On Cartesian line sampling with anisotropic total variation regularization

Clarice Poon *†

February 1, 2016

Abstract

This paper considers the use of the anisotropic total variation seminorm to recover a two dimensional vector $x \in \mathbb{C}^{N \times N}$ from its partial Fourier coefficients, sampled along Cartesian lines. We prove that if $(x_{k,j} - x_{k-1,j})_{k,j}$ has at most s_1 nonzero coefficients in each column and $(x_{k,j} - x_{k,j-1})_{k,j}$ has at most s_2 nonzero coefficients in each row, then, up to multiplication by log factors, one can exactly recover x by sampling along s_1 horizontal lines of its Fourier coefficients and along s_2 vertical lines of its Fourier coefficients. Finally, unlike standard compressed sensing estimates, the log factors involved are dependent on the separation distance between the nonzero entries in each row/column of the gradient of x and not on N^2, the ambient dimension of x.

1 Introduction

Research into compressed sensing has resulted in several examples in which one can recover an s-sparse vector of length N from $\mathcal{O}(s \log N)$ randomly chosen linear measurements. One of the first examples of this, by Candès, Romberg and Tao, is the recovery of a gradient sparse vector from samples of its Fourier transform by means of solving a total variation regularization problem. This is perhaps one of the most well known and influential results in compressed sensing because of its links with applications, in particular, this result motivated the use of total variation regularization to reduce the sampling cardinality in many imaging applications, such as Electron Microscopy [9], Magnetic Resonance Imaging (MRI) [11], Optical Deflectometric Tomography [7], Phase-Contrast Tomography [6] and Radio Interferometry [16]. However, while studies into uniformly random sampling provide some insight into how total variation regularization can allow one to subsample the Fourier transform, there are two further aspects that one should consider.

1. Dense sampling at low frequencies and sparsity structure. It was observed in [11, 10] that one can obtain far superior results via *variable density sampling* where one samples more densely at low frequencies. This effect is demonstrated in Figure 1, where we compare the reconstruction of the boat test image from 12.3% of its Fourier coefficients via different sampling maps. On the theoretical side, one particular type of variable density sampling was first studied by Krahmer and Ward in [8] and later in [13]. The analysis of [13] showed that compared with sampling uniformly at random, one of the advantages offered by sampling more densely at low frequencies is improved robustness to inexact sparsity and noise. However, an important reason for the effectiveness of sampling densely at low frequencies is that although the

*CEREMADE, Université Paris-Dauphine †Email:cmhsp2@cam.ac.uk

sampling cardinality of $\mathcal{O}(s \log N)$ is optimal for the recovery of s-sparse vectors, one can further reduce this sampling cardinality by placing a structure assumption on the vector to be recovered. This observation was made by Candès and Fernandez-Granda [3] in the context of recovering a superposition of diracs in super-resolution, and in the case of total variation regularization of one dimensional signals, by exploiting the results of [3] and [14], the following result was proved in [13]:

Theorem 1.1. *Let $N \in \mathbb{N}$, let $\epsilon \in [0,1]$ and let $M \in \mathbb{N}$ be such that $N/4 \geq M \geq 10$. Let A be the discrete Fourier transform on $\mathbb{C}^N$ (defined in Section 1.3).*

- *Let $x \in \mathbb{C}^N$ and $\Delta \subset \{1, \ldots, N\}$ be of cardinality s and suppose that*
$$\min_{k,j \in \Delta, k \neq j} \frac{|k-j|}{N} \geq \frac{2}{M}.$$
- *Let $\Omega = \{0\} \cup \Omega'$ where $\Omega' \subset \{-M, \ldots, M\}$ consist of m indices chosen uniformly at random with*
$$m \gtrsim \max\left\{\log^2\left(\frac{M}{\epsilon}\right), \, s \cdot \log\left(\frac{s}{\epsilon}\right) \cdot \log\left(\frac{M}{\epsilon}\right)\right\}.$$

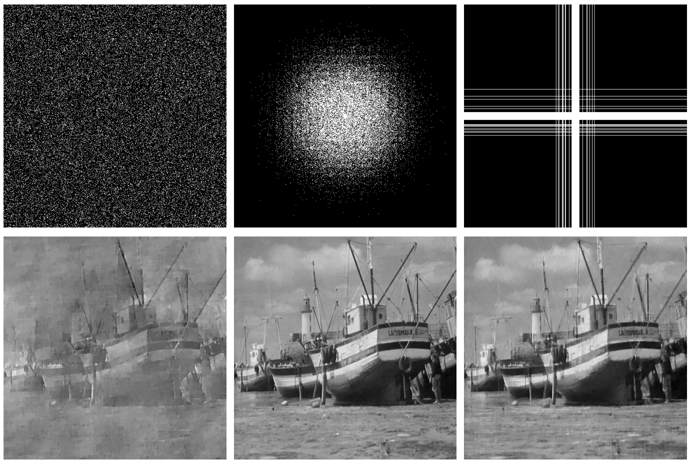
Then with probability exceeding $1-\epsilon$, given $y = P_\Omega A x + \eta$ and $\|\eta\|_2 \leq \delta \cdot \sqrt{m}$, any solution ξ to
$$\min_{x \in \mathbb{C}^N} \|x\|_{TV} \text{ subject to } \|P_\Omega A x - y\|_2 \leq \delta \cdot \sqrt{m} \tag{1.1}$$
satisfies
$$\frac{\|x - \xi\|_2}{\sqrt{N}} \lesssim \frac{N^2}{M^2} \cdot \left(\delta \cdot s + \sqrt{s} \cdot \|P_{\Delta^c} D x\|_1\right).$$
If $m = 2M+1$, then the error bound holds with probability 1.

So, if $x \in \mathbb{C}^N$ is s-gradient sparse with a minimum separation of $2/s$. Then, x can be exactly recovered from $2s+1$ Fourier coefficients. However, random sampling guarantees recovery only with $\mathcal{O}(s \log N)$ samples. Thus, one can further reduce the number of samples required by choosing the samples in accordance to the underlying sparsity structure.

2. The need to understand more realistic sampling patterns. Sampling in applications such as MRI is constrained to sampling along smooth trajectories, such as radial lines, spirals or Cartesian lines [10, 15]; on the other hand, the majority of results in compressed sensing describe only the effects of pointwise sampling. To our knowledge, the only theoretical result in this direction is by Boyer et al. in [2] where they consider the use of wavelet regularization while sampling along horizontal (or vertical) Cartesian lines in the Fourier domain.

1.1 This paper's contribution

The purpose of this paper is to present a two dimensional version of Theorem 1.1, where we consider how one can efficiently sample the Fourier transform along Cartesian lines by taking into account the sparsity structure in the gradient of the underlying vector. The Cartesian sampling pattern studied in this paper is one of the sampling patterns which has been empirically studied in the application of compressed sensing to MRI [10, 15]. Thus, the result of this paper provides further justification and insight into the use of compressed sensing in MRI. The main result of this paper is presented and discussed in Section 3 and its proof is presented in Section 4.



Rel. Err. 28.9% Rel. Err. 6.6% Rel. Err. 7.6%

Figure 1: The top row shows three different sampling maps, covering 12.3% of the Fourier coefficients. Note that the zeroth Fourier frequency corresponds to the centre of each map. The bottom row shows the corresponding reconstructions.

1.2 Related works – wavelet regularization

The link between success of dense sampling at low frequencies and the correspondence between such sampling patterns and the underlying sparsity structure has previously been investigated in the context of (orthogonal) wavelet regularization with Fourier sampling by Adcock et al. [1]. In the context of wavelet regularization, the relevant sparsity structure is the sparsity of the underlying wavelet coefficients within each wavelet scale and the results of [1] provide a link between the distribution of the Fourier samples and the wavelet sparsity at each scale, while the result of [12] demonstrate that one can recover the first N wavelet coefficients of the lowest scales from $\mathcal{O}(N)$ Fourier coefficients of the lowest frequencies by ℓ^1 wavelet regularization. Thus, the notion that ℓ^1 regularization can allow for sampling rates without the log factor is relevant in greater generality.

In [2], Boyer et al. investigated this structure dependency in the case of wavelet regularization with Cartesian line sampling in the Fourier domain. In particular, they proved that one can guarantee stable error bounds provided that the number of horizontal lines within each block of Fourier coefficients is proportional, up to log factors, with the sparsity in each column of the wavelet transform within the corresponding wavelet scale.

1.3 Notation

Let $N \in \mathbb{N}$, and let $[N] = \{-\lceil N/2 \rceil + 1, \dots, \lfloor N/2 \rfloor\}$. Let $A : \mathbb{C}^N \to \mathbb{C}^N$ be the discrete Fourier transform with

$$Az = \left(\sum_{j=1}^{N} z_j e^{-2\pi i k j/N} \right)_{k \in [N]}, \qquad z = (z_j)_{j=1}^N \in \mathbb{C}^N.$$

Let $D : \mathbb{C}^N \to \mathbb{C}^N$ be the finite differences operator defined for each $z \in \mathbb{C}^N$ as $Dz = (z_j - z_{j-1})_{j=1}^N$ where $z_0 := z_N$. Given any $\Lambda \subset \mathbb{Z}$ and $m \in \mathbb{N}$, $\Omega \sim \mathrm{Unif}(\Lambda, m)$ means that Ω consists of m elements of Λ drawn uniformly at random (without replacement).

Given any $z \in \mathbb{C}^N$, and any operator $V : \mathbb{C}^N \to \mathbb{C}^N$, Vz is a column vector whenever z is a column vector and a row vector whenever z is a row vector.

Given any $z \in \mathbb{C}^{N\times N}$, let $z^{[\mathrm{col},j]} = (z_{k,j})_{k\in[N]} \in \mathbb{C}^N$ denote the j^{th} column of z and let $z^{[\mathrm{row},k]} = (z_{k,j})_{j\in[N]} \in \mathbb{C}^N$ denote the k^{th} row of z.

Let $\tilde{A} : \mathbb{C}^{N\times N} \to \mathbb{C}^{N\times N}$ with

$$\tilde{A}z = \left(\sum_{k_1=1}^{N} \sum_{k_2=1}^{N} z_{k_1,k_2} e^{-2\pi i (k_1 n_1 + k_2 n_2)/N} \right)_{n_1,n_2\in[N]}$$

and let $\tilde{D}_1 : \mathbb{C}^{N\times N} \to \mathbb{C}^{N\times N}$ and $\tilde{D}_2 : \mathbb{C}^{N\times N} \to \mathbb{C}^{N\times N}$, with

$$\tilde{D}_1 z = (z_{k,j} - z_{k-1,j})_{k,j=1}^{N}, \quad \tilde{D}_2 z = (z_{k,j} - z_{k,j-1})_{k,j=1}^{N}.$$

Let $\tilde{D}z = (\tilde{D}_1 z, \tilde{D}_2 z)$, $\left\| \tilde{D}z \right\|_2 = \sqrt{\left\| \tilde{D}_1 z \right\|_2^2 + \left\| \tilde{D}_2 z \right\|_2^2}$, $\left\| \tilde{D}z \right\|_1 = \left\| \tilde{D}_1 z \right\|_1 + \left\| \tilde{D}_2 z \right\|_1$.

Given any $\Omega \subset \mathbb{Z} \times \mathbb{Z}$, let $\tilde{P}_\Omega : \mathbb{C}^{N\times N} \to \mathbb{C}^{N\times N}$ with

$$\tilde{P}_\Omega z = y, \quad y_j = \begin{cases} z_j & j \in \Omega, \\ 0 & j \notin \Omega. \end{cases}$$

Let $\|\cdot\|_{TV}$ denote the anisotropic total variation norm with

$$\|z\|_{TV} := \left\| \tilde{D}_1 z \right\|_1 + \left\| \tilde{D}_2 z \right\|_1, \qquad \forall z \in \mathbb{C}^{N\times N}$$

and given $\Delta_1, \Delta_2 \subset \{1, \dots, N\}^2$, let

$$\|z\|_{TV,\Delta_1,\Delta_2} := \left\| \tilde{P}_{\Delta_1} \tilde{D}_1 z \right\|_1 + \left\| \tilde{P}_{\Delta_2} \tilde{D}_2 z \right\|_1, \qquad \forall z \in \mathbb{C}^{N\times N}.$$

Given $a, b \in \mathbb{R}$, we write $a \lesssim b$ if there exists some constant $C > 0$ (independent of all variables under consideration) such that $a \le C \cdot b$.

2 Key concepts

In Theorem 1.1, the sparsity structure considered is the separation between the discontinuities of the underlying signal. When considering the recovery of some vector $x \in \mathbb{C}^{N\times N}$ by sampling along Cartesian lines of its Fourier transform, our main result will demonstrate how one should subsample depends on the sparsity and the minimum separation distance within each column of $\tilde{D}_1 x$ and each row of $\tilde{D}_2 x$. We first present three definitions that our main result will depend on.

Definition 2.1 (Sparsity). *Let $\Delta \subset \{1, \dots, N\}^2$. The column cardinality of Δ is*

$$s_1 = \max_{j=1}^{N} |\{k : (k,j) \in \Delta\}|.$$

The row cardinality of Δ is

$$s_2 = \max_{k=1}^{N} |\{j : (k,j) \in \Delta\}|.$$

Definition 2.2 (Minimum separation distance)**.** *Let $N \in \mathbb{N}$ and let $\Delta \subset \{1, \ldots, N\}^2$. The minimum separation distance of its rows is defined to be*

$$\nu_{\text{row}}(\Delta, N) = \min_{n=1}^{N} \min \left\{ \frac{|j-k|}{N} : (j,n), (k,n) \in \Delta, j \neq k \right\},$$

and minimum separation distance of its columns is defined to be

$$\nu_{\text{col}}(\Delta, N) = \min_{n=1}^{N} \min \left\{ \frac{|j-k|}{N} : (n,j), (n,k) \in \Delta, j \neq k \right\}.$$

Definition 2.3. *We say that x has T_1 distinct column supports if*

$$T_1 = \left| \left\{ (x_{k,j})_{k=1}^{N} : j = 1, \ldots, N \right\} \right|,$$

and say that x has T_2 distinct row supports if

$$T_2 = \left| \left\{ (x_{k,j})_{j=1}^{N} : k = 1, \ldots, N \right\} \right|.$$

3 Main theorem

Let $x \in \mathbb{C}^{N \times N}$ and let $\Delta_1, \Delta_2 \subset \{1, \ldots, N\}^2$. Suppose that $\tilde{P}_{\Delta_1} \operatorname{sgn}(\tilde{D}_1 x)$ has T_1 distinct column supports with a minimum separation of $2/M_1$ along its columns, and $\tilde{P}_{\Delta_2} \operatorname{sgn}(\tilde{D}_2 x)$ has T_2 distinct supports with a minimum separation of $2/M_2$ along its rows. Suppose that Δ_1 has column cardinality s_1 and Δ_2 has row cardinality s_2. Assume also that for $i = 1, 2$,

$$s_i \log(T_i s_i / \epsilon) \geq \log(T_i M_i / \epsilon).$$

Theorem 3.1. *Let $\epsilon \in (0,1)$ and let $\Omega = \{0\} \cup \{\Omega_1 \times [N]\} \cup \{[N] \times \Omega_2\}$ and let $m = |\Omega|$, where*

$$\Omega_1 \sim \operatorname{Unif}([M_1], m_1), \qquad m_1 \gtrsim s_1 \log(T_1 s_1/\epsilon) \log(T_1 M_1/\epsilon),$$

and

$$\Omega_2 \sim \operatorname{Unif}([M_2], m_2), \qquad m_2 \gtrsim s_2 \log(T_2 s_2/\epsilon) \log(T_2 M_2/\epsilon).$$

Let $\xi = P_\Omega x + \eta$ with $\|\eta\|_2 \leq \delta \cdot \sqrt{m}$, and suppose that $\hat{x}$ is a minimizer of

$$\min_{z \in \mathbb{C}^{N \times N}} \|z\|_{TV} \text{ subject to } \left\| \tilde{P}_\Omega \tilde{A} z - \xi \right\|_2 \leq \delta \cdot \sqrt{m}. \tag{3.1}$$

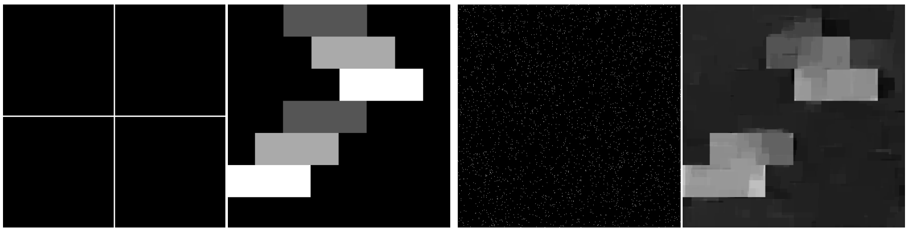
Then, with probability exceeding $1 - \epsilon$,

$$\left\| \tilde{D}(x - \hat{x}) \right\|_2 \lesssim \frac{N^2}{M_0^2} \left((m_0 N)^{-1/2} \sqrt{m} \delta + \|x\|_{TV, \Delta_1^c, \Delta_2{}^c} \right), \tag{3.2}$$

and

$$\|x - \hat{x}\|_2 \lesssim \frac{N^2}{M_0^2} \left((m/m_0)^{1/2} \sqrt{s} \delta + \sqrt{s} \|x\|_{TV, \Delta_1{}^c, \Delta_2{}^c} \right), \tag{3.3}$$

where $s = \max\{s_1, s_2\}$, $m_0 = \min\{m_1, m_2\}$, and $M_0 = \min\{M_1, M_2\}$. If $\Omega_1 = [M_1]$ and $\Omega_2 = [M_2]$, then (3.2) and (3.3) hold with probability one.



(a) Samp. map (a) Recov. Rel Err. 0% (b) Samp. map (b) Recov. Rel. Err. 49.78%

Figure 2: In (a), the test image (500×500) can be perfectly recovered from 1.2% of its Fourier coefficients, indexed by the Cartesian lines passing through the low frequencies. In (b), the reconstruction obtained from sampling 1.2% of the Fourier coefficients uniformly at random is shown.

3.1 Remarks on the main result

Removal of the $\log N$ factor: Suppose that $x \in \mathbb{C}^{N\times N}$ is such that the support of $\tilde{D}^{[\mathrm{col}]}x$ consists of M_1 lines which have a minimum separation of $1/M_1$ and the support of $\tilde{D}^{[\mathrm{row}]}x$ consists of M_2 lines which have a minimum separation of $1/M_2$. Then, one can perfectly recover x from sampling the Fourier transform of x along $\mathcal{O}(M_1)$ horizontal lines and $\mathcal{O}(M_2)$ vertical lines. On the other hand, in this case, the sparsity is $(M_1+M_2)N$ and one is guaranteed exact recovery from sampling uniformly at random when one observes $\mathcal{O}((M_1+M_2)N\log N)$ samples. Figure 2 illustrates this effect by showing the recovery of an image from 1.2% of its Fourier coefficients. The test image can be perfect reconstructed from sampling the low frequency Cartesian lines, on the other hand, sampling 1.2 % of the Fourier coefficients uniformly at random yields a poor reconstruction. Note that this image can in fact be recovered from 3% of its Fourier coefficients drawn uniformly at random, it is simply that one can sample less by considering the sparsity structure of the test image.

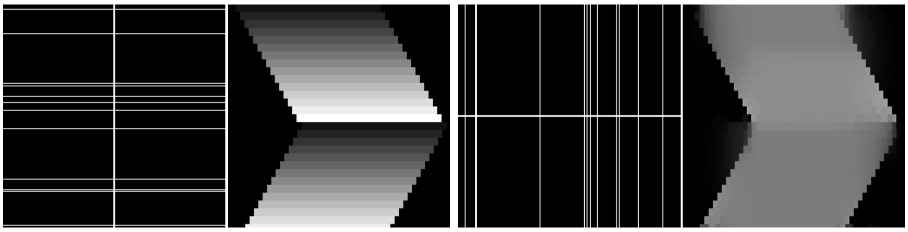
The importance of structure dependency: Note that from Theorem 3.1, the range that one should sample from is dependent on the minimum separation between the discontinuities in the corresponding direction, and the number of samples that one should draw is up to log factors dependent on the maximum sparsity in each row or each column of the corresponding direction. It is important to consider this structure dependency when devising a sampling scheme – see Figure 3.

4 Proof

As is now standard in compressed sensing, the proof of Theorem 3.1 consists of showing the existence of some dual certificate [5, 4]. Following the arguments in [13], one can show that given $x \in \mathbb{C}^{N\times N}$ with $\mathrm{supp}(\tilde{D}_1 x) = \Delta_1$ and $\mathrm{supp}(\tilde{D}_2 x) = \Delta_2$, x is the unique solution of (3.1) if $0 \in \Omega$, and the following two conditions hold:

(i) For $i = 1, 2$, there exists some dual certificate $\rho_i, \in \mathrm{ran}(\tilde{A}^* \tilde{P}_\Omega)$ such that $\|\rho_i\|_\infty$ and $\tilde{P}_{\Delta_i}\rho_i = \mathrm{sgn}(\tilde{D}_i x)$,

(ii) $\tilde{P}_\Omega \tilde{A} \tilde{P}_{\Delta_i}$ is injective for $i = 1, 2$.



(a) Samp. map (a) Recov. Rel. Err. 0% (b) Samp. map (b) Recov. Rel. Err. 44.8%

Figure 3: This figure shows two sampling maps and their corresponding reconstructions (256×256). Both sampling maps cover 5.4% of the Fourier coefficients. The sampling map (a) is constructed by sampling along 12 lines uniformly at random in the horizontal direction and along the 2 lowest frequency lines in the vertical direction. The sampling map (b) is obtained in the same manner, but sampled in the opposite orientations with 12 random lines in the vertical direction and 2 of the lowest frequency lines in the horizontal direction.

However, instead of directly showing the existence of one dual certificate defined on $\mathbb{C}^{N\times N}$, we will exploit the fact that we are given Fourier samples along Cartesian lines and apply Proposition 4.2 to show that it suffices to prove the existence a sequence of one dimensional certificates defined on $\mathbb{C}^N$.

Lemma 4.1. *Let $N \in \mathbb{N}$, $\Omega \subset [N]$ and let $z \in \mathbb{C}^{N\times N}$. Then,*

$$\left\|\tilde{P}_{\Omega\times[N]}\tilde{A}z\right\|_2^2 = N\sum_{k=1}^N \left\|P_\Omega A z^{[\mathrm{col},k]}\right\|_2^2, \qquad \left\|\tilde{P}_{[N]\times\Omega}\tilde{A}z\right\|_2^2 = N\sum_{k=1}^N \left\|P_\Omega A z^{[\mathrm{row},k]}\right\|_2^2.$$

Proof. We first note that

$$\begin{aligned}(\tilde{A}z)_{n_1,n_2} &= \sum_{k_2=1}^N \left(\sum_{k_1=1}^N z_{k_1,k_2} e^{-2\pi i k_1 n_1/N}\right) e^{-2\pi i k_2 n_2/N} \\ &= \sum_{k_2=1}^N (Az^{[\mathrm{col},k_2]})_{n_1} e^{-2\pi i k_2 n_2/N} = (A\beta^{[n_1]})_{n_2},\end{aligned}$$

where $\beta^{[n_1]} = \left((Az^{[\mathrm{col},k]})_{n_1}\right)_{k=1}^N$. Then, by using this identity, we have that

$$\begin{aligned}\left\|\tilde{P}_{\Omega\times[N]}\tilde{A}z\right\|_2^2 &= \sum_{n_1\in\Omega}\sum_{n_2\in[N]} \left|(\tilde{A}z)_{n_1,n_2}\right|^2 = \sum_{n_1\in\Omega}\sum_{n_2\in[N]} \left|(A\beta^{[n_1]})_{n_2}\right|^2 \\ &= \sum_{n_1\in\Omega}\left\|A\beta^{[n_1]}\right\|_2^2 = N\sum_{n_1\in\Omega}\left\|\beta^{[n_1]}\right\|_2^2 = N\sum_{k=1}^N\sum_{n_1\in\Omega}\left|(Az^{[\mathrm{col},k]})_{n_1}\right|^2 \\ &= N\sum_{k=1}^N\left\|P_\Omega Az^{[\mathrm{col},k]}\right\|_2^2,\end{aligned} \tag{4.1}$$

where we have applied in the second line the fact that $N^{-1/2}A$ is unitary. Finally, by a symmetric argument,

$$\left\|\tilde{P}_{[N]\times\Omega}\tilde{A}z\right\|_2^2 = N\sum_{k=1}^{N}\left\|P_\Omega A z^{[\mathrm{row},k]}\right\|_2^2.$$

$\square$

Proposition 4.2 (Dual certificates)**.** *Let $x \in \mathbb{C}^{N\times N}$. Let $\Delta_1, \Delta_2 \subset \{1,\dots,N\}^2$. Let $\Omega_1, \Omega_2 \subset [N]$ and let $m_1 = |\Omega_1|$ and $m_2 = |\Omega_2|$. Let*

$$\Omega = \{\Omega_1 \times [N]\} \cup \{[N] \times \Omega_2\}.$$

Let $m = |\Omega|$. Let $\xi = P_\Omega x + \eta$ with $\|\eta\|_2 \le \delta\cdot\sqrt{m}$, and suppose that $\hat{x}$ is a minimizer of

$$\min_{z\in\mathbb{C}^{N\times N}} \|z\|_{TV} \text{ subject to } \left\|\tilde{P}_\Omega\tilde{A}z - \xi\right\|_2 \le \delta\cdot\sqrt{m}. \tag{4.2}$$

For $n = 1, 2$, let $\Delta_{n,j} = \{k : (k,j) \in \Delta_n\} \subset \{1,\dots,N\}$ and let $s_n = \max_{j=1}^{N}|\Delta_{n,j}|$. Assume that the following conditions hold.

(i) For each $j \in \{1,\dots,N\}$,

$$m_1^{-1/2}\inf_{\mathrm{supp}(x)=\Delta_{1,j},\|x\|_2=1}\|P_{\Omega_1}Ax\|_2 \ge c_1 > 0.$$

(ii) For each $j \in \{1,\dots,N\}$,

$$m_2^{-1/2}\inf_{\mathrm{supp}(x)=\Delta_{2,j},\|x\|_2=1}\|P_{\Omega_2}Ax\|_2 \ge c_1 > 0.$$

(iii) For each $j \in \{1,\dots,N\}$, there exists $\rho_j = m_1^{-1/2}A^*P_{\Omega_1}w_j \in \mathbb{C}^N$ such that

$$P_{\Delta_{1,j}}\rho_j = P_{\Delta_{1,j}}\mathrm{sgn}(\tilde{D}_1x)^{[\mathrm{col},j]}, \qquad \left\|P_{\Delta_{1,j}}^\perp\rho_j\right\|_\infty \le c_2 < 1, \qquad \sum_{j=1}^{N}\|w_j\|_2 \le L^2.$$

(iv) For each $j \in \{1,\dots,N\}$, there exists $\tau_j = m_2^{-1/2}A^*P_{\Omega_2}u_j \in \mathbb{C}^N$ such that

$$P_{\Delta_{2,j}}\tau_j = P_{\Delta_{2,j}}\mathrm{sgn}(\tilde{D}_2x)^{[\mathrm{col},j]}, \qquad \left\|P_{\Delta_{2,j}}^\perp\tau_j\right\|_\infty \le c_2 < 1, \qquad \sum_{j=1}^{N}\|u_j\|_2 \le L^2.$$

Then,

$$\left\|\tilde{D}(x-\hat{x})\right\|_2 \lesssim C_1\cdot\delta\cdot + C_2\cdot\|x\|_{TV,{\Delta_1}^c,{\Delta_2}^c},$$

and

$$\|x-\hat{x}\|_2 \lesssim C_3\cdot\delta\cdot\sqrt{\frac{sm}{m_0}} + C_2\cdot\sqrt{s}\cdot\|x\|_{TV,{\Delta_1}^c,{\Delta_2}^c},$$

where $s = \max\{s_1, s_2\}$, $m_0 = \min\{m_1, m_2\}$,

$$C_1 = (m_0N)^{-1/2}m^{1/2}(1+L)(1+c_1^{-1})(1-c_2)^{-1}, \qquad C_2 = (1+c_1^{-1})(1-c_2)^{-1},$$

and $C_3 = c_1^{-1}(1 + L(1-c_2)^{-1}N^{-1/2})$.

Proof of Proposition 4.2. First, suppose that $z \in \mathbb{C}^{N\times N}$ is such that $\left\|\tilde{P}_{\Omega_1\times[N]}\tilde{A}z\right\|_2 \leq \sqrt{m}\delta$. By applying assumption (i) and Lemma 4.1,

$$\begin{aligned}
&c_1^2\left\|\tilde{P}_{\Delta_1}z\right\|_2^2 = c_1^2\sum_{j=1}^N \left\|P_{\Delta_{1,j}}z^{[\mathrm{col},j]}\right\|_2^2 \leq \sum_{j=1}^N m_1^{-1}\left\|P_{\Omega_1}AP_{\Delta_{1,j}}z^{[\mathrm{col},j]}\right\|_2^2 \\
&= (m_1N)^{-1}\left\|\tilde{P}_{\Omega_1\times[N]}\tilde{A}\tilde{P}_{\Delta_1}z\right\|_2^2 \leq (m_1N)^{-1}\left(\left\|\tilde{P}_{\Omega_1\times[N]}\tilde{A}z\right\|_2 + \left\|\tilde{P}_{\Omega_1\times[N]}\tilde{A}\tilde{P}_{\Delta_1}^{\perp}z\right\|_2\right)^2 \\
&\leq (m_1N)^{-1}\left(\sqrt{m}\delta + \max_{j\notin\Delta_1}\left\|\tilde{P}_{\Omega_1\times[N]}\tilde{A}e_j\right\|_2\left\|\tilde{P}_{\Delta_1}^{\perp}z\right\|_1\right)^2.
\end{aligned}$$

Note that

$$(m_1N)^{-1/2}\max_{j\notin\Delta_1}\left\|\tilde{P}_{\Omega_1\times[N]}\tilde{A}e_j\right\|_2 = 1.$$

It thus follows that

$$\left\|\tilde{P}_{\Delta_1}z\right\|_2 \leq \frac{\sqrt{m}\delta}{c_1\sqrt{Nm_1}} + \frac{\left\|\tilde{P}_{\Delta_1}^{\perp}z\right\|_1}{c_1}. \tag{4.3}$$

Similarly, it follows from assumption (ii) that for any $z \in \mathbb{C}^{N\times N}$ with $\left\|\tilde{P}_{[N]\times\Omega_2}\tilde{A}z\right\|_2 \leq \sqrt{m}\delta$,

$$\left\|\tilde{P}_{\Delta_2}z\right\|_2 \leq \frac{\sqrt{m}\delta}{c_1\sqrt{Nm_2}} + \frac{\left\|\tilde{P}_{\Delta_2}^{\perp}z\right\|_1}{c_1}. \tag{4.4}$$

Let $h = \hat{x} - x$, and observe that since $\hat{x}$ and x both satisfy the constraint of (4.2),

$$\left\|\tilde{P}_{\Omega_1\times[N]}\tilde{A}h\right\|_2 \leq \left\|\tilde{P}_{\Omega_1\times[N]}\tilde{A}x - \xi\right\|_2 + \left\|\tilde{P}_{\Omega_1\times[N]}\tilde{A}\hat{x} - \xi\right\|_2 \leq 2\sqrt{m}\delta.$$

Note also that

$$(\tilde{A}\tilde{D}_1h)_{k,j} = (1 - e^{-2\pi ik/N})(\tilde{A}h)_{k,j}, \qquad (\tilde{A}\tilde{D}_2h)_{k,j} = (1 - e^{-2\pi ij/N})(\tilde{A}h)_{k,j}.$$

Therefore, $\left\|\tilde{P}_{\Omega_1\times[N]}\tilde{A}\tilde{D}_1h\right\|_2 \leq 2\left\|\tilde{P}_{\Omega\times[N]}\tilde{A}h\right\|_2 \leq 4\delta\sqrt{m}$. Similarly, $\left\|\tilde{P}_{[N]\times\Omega_2}\tilde{A}\tilde{D}_2h\right\|_2 \leq 4\delta\sqrt{m}$. So, we can apply the bounds (4.3) and (4.4) to obtain

$$\begin{aligned}
\left\|\tilde{P}_{\Delta_1}\tilde{D}_1h\right\|_2 &\leq \frac{4\sqrt{m}\delta}{c_1\sqrt{Nm_1}} + \frac{4\left\|\tilde{P}_{\Delta_1}^{\perp}\tilde{D}_1h\right\|_1}{c_1}, \\
\left\|\tilde{P}_{\Delta_2}\tilde{D}_2h\right\|_2 &\leq \frac{4\sqrt{m}\delta}{c_1\sqrt{Nm_2}} + \frac{4\left\|\tilde{P}_{\Delta_2}^{\perp}\tilde{D}_2h\right\|_1}{c_1}.
\end{aligned} \tag{4.5}$$

We now proceed to derive upper bounds for $\left\|\tilde{P}_{\Delta_1}^{\perp}\tilde{D}_1h\right\|_1$ and $\left\|\tilde{P}_{\Delta_2}^{\perp}\tilde{D}_2h\right\|_1$: By Hölder's inequality and the triangle inequality,

$$\begin{aligned}
&\left\|\tilde{D}_1\hat{x}\right\|_1 = \left\|P_{\Delta_1}\tilde{D}_1(x+h)\right\|_1 + \left\|P_{\Delta_1}^{\perp}\tilde{D}_1(x+h)\right\|_1 \\
&\geq \left\|P_{\Delta_1}\tilde{D}_1x\right\|_1 + \mathrm{Re}\,\langle P_{\Delta_1}\tilde{D}_1h, \mathrm{sgn}(\tilde{D}_1x)\rangle + \left\|P_{\Delta_1}^{\perp}\tilde{D}_1h\right\|_1 - \left\|P_{\Delta_1}^{\perp}\tilde{D}_1x\right\|_1.
\end{aligned}$$

Rearranging the terms yields

$$\left\|P_{\Delta_1}^{\perp}\tilde{D}_1 h\right\|_1 \leq \left\|\tilde{D}_1\hat{x}\right\|_1 - \left\|\tilde{D}_1 x\right\|_1 + 2\left\|P_{\Delta_1}^{\perp}\tilde{D}_1 x\right\|_1 + \left|\langle P_{\Delta_1}\tilde{D}_1 h, \operatorname{sgn}(\tilde{D}_1 x)\rangle\right|.$$

Similarly,

$$\left\|P_{\Delta_2}^{\perp}\tilde{D}_2 h\right\|_1 \leq \left\|\tilde{D}_2\hat{x}\right\|_1 - \left\|\tilde{D}_2 x\right\|_1 + 2\left\|P_{\Delta_2}^{\perp}\tilde{D}_2 x\right\|_1 + \left|\langle P_{\Delta_2}\tilde{D}_2 h, \operatorname{sgn}(\tilde{D}_2 x)\rangle\right|.$$

Now, since $\hat{x}$ is a minimizer of (4.2),

$$\left\|\tilde{D}_1\hat{x}\right\|_1 + \left\|\tilde{D}_2\hat{x}\right\|_1 \leq \left\|\tilde{D}_1 x\right\|_1 + \left\|\tilde{D}_2 x\right\|_1.$$

So,

$$\begin{aligned} &\left\|P_{\Delta_1}^{\perp}\tilde{D}_1 h\right\|_1 + \left\|P_{\Delta_2}^{\perp}\tilde{D}_2 h\right\|_1 \\ &\leq 2\|x\|_{TV,\Delta_1^c,\Delta_2^c} + \left|\langle P_{\Delta_1}\tilde{D}_1 h, \operatorname{sgn}(\tilde{D}_1 x)\rangle\right| + \left|\langle P_{\Delta_2}\tilde{D}_2 h, \operatorname{sgn}(\tilde{D}_2 x)\rangle\right|. \end{aligned} \tag{4.6}$$

We now proceed to bound $\left|\langle P_{\Delta_1}\tilde{D}_1 h, \operatorname{sgn}(\tilde{D}_1 x)\rangle\right|$. Let $y = \tilde{D}_1 x$ and let $z = \tilde{D}_1 h$. By using the existence of $\rho_j = m_1^{-1/2} A^* P_{\Omega_1} w_j \in \mathbb{C}^N$ for $j = 1, \dots, N$ (from assumption (iii)), we have the following bound.

$$\begin{aligned} &\left|\langle \tilde{P}_{\Delta_1} z, \operatorname{sgn}(y)\rangle\right| = \left|\sum_{j=1}^{N}\langle P_{\Delta_{1,j}} z^{[\mathrm{col},j]}, \operatorname{sgn}(y)^{[\mathrm{col},j]}\rangle\right| \\ &= \left|\sum_{j=1}^{N}\langle P_{\Delta_{1,j}} z^{[\mathrm{col},j]}, \operatorname{sgn}(y)^{[\mathrm{col},j]} - \rho_j\rangle + \langle z^{[\mathrm{col},j]}, \rho_j\rangle - \langle P_{\Delta_{1,j}}^{\perp} z^{[\mathrm{col},j]}, \rho_j\rangle\right| \\ &\leq \left|\sum_{j=1}^{N}\langle m_1^{-1/2} P_{\Omega_1} A z^{[\mathrm{col},j]}, w_j\rangle\right| + \sum_{j=1}^{N}\left\|P_{\Delta_{1,j}}^{\perp} z^{[\mathrm{col},j]}\right\|_1 \left\|P_{\Delta_{1,j}}^{\perp}\rho_j\right\|_\infty \\ &\leq \sqrt{\sum_{j=1}^{N} m_1^{-1}\left\|P_{\Omega_1} A z^{[\mathrm{col},j]}\right\|_2^2}\sqrt{\sum_{j=1}^{N}\|w_j\|_2^2} + c_2\left\|\tilde{P}_{\Delta_1}^{\perp} z\right\|_1, \end{aligned}$$

where we have applied the Cauchy-Schwarz inequality to obtain the last line. Recall from Lemma 4.1 that

$$\begin{aligned} &\sqrt{\sum_{j=1}^{N} m_1^{-1}\left\|P_{\Omega_1} A z^{[\mathrm{col},j]}\right\|_2^2} = (m_1 N)^{-1/2}\left\|\tilde{P}_{\Omega_1\times[N]}\tilde{A} z\right\|_2 \\ &= (m_1 N)^{-1/2}\left\|\tilde{P}_{\Omega_1\times[N]}\tilde{A} D h\right\|_2 \leq 4(m_1 N)^{-1/2} m^{1/2}\delta. \end{aligned}$$

Hence it follows that

$$\left|\langle \tilde{P}_{\Delta_1}\tilde{D}_1 h, \operatorname{sgn}(\tilde{D}_1 x)\rangle\right| \leq 4L(m_1 N)^{-1/2} m^{1/2}\delta + c_2\left\|\tilde{P}_{\Delta_1}^{\perp}\tilde{D}_1 h\right\|_1.$$

A similar argument also yields

$$\left|\langle \tilde{P}_{\Delta_2}\tilde{D}_2 h, \operatorname{sgn}(\tilde{D}_2 x)\rangle\right| \leq 4L(m_2 N)^{-1/2} m^{1/2}\delta + c_2\left\|\tilde{P}_{\Delta_2}^{\perp}\tilde{D}_2 h\right\|_1.$$

By plugging these two estimates back into (4.6) and rearranging, we have that

$$\begin{aligned}&\left\|\tilde{P}_{\Delta_1}^{\perp}\tilde{D}_1 h\right\|_1+\left\|\tilde{P}_{\Delta_2}^{\perp}\tilde{D}_2 h\right\|_1\\&\leq(1-c_2)^{-1}\left(4LN^{-1/2}m^{1/2}(m_1^{-1/2}+m_2^{-1/2})\delta+\|x\|_{TV,\Delta_1^c,\Delta_2^c}\right).\end{aligned}\tag{4.7}$$

Combining this estimate with (4.5) yields the required bound on $\left\|\tilde{D}(\hat{x}-x)\right\|_2$.

To derive the bound on $\|\hat{x}-x\|_{TV}$, first let $z=\tilde{D}_1(\hat{x}-x)$. Note that by the Cauchy-Schwarz inequality,

$$\left\|\tilde{P}_{\Delta_1}z\right\|_1=\sum_{j=1}^{N}\left\|P_{\Delta_{1,j}}z^{[\mathrm{col},j]}\right\|_1\leq\sqrt{s_1}\sum_{j=1}^{N}\left\|P_{\Delta_{1,j}}z^{[\mathrm{col},j]}\right\|_2$$

By applying condition (i), $\left\|\tilde{P}_{\Delta_1}z\right\|_1$ is upper bounded by

$$\begin{aligned}&\frac{\sqrt{s_1}}{c_1\sqrt{m_1}}\sum_{j=1}^{N}\left\|P_{\Omega_1}AP_{\Delta_{1,j}}z^{[\mathrm{col},j]}\right\|_2\leq\frac{\sqrt{s_1}}{c_1\sqrt{m_1}}\sum_{j=1}^{N}\left(\left\|P_{\Omega_1}Az^{[\mathrm{col},j]}\right\|_2+\left\|P_{\Omega_1}AP_{\Delta_{1,j}}^{\perp}z^{[\mathrm{col},j]}\right\|_2\right)\\&\leq\frac{\sqrt{s_1}}{c_1\sqrt{m_1}}\sqrt{N}\sqrt{\sum_{j=1}^{N}\|P_{\Omega_1}Az^{[\mathrm{col},j]}\|_2^2}+\frac{\sqrt{s_1}}{c_1\sqrt{m_1}}\max_{j=1}^{N}\max_{l\notin\Delta_{1,j}}\|P_{\Omega_1}Ae_l\|_2\sum_{j=1}^{N}\left\|P_{\Delta_{1,j}}^{\perp}z^{[\mathrm{col},j]}\right\|_1\\&\leq\frac{\sqrt{s_1}}{c_1\sqrt{m_1}}\left\|\tilde{P}_{\Omega_1\times[N]}\tilde{A}z\right\|_2+\frac{\sqrt{s_1}}{c_1}\left\|\tilde{P}_{\Delta_1}^{\perp}z\right\|_1\leq\frac{4\delta\sqrt{m}\sqrt{s_1}}{c_1\sqrt{m_1}}+\frac{\sqrt{s_1}}{c_1}\left\|\tilde{P}_{\Delta_1}^{\perp}z\right\|_1.\end{aligned}$$

By a symmetric argument,

$$\left\|\tilde{P}_{\Delta_2}z\right\|_1\leq\frac{4\delta\sqrt{m}\sqrt{s_2}}{c_1\sqrt{m_2}}+\frac{\sqrt{s_2}}{c_1}\left\|\tilde{P}_{\Delta_2}^{\perp}z\right\|_1$$

Therefore, by combining with the bound from (4.7),

$$\|\hat{x}-x\|_{TV}\lesssim C\frac{\sqrt{sm}}{\sqrt{m_0}}\delta+\left(1+\frac{\sqrt{s}}{c_1}\right)(1-c_2)^{-1}\|x\|_{TV,\Delta_1{}^c,\Delta_2{}^c},$$

where $s=\max\{s_1,s_2\}$, $m_0=\min\{m_1,m_2\}$, and $C=c_1^{-1}(1+L(1-c_2)^{-1}N^{-1/2})$. Finally, recall that due to the Poincaré inequality, any zero mean image $X\in\mathbb{C}^{N\times N}$ satisfies

$$\|X\|_2\leq\|X\|_{TV}.\tag{4.8}$$

Since $0\in\Omega$, $\left|\sum_j(x-\hat{x})_j\right|\leq2\delta\sqrt{m}$, so, by letting $X_j=(\hat{x}-x)_j-\frac{1}{N^2}\sum_j(x-\hat{x})_j$ for each $j\in\{1,\ldots,N\}^2$, (4.8) and the triangle inequality yields

$$\|\hat{x}-x\|_2\leq\delta+\|\hat{x}-x\|_{TV},$$

and hence, the conclusion follows.

□

Proof of Theorem 3.1. To prove this theorem, we simply need to show that conditions (i) to (iv) of Proposition 4.2 hold with high probability. Note that these conditions were studied in [13]: we recall from Lemmas 4.25 of [13] that given any $\Delta \subset \{1, \ldots, N\}$ with a minimum separation distance of $1/M$, if $\Omega \subset [M]$ consists m indices chosen uniformly at random with

$$m \gtrsim \max\left\{\log^2(M/\epsilon), |\Delta| \log(|\Delta|/\epsilon)\log(M/\epsilon)\right\},$$

then the following hold with probability exceeding $1 - \epsilon$.

1. For all $x \in \mathbb{C}^N$,

$$m^{-1/2}\|P_\Omega A P_\Delta x\|_2 \geq \frac{3}{2\sqrt{5}}\|P_\Delta x\|_2,$$

2. There exists $\rho = m^{-1/2}A^* P_\Omega w$ with $\|w\| \lesssim \sqrt{|\Delta|}$ such that

$$P_\Delta \rho = x_0, \qquad \left\|P_\Delta^\perp \rho\right\|_\infty \leq c(M)$$

   where

$$c(M) := \max\left\{0.99993, 1 - \frac{0.92(M^2-1)}{N^2}\right\}.$$

Furthermore, these two conditions hold with probability 1 if $\Omega = [M_1]$.

Therefore, by applying the above fact T_1 times and applying the union bound, conditions (i) and (iii) of Proposition 4.2 (with $c_1^{-1} = 2\sqrt{5}/3$, $c_2 = c(M_1)$ and $L \lesssim \sqrt{s_1}$) hold with probability exceeding $1 - T_1\epsilon$ provided that Ω_1 is chosen uniformly at random with

$$\Omega_1 \subset [M_1], \qquad |\Omega_1| \gtrsim \max\left\{\log^2(M_1/\epsilon), s_1\log(s_1/\epsilon)\log(M_1/\epsilon)\right\}, \tag{4.9}$$

and they hold with probability 1 if $\Omega_1 = [M_1]$. Similarly, conditions (ii) and (iv) of Proposition 4.2 (with $c_1^{-1} = 2\sqrt{5}/3$, $c_2 = c(M_2)$ and $L \lesssim \sqrt{s_2}$) hold with probability exceeding $1 - T_2\epsilon$ if Ω_2 is chosen uniformly at random with

$$\Omega_2 \subset [M_2], \qquad |\Omega_2| \gtrsim \max\left\{\log^2(M_2/\epsilon), s_1\log(s_2/\epsilon)\log(M_2/\epsilon)\right\}, \tag{4.10}$$

and they hold with probability 1 if $\Omega_2 = [M_2]$. So, by applying the union bound once more, conditions (i) to (iv) with $c_1^{-1} = 2\sqrt{5}/3$, $c_2 = \max\{c(M_1), c(M_2)\}$ and $L \lesssim \max\left\{\sqrt{s_1}, \sqrt{s_2}\right\}$ are satisfied with probability exceeding $1 - T_1\epsilon - T_2\epsilon$ provided that Ω_1 and Ω_2 are chosen uniformly at random such that (4.9) and (4.10) hold. $\square$

5 Conclusion

In this paper, we have derived recovery guarantees for total variation regularized solutions when given partial measurements of the Fourier transform taken along Cartesian lines. In particular, we established a link between the sparsity structure and the sampling pattern by proving that the number of Cartesian lines required for accurate recovery is dependent not only on the gradient sparsity of the underlying vector, but also on the separation distance between the discontinuities.

6 Acknowledgements

The author acknowledges support from the Fondation Science Mathematique de Paris.

References

[1] 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.

[2] C. Boyer, J. Bigot, and P. Weiss. Compressed sensing with structured sparsity and structured acquisition. *arXiv preprint arXiv:1505.01619*, 2015.

[3] E. J. Candès and C. Fernandez-Granda. Towards a mathematical theory of super-resolution. *Communications on Pure and Applied Mathematics*, 67(6):906–956, 2014.

[4] E. J. Candes and Y. Plan. A probabilistic and ripless theory of compressed sensing. *Information Theory, IEEE Transactions on*, 57(11):7235–7254, 2011.

[5] E. J. Candès, J. Romberg, and T. Tao. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *Information Theory, IEEE Transactions on*, 52(2):489–509, 2006.

[6] W. Cong, J. Yang, and G. Wang. Differential phase-contrast interior tomography. *Physics in medicine and biology*, 57(10):2905, 2012.

[7] A. Gonzalez, L. Jacques, C. De Vleeschouwer, and P. Antoine. Compressive optical deflectometric tomography: A constrained total-variation minimization approach. *Inverse Problems & Imaging*, 8(2), 2014.

[8] F. Krahmer and R. Ward. Stable and robust sampling strategies for compressive imaging. *Image Processing, IEEE Transactions on*, 23(2):612–622, 2014.

[9] R. Leary, Z. Saghi, P. A. Midgley, and D. J. Holland. Compressed sensing electron tomography. *Ultramicroscopy*, 131:70–91, 2013.

[10] M. Lustig, D. Donoho, and J. M. Pauly. Sparse mri: The application of compressed sensing for rapid mr imaging. *Magnetic resonance in medicine*, 58(6):1182–1195, 2007.

[11] M. Lustig, D. L. Donoho, J. M. Santos, and J. M. Pauly. Compressed Sensing MRI. *IEEE Signal Process. Mag.*, 25(2):72–82, March 2008.

[12] C. Poon. A consistent and stable approach to generalized sampling. *Journal of Fourier Analysis and Applications*, pages 1–35, 2014.

[13] C. Poon. On the role of total variation in compressed sensing. *SIAM Journal on Imaging Sciences*, 8(1):682–720, 2015.

[14] G. Tang, B. N. Bhaskar, P. Shah, and B. Recht. Compressive sensing off the grid. In *Communication, Control, and Computing (Allerton), 2012 50th Annual Allerton Conference on*, pages 778–785. IEEE, 2012.

[15] H. Wang, D. Liang, and L. Ying. Pseudo 2d random sampling for compressed sensing mri. In *Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE*, pages 2672–2675. IEEE, 2009.

[16] Y. Wiaux, L. Jacques, G. Puy, A. Scaife, and P. Vandergheynst. Compressed sensing imaging techniques for radio interferometry. *Monthly Notices of the Royal Astronomical Society*, 395(3):1733–1742, 2009.