as explained in Section III-A, we are able to decimate $\boldsymbol{\theta}$. For that purpose, we modified a decimation algorithm introduced in [10]. The design details of the decimation matrix are summarized in Algorithm 1. Notice that a matrix version $\tilde{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$ is required in Line 3 to create super-pixels. Then, Algorithm 1 generates a random point $(\hat{i}, \hat{j})$ in Line 8 and creates a set $\mathcal{B}$ that contains the spatial coordinates of a $2^\eta \times 2^\eta$ super-pixel whose top-left corner is the point $(\hat{i}, \hat{j})$. Therefore, Algorithm 1 calculates the average of the normalized intensities of $\tilde{\boldsymbol{\theta}}_{\mathcal{B}}$ in order to determine the similarity with each point of $\tilde{\boldsymbol{\theta}}$ contained in $\mathcal{B}$, considering as metric its mean squared error, as shown in Line 11. Thus, this spatial similarity is used to create the super-pixels of size $2^\eta \times 2^\eta$, as long as, the maximum mean squared error is smaller than the fixed threshold σ. Further, it is important to remark that the rows of $\mathbf{M}$ are normalized by the cardinality of $\mathcal{B}$, as shown in Line 13. This decimation approach allows to group the zero coefficients of the sparse representation, which leads to reducing the number of unknowns.

Figure 3 shows an example of a decimation grid of the sparse image $\tilde{\boldsymbol{\theta}}$. Note that, the non-zero coefficients are not decimated.

IV. Reconstruction Algorithm

This section presents the proposed reconstruction algorithm that exploits the sparsity of $\boldsymbol{\theta}$ in order to substantially reduce the total number of unknowns in XC. More precisely, considering the fact that state-of-the-art algorithms cannot be directly applied to solve (5), then a modified version of the Sparse Phase Retrieval via Smoothing Function (SPRSF) introduced in [13] is presented. Specifically, the proposed method is summarized in Algorithm 2.

Algorithm 2 starts with a proper initialization, which is summarized from Line 2 to 7. Note that in Line 2, Algorithm 2 estimates the non-zero coefficients of $\boldsymbol{\theta}$ following the methodology explained in Section III-A. In this way, the matrix $\mathbf{M}$ can be created, as shown in Line 3 which allows to calculate the

Algorithm 1 Algorithm to determine the matrix $\mathbf{M}$

1: **Input:** Sparse signal $\boldsymbol{\theta}$, threshold parameter σ, super-pixel size parameter η.
2: Let $\mathbf{s} = [2^\eta, 2^{\eta-1}, \dots, 1]$ be super-pixel sizes and initialize $\alpha = 0$, $\rho = 0$.
3: $\tilde{\boldsymbol{\theta}} \in \mathbb{R}^{N_1 \times N_2} \leftarrow$ matrix version of $\boldsymbol{\theta}$
4: Define the set of coordinate points as $\mathcal{I} = \{(i,j) | \forall i,j \in \{1, \cdots, N_1\}, \{1, \cdots, N_2\}\}$
5: **while** $card(\mathcal{I}) > 0$ **do**
6: $\quad \hat{\mathcal{I}} = \mathcal{I}$ Generate a new set of available points
7: $\quad$ **while** $card(\hat{\mathcal{I}}) > 0$ **do**
8: $\quad\quad$ Generate a random point $(\hat{i}, \hat{j}) \in \hat{\mathcal{I}}$
9: $\quad\quad$ Build a super-pixel $\mathcal{B} = \{(i,j) | i = [\hat{i}, \dots, \hat{i} + \mathbf{s}(\alpha)], j = [\hat{j}, \dots, \hat{j} + \mathbf{s}(\alpha)]\}$ of size $\mathbf{s}(\alpha)$ and a new top-left corner point $(\hat{i}, \hat{j}) \in \hat{\mathcal{I}}$
10: $\quad\quad \mathbf{p} \leftarrow E\{\tilde{\boldsymbol{\theta}}_{\mathcal{B}}\}$ ▷ Calculate the average normalized of $\tilde{\boldsymbol{\theta}}$ in $\mathcal{B}$
11: $\quad\quad$ **if** $max(MSE(\mathbf{p}, \tilde{\boldsymbol{\theta}}_{\mathcal{B}})) < \sigma$ **then**
12: $\quad\quad\quad \boldsymbol{\Gamma} \leftarrow \mathbf{0}_{N_1 \times N_2}$
13: $\quad\quad\quad \boldsymbol{\Gamma}_{(i,j)} \leftarrow \frac{1}{card(\mathcal{B})}$ for $(i,j) \in \mathcal{B}$ ▷ Super-pixel indicator
14: $\quad\quad\quad (\mathbf{M})_\rho = vec(\boldsymbol{\Gamma})$ ▷ Assigning new row to $\mathbf{M}$
15: $\quad\quad\quad \rho = \rho + 1$ ▷ Update super-pixels counter
16: $\quad\quad\quad \mathcal{I} = \mathcal{I} - \mathcal{B}$ ▷ Update available points
17: $\quad\quad\quad \hat{\mathcal{I}} = \hat{\mathcal{I}} - \mathcal{B}$ ▷ Update eligible top-left points
18: $\quad\quad$ **else**
19: $\quad\quad\quad \hat{\mathcal{I}} = \hat{\mathcal{I}} - (\hat{i}, \hat{j})$ ▷ Remove the corner point $(\hat{i}, \hat{j}) \in \hat{\mathcal{I}}$
20: $\quad \alpha = \alpha + 1$ ▷ Change super-pixel size index
21: **Output:** matrix $\mathbf{M}$

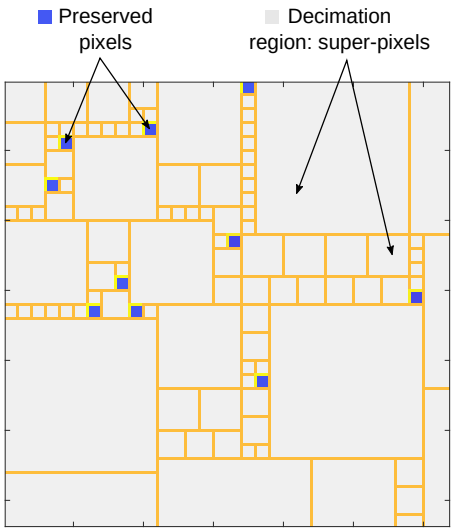

Fig. 3: Illustration of the multi-resolution strategy by grouping sets of zero-pixels of the sparse representation. Non-zero pixels are not grouped.

decimated sparse version $\boldsymbol{\theta}_e$ in Line 7. Lastly, under successive refinements based upon a sequence of gradient iterations $\hat{\boldsymbol{\theta}}_e$ is estimated in Line 8. These gradient iterations can be used with traditional sparse-based reconstructions methods such as Sparse Wirtinger flow (SWF) [12], Sparse Truncated Amplitude flow (SPARTA) [14] or SPRSF [13]. Notice that, the computation of the matrix $\hat{\mathbf{A}} \in \mathbb{C}^{m \times r}$ in Line 4 is needed. In fact, $\hat{\mathbf{A}}$ allows Algorithm 2 to reduce the reconstruction time, because $r \ll n$. Finally, the overall computational complexity of Algorithm 2 is $\mathcal{O}(Tr^2)$ in contrast to $\mathcal{O}(Tn^2)$ of algorithms that do not decimate the sparse representation.

Algorithm 2 Multi-resolution reconstruction method for XC

1: **Input:** Data $\{(\mathbf{a}_i;\mathbf{y}_i)\}_{i=1}^m$, sparsity level k. The step size $\tau \in (0,1)$, control variables $\gamma, \gamma_1 \in (0,1)$, $\mu_{(0)} \in \mathbb{R}_{++}$, number of iterations T, threshold parameter σ, super-pixel size parameter η.
2: **Initialization:** S_0 is set to be the set of k largest indices of $\{\frac{1}{m}\sum_{i=1}^m q_i^2 a_{i,j}^2\}_{1\leq j\leq n}$, with $q_i = \sqrt{\mathbf{y}_i}$.
3: $\mathbf{M} \leftarrow$ **Algorithm 1** (S_0, σ, η)
4: Compute $\hat{\mathbf{A}} = [\hat{a}_1^T, \cdots, \hat{a}_{nL}^T]^H = \mathbf{A}\mathbf{F}^H\mathbf{M}^T$
5: Let $\tilde{\boldsymbol{\theta}}^{(0)}$ be the leading eigenvector of $\mathbf{Y} := \frac{1}{m}\sum_{i\in I_0}\sqrt{q_i}\frac{\hat{\mathbf{a}}_{i,S_0}\hat{\mathbf{a}}_{i,S_0}^H}{\|\hat{\mathbf{a}}_{i,S_0}\|_2^2}$
6: Define the initial point as $\hat{\boldsymbol{\theta}}^{(0)} := \lambda_0\tilde{\boldsymbol{\theta}}^{(0)}$, where $\lambda_0 := \sqrt{\frac{\sum_{i=1}^m q_i^2}{m}}$.
7: $\boldsymbol{\theta}_e \leftarrow \mathbf{M}\hat{\boldsymbol{\theta}}^{(0)}$
8: $\hat{\boldsymbol{\theta}}_e \leftarrow$ **reconstruction_algorithm**$(\boldsymbol{\theta}_e, \hat{\mathbf{A}}, \mathbf{y})$
9: **Output:** $\mathbf{F}^H\mathbf{M}^T\hat{\boldsymbol{\theta}}_e$

V. Simulations and Results

In this section, the performance of the proposed reconstruction algorithm and the decimation strategy are evaluated through three different tests: first, we determine the performance of some admissible random variables to estimate the non-zero coefficients of $\boldsymbol{\theta}$. Then, we compare the reconstruction time of the proposed method against state-of-the-art alternatives, and finally we examine the reconstruction quality attained with the proposed algorithm in terms of the Structural Similarity Index Measure (SSIM) due to it is regarded as a reliable indicator of image quality degradation, measuring from error visibility to structural similarity.

For the first test, the complex signal is generated for various values of $n \in [128, 4096]$. For tests two and three, the complex signal is $\mathbf{x} \sim \mathcal{N}(0, \mathbf{I}_{1024}) + j\mathcal{N}(0, \mathbf{I}_{1024})$. The default values of parameters of Algorithm 2 were determined using a cross-validation strategy. They were fixed as $\tau = 0.9$, $\gamma = 5, \gamma_1 = 0.015$, $\mu_{(0)} = 0.04$ $T = 500$, $\eta = 4$, $\sigma = 1e^{-3}$. Furthermore, in order to validate that our super-pixel approach can be used with any sparse reconstruction algorithm, we modified two algorithms of the state-of-the-art SWF and SPARTA, to introduce the decimation matrix $\mathbf{M}$.

A. Non-zero coefficients estimation

This section presents the performance to estimate the non-zero coefficients of $\boldsymbol{\theta}$ based on admissible random variables given by

$$d_1 = \begin{cases} -1 & \text{prob. } 1/4 \\ 1 & \text{prob. } 1/4 \\ j & \text{prob. } 1/4 \\ -j & \text{prob. } 1/4 \end{cases}, d_2 = \begin{cases} 0 & \text{prob. } 1/2 \\ 1 & \text{prob. } 1/2 \end{cases}, d_3 = \begin{cases} -1 & \text{prob. } 1/2 \\ 1 & \text{prob. } 1/2 \end{cases}. \tag{7}$$

We performed 100 trials to empirically investigate the success rate of the admissible random variables in (7) to estimate the non-zero coefficients of $\boldsymbol{\theta}$ as explained in Section III-A. The numerical results are summarized in Fig. 4 where the color represents the estimation percentage obtained for the different admissible random variables in (7). In particular, note that d_2 attains the highest performance compared with d_1 and d_3. This observation validates the theoretical result established in Lemma 1, because $\mathbb{E}[d_2] \neq 0$.

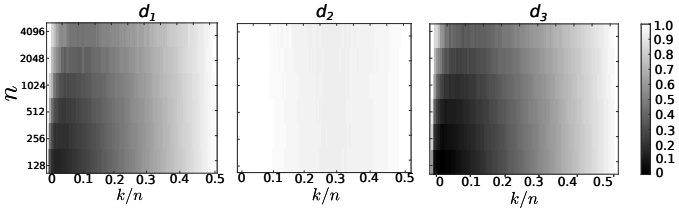

Fig. 4: Performance of some admissible random variables estimating the non-zero coefficients of the sparse representation.

B. Reconstruction time comparison

This section summarizes the reconstruction time comparison between the proposed algorithm and non-decimation-based approaches. Specifically, we compare Algorithm 2 with its non-decimation-based version proposed in [13]. Also, to demonstrate that our super-pixel methodology is independent of the reconstruction methodology, the algorithms SWF and SPARTA were adapted to solve (5). Specifically, Line 8 in Algorithm 2 can be replaced by SPARTA or SWF to solve (5). Figure 5 shows the attained results as a function of the compression rate. From Fig. 5 we can find that Algorithm 2 presents a reconstruction time reduction in at least 80% compared with the original version. Also, the decimation-based versions of SPARTA and SWF report a reconstruction time reduction with respect to their non-decimation-based versions of at least 77% and 84%, respectively.

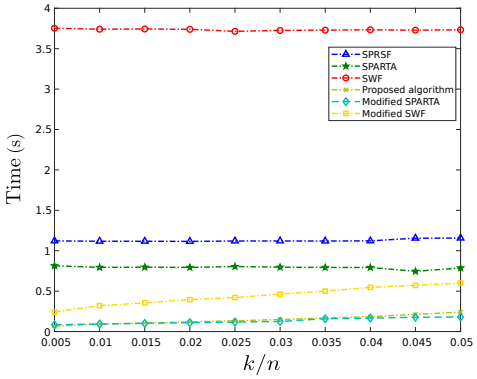

Fig. 5: Reconstruction time of the decimation-based algorithms vs the sparsity level of $\boldsymbol{\theta}$.

C. Reconstruction quality

This section presents the reconstruction quality attained with the proposed decimation algorithm, summarized in Fig. 6. Observe that the proposed reconstruction algorithm attains a gain of up to 6% in terms of the Structural Similarity Index Measure (SSIM) compared with the traditional method. Moreover, the decimation-based versions of SPARTA and SWF show an improvement in the reconstruction quality in up to 5% and 3% in terms of SSIM, respectively. This comes from the fact that the super-pixel approach reduces the number of unknowns implying that the non-zero coefficients

are better estimated compared with traditional sparse-based reconstruction methods.

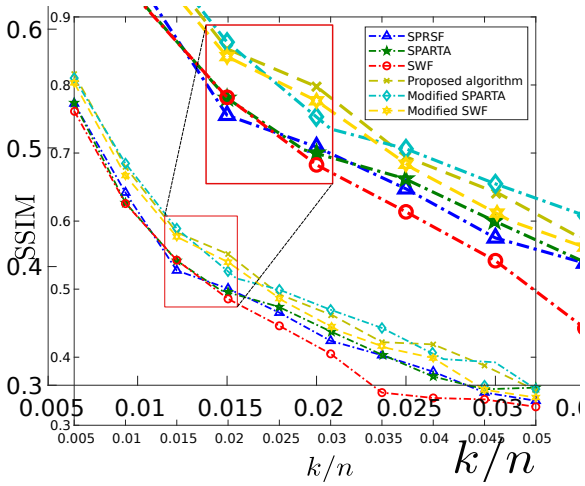

Fig. 6: Reconstruction quality attained with the decimation-based reconstruction algorithm.

VI. Conclusions

This work presented an algorithm to solve the phase retrieval in X-ray crystallography that is able to exploit the sparsity of $\boldsymbol{\theta}$ to reduce the total number of unknowns in the inverse problem. This algorithm groups sets of zero-valued pixels, called super-pixels, in the sparse representation to map a high-spatial-resolution image to one of lower resolution, reducing the total number of unknowns. This fact leads to a reconstruction time reduction of at least 80% and improves the reconstruction quality in up to 6% in terms of the Structural Similarity Index Measure (SSIM). Also, through simulations we validated that our super-pixel approach can be used by traditional sparse-based reconstruction methods, to reduce the reconstruction time in XC.

VII. Acknowledgment

The authors would like to thank the support provided by the Vicerrectoría de Investigación y extensión de Universidad Industrial de Santander under the VIE code 2467, entitled Optimización de aperturas codificadas para la reconstrucción de patrones de difracción de alta resolución en cristalografía de rayos X, utilizando técnicas de superresolución de imágenes.

Appendix A
Proof of Lemma 1

For two integers a and b we use $a \overset{n}{\equiv} b$ to denote congruence of a and b modulo n (n divides $a-b$). Set $\omega = e^{\frac{2\pi j}{n}}$ to be the nth root of unity so that $\mathbf{f}_i = [\omega^{-0(i-1)}, \cdots, \omega^{-(n-1)(i-1)}]$. Then, taking $z_p = \frac{1}{n}\sum_{q=1}^{n} Z_{q,p}$ we obtain that

$$z_p = \sum_{a,b,c,e=1}^{n} d_a \overline{d}_b d_c \overline{d}_e x_a \overline{x}_b \omega^{(c-e)(p-1)} \left(\frac{1}{n} \sum_{i=1}^{n} \omega^{(b-a+e-c)(i-1)} \right)$$
$$= \sum_{a=1}^{n}\sum_{b=1}^{n}\sum_{c=1}^{n}\sum_{e=1}^{n} d_a \overline{d}_b d_c \overline{d}_e x_a \overline{x}_b \omega^{(c-e)(p-1)} \mathbb{1}_{\{b+e \overset{n}{\equiv} a+c\}}, \tag{8}$$

where the notation $\mathbb{1}_{\{b+e \overset{n}{\equiv} a+c\}}$ means that $\frac{1}{n}\sum_{i=1}^{n} \omega^{(b-a+e-c)(i-1)}$ is different from zero, in this case it is equal to 1, only when the condition $b+e \overset{n}{\equiv} a+c$ is satisfied. Define $\mu = \mathbb{E}[d]$ and $h = d - \mu$. Therefore, from (8) we have that $\mathbb{E}[z_p]$ can be written as

$$\sum_{a,b,c,e=1}^{n} \left(\mathbb{E}\left[h_a \overline{h}_b h_c \overline{h}_e\right] x_a \overline{x}_b + |\mu|^4 x_a \overline{x}_b \right) \omega^{(c-e)(p-1)} \mathbb{1}_{\{b+e \overset{n}{\equiv} a+c\}}. \tag{9}$$

Observe that $\mathbb{E}\left[h_a \overline{h}_b h_c \overline{h}_e\right] = 0$ unless $(a = e, b = c)$, or $(a = b, c = e, a \neq c)$, where these two conditions also satisfy that $b + e \overset{n}{\equiv} a + c$. Thus, defining $\mathbb{E}\left[\overline{h}_a h_b \overline{h}_c h_e\right] = \mathbb{E}\left[|h_a|^2 |h_b|^2\right] = v \leq 1$ we have that

- $(a = e, b = c)$: Then, the term in (9) can be expressed as
$$\left(\sum_{a=1}^{n} v\omega^{-(a-1)(p-1)} x_a\right)\left(\sum_{b=1}^{n} v\omega^{(b-1)(p-1)} \overline{x}_b\right) = v^2 |\theta_p|^2. \tag{10}$$
- $(a = b, c = e, a \neq c)$: Thus, the term in (9) can be rewritten as
$$\sum_{a=1}^{n} \left(v|x_a|^2\right) \sum_{c \neq a}^{n} v = v^2(n-1)\|\mathbf{x}\|_2^2. \tag{11}$$

Thus, by combining (9), (10), and (11) it can be concluded that

$$\mathbb{E}[z_p] = v^2(n-1)\|\mathbf{x}\|_2^2 + (v^2 + |\mu|^4)|\theta_p|^2. \tag{12}$$

Since v is always greater than zero, it is clear from (12) that when $\mathbb{E}[d] \neq 0$, the non-zero coefficients can be better estimated.

References

[1] A. Walther, "The question of phase retrieval in optics," *Optica Acta: International Journal of Optics*, vol. 10, no. 1, pp. 41–49, 1963.
[2] C. Fienup and J. Dainty, "Phase retrieval and image reconstruction for astronomy," *Image Recovery: Theory and Application*, vol. 231, p. 275, 1987.
[3] J. Miao, P. Charalambous, J. Kirz, and D. Sayre, "Extending the methodology of x-ray crystallography to allow imaging of micrometre-sized non-crystalline specimens," *Nature*, vol. 400, no. 6742, p. 342, 1999.
[4] R. W. Harrison, "Phase problem in crystallography," *JOSA a*, vol. 10, no. 5, pp. 1046–1055, 1993.
[5] M. Smyth and J. Martin, "x ray crystallography," *Journal of Clinical Pathology*, vol. 53, no. 1, p. 8, 2000.
[6] R. P. Millane, "Phase retrieval in crystallography and optics," *JOSA A*, vol. 7, no. 3, pp. 394–411, 1990.
[7] G. Wang, G. B. Giannakis, and Y. C. Eldar, "Solving systems of random quadratic equations via truncated amplitude flow," *IEEE Transactions on Information Theory*, vol. 64, no. 2, pp. 773–794, 2018.
[8] E. J. Candes, X. Li, and M. Soltanolkotabi, "Phase retrieval via wirtinger flow: Theory and algorithms," *IEEE Transactions on Information Theory*, vol. 61, no. 4, pp. 1985–2007, 2015.
[9] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, "Slic superpixels," Tech. Rep., 2010.
[10] H. Garcia, C. V. Correa, and H. Arguello, "Multi-resolution compressive spectral imaging reconstruction from single pixel measurements," *IEEE Transactions on Image Processing*, vol. 27, no. 12, pp. 6174–6184, 2018.
[11] E. J. Candes, X. Li, and M. Soltanolkotabi, "Phase retrieval from coded diffraction patterns," *Applied and Computational Harmonic Analysis*, vol. 39, no. 2, pp. 277–299, 2015.
[12] Z. Yuan, H. Wang, and Q. Wang, "Phase retrieval via sparse wirtinger flow," *Journal of Computational and Applied Mathematics*, 2019.
[13] S. Pinilla, J. Bacca, and H. Arguello, "Sprsf: Sparse phase retrieval via smoothing function," *arXiv preprint arXiv:1807.09703*, 2018.
[14] G. Wang, L. Zhang, G. B. Giannakis, M. Akçakaya, and J. Chen, "Sparse phase retrieval via truncated amplitude flow," *IEEE Transactions on Signal Processing*, vol. 66, no. 2, pp. 479–491, 2018.