On compressed sensing with generative neural networks and Fourier measurements

Aaron Berk

Postdoctoral researcher
Department of Mathematics & Statistics
McGill University

CRM Applied Math Seminar
23 January 2023

Acknowledgment

Simone Brugiapaglia
Mathematics & Statistics
Concordia University
Montréal, QC, Canada

Tim Hoheisel
Mathematics & Statistics
McGill University
Montréal, QC, Canada

Babhru Joshi, Matthew Scott, Yaniv Plan & Özgür Yılmaz
Mathematics
University of British Columbia
Vancouver, BC, Canada

CENTRE DE RECHERCHES MATHÉMATIQUES
IVADO
Natural Sciences and Engineering Research Council of Canada
Conseil de recherches en sciences naturelles et en génie du Canada
Canada

Some motivation

FDA Clears Compressed Sensing MRI Acceleration Technology From Siemens Healthineers

2017-02-21

· New technology employs iterative reconstruction to produce high-quality MR images at a rapid rate with zero diagnostic information loss

· Compressed Sensing Cardiac Cine – the technology's first application – enables diagnostic cardiac imaging of patients with arrhythmias or respiratory problems

Siemens Healthineers has announced that the Food and Drug Administration (FDA) has cleared the company's revolutionary Compressed Sensing technology, which slashes the long acquisition times associated with magnetic resonance imaging (MRI) to enable dramatically shortened scans. For example, cardiac cine imaging with Compressed Sensing can be completed in just 16 seconds rather than the traditional four minutes, courtesy of an algorithm that reduces the required amount of data.[1] With products and solutions such as Compressed Sensing, as well as a new name that underscores the company's pioneering spirit and engineering expertise, Siemens Healthineers – the separately managed healthcare business of Siemens AG – is helping to enable healthcare providers worldwide to meet current challenges and excel in their respective environments.

Figure: FDA approval for Siemens CS-MRI;
acquisition time *ca.* 2017: 4 minutes $\rightarrow$ 16 seconds ($15\times$ speed-up)

Standard set-up

In standard compressed sensing

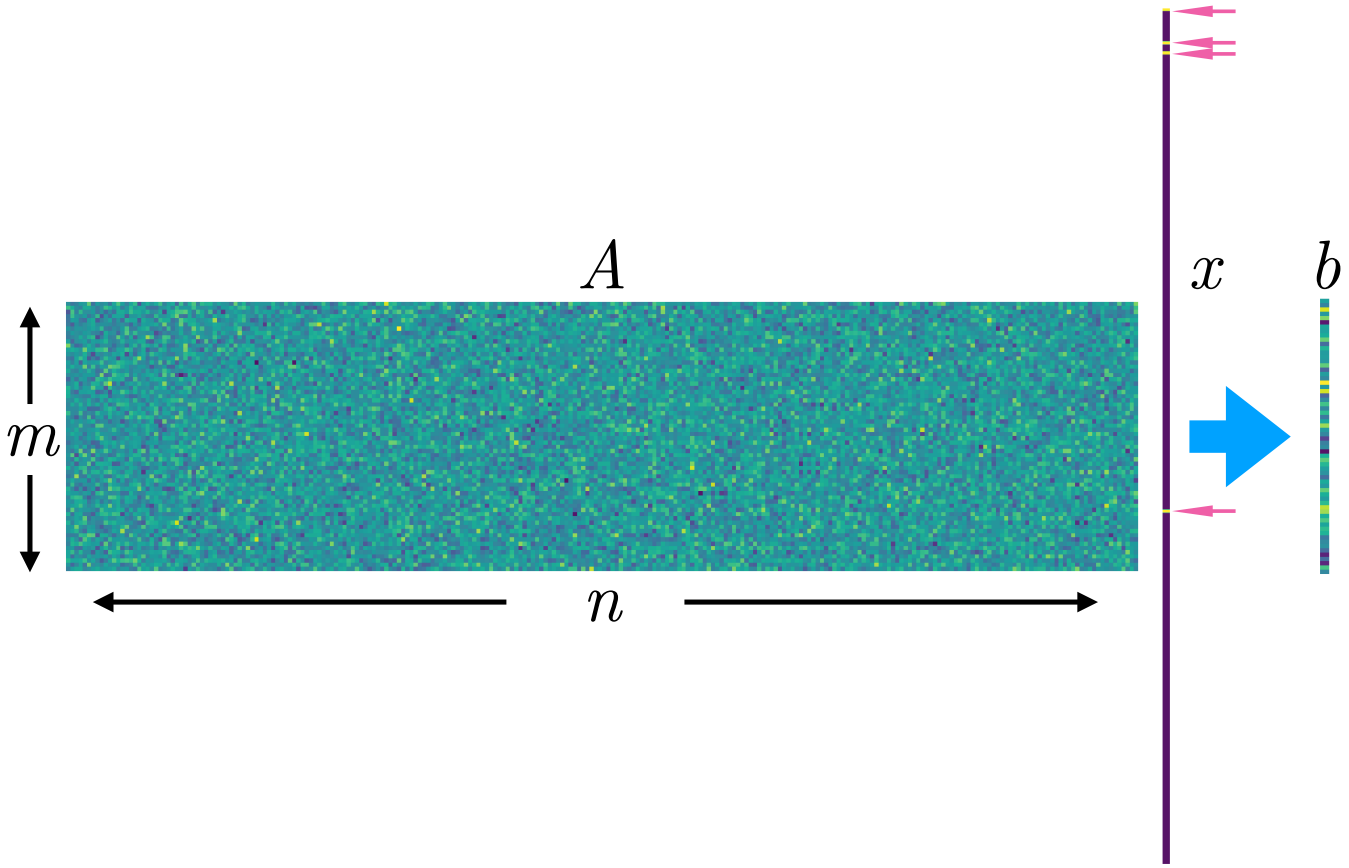
- $A \in \mathbb{C}^{m \times n}, m \ll n$, *"incoherent rows"*
  $\rightarrow$ **random measurements**: *e.g.*, A has iid Gaussian entries
- $x_0 \in \mathbb{R}^N$ is s **sparse** (in a basis) and **unknown**
  $\rightarrow$ **structural proxy**: ℓ_1 norm
- A and b are **known**:

$$b = Ax_0 + \eta, \quad \eta \in \mathbb{C}^m$$

Goal: `optimize` to recover x_0 from (b, A).

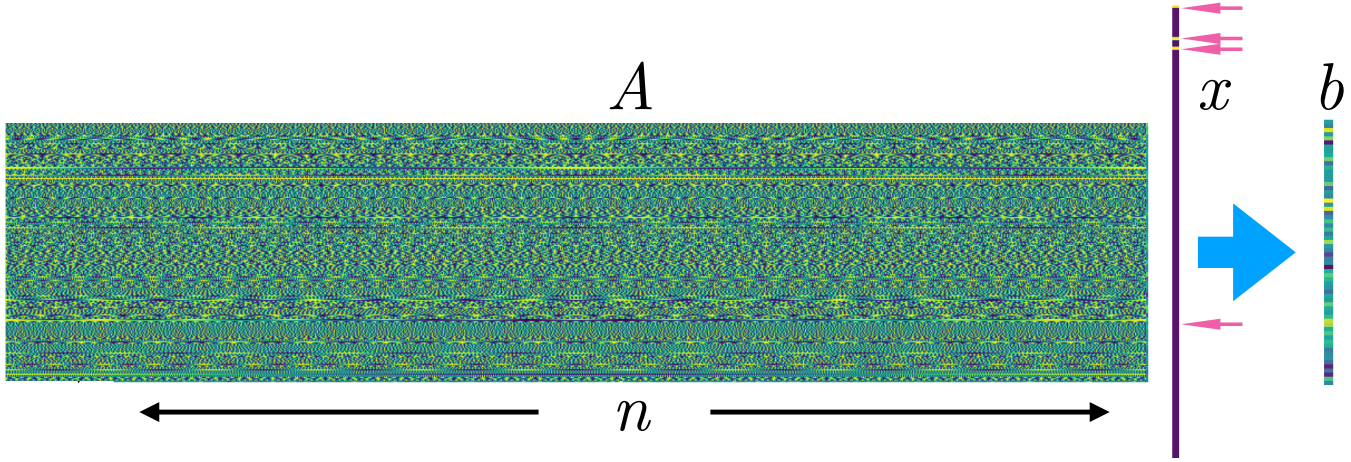
Practical sensing matrices

CS has a clean, well-developed theory for Gaussian sensing matrices, which are **impractical** for applications like MRI.



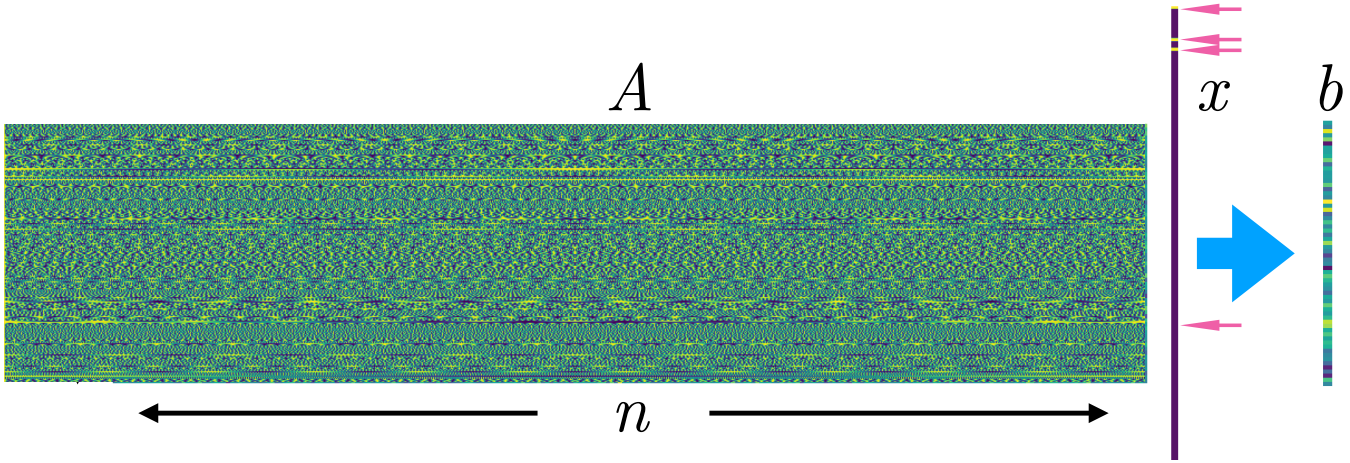
Practical sensing matrices

At the least, we want **structured** random matrices, like **subsampled isometries**.



Practical sensing matrices

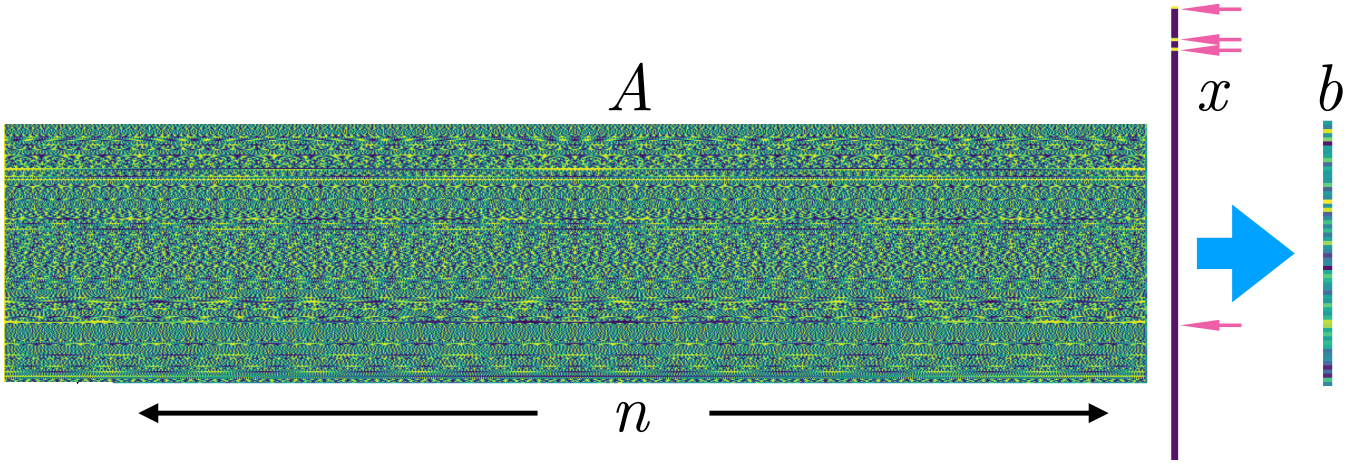
The best known sample complexity for **structured** random matrices is $m \gtrsim s \log s \log^2 n$ (due to Bourgain [4]).



Candès et al. [5], Donoho [6], Rauhut [11], Rudelson and Vershynin [12]

Practical sensing matrices

Further developments and theory for even more realistic sensing matrices exists (*e.g.*, see Adcock and Hansen [1]).



Candès et al. [5], Donoho [6], Rauhut [11], Rudelson and Vershynin [12]

Compressed sensing with generative models

An ingenuity of Bora et al. [3] replaced the **structural proxy** of sparsity/ℓ_1 norm with a **generative neural network** (GNN).

Compressed Sensing using Generative Models

Ashish Bora [1] **Ajil Jalal** [2] **Eric Price** [1] **Alexandros G. Dimakis** [2]

Abstract

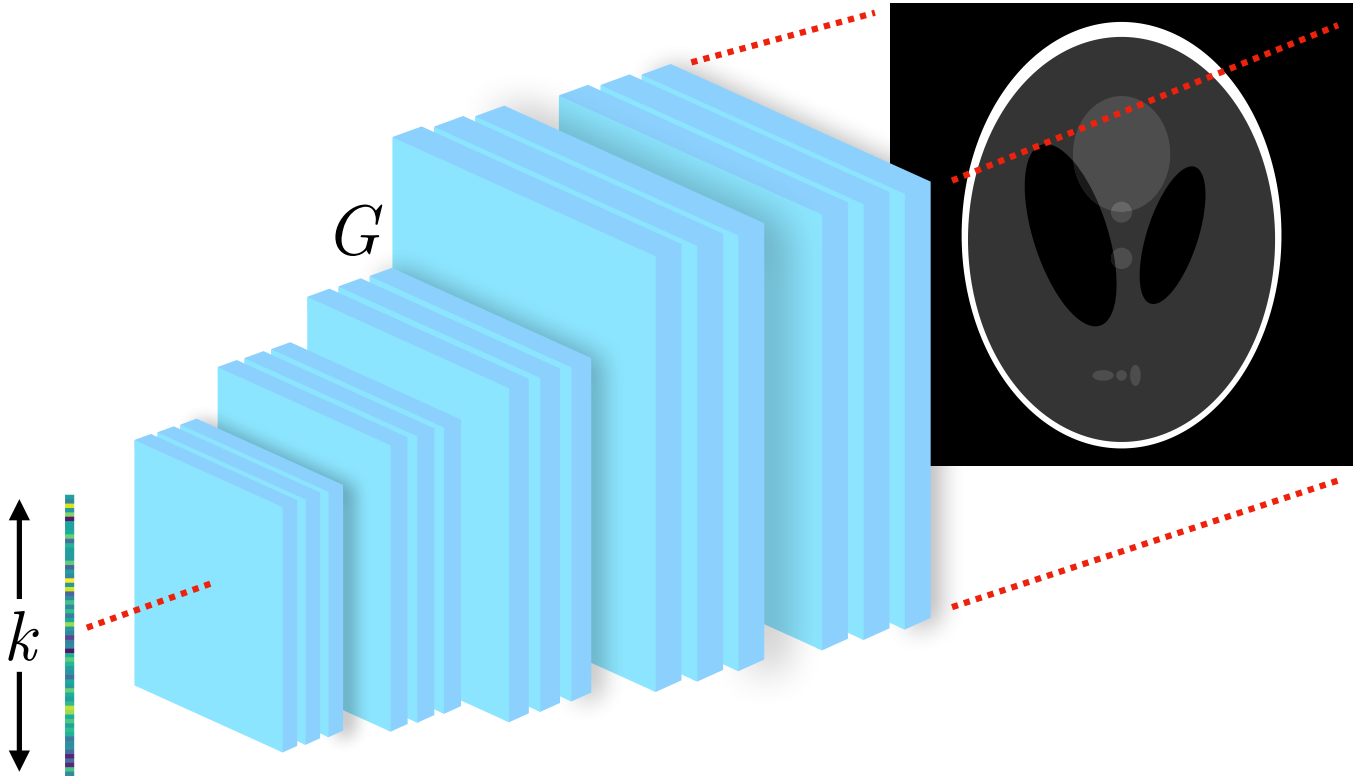
The goal of compressed sensing is to estimate a vector from an underdetermined system of noisy linear measurements, by making use of prior knowledge on the structure of vectors in the relevant domain. For almost all results in this literature, the structure is represented by sparsity

of the unknown vector x^*. We need to assume that the unknown vector is "natural," or "simple," in some application-dependent way.

The most common structural assumption is that the vector x^* is k-sparse in some known basis (or approximately k-sparse). Finding the sparsest solution to an underdetermined system of linear equations is NP-hard, but still con-

Compressed sensing with generative models

These networks characterize structure by mapping low-dimensional latent codes to high-dimensional signals.



Why use GNNs?

GNNs are effective at capturing the intrinsic structure of natural images.

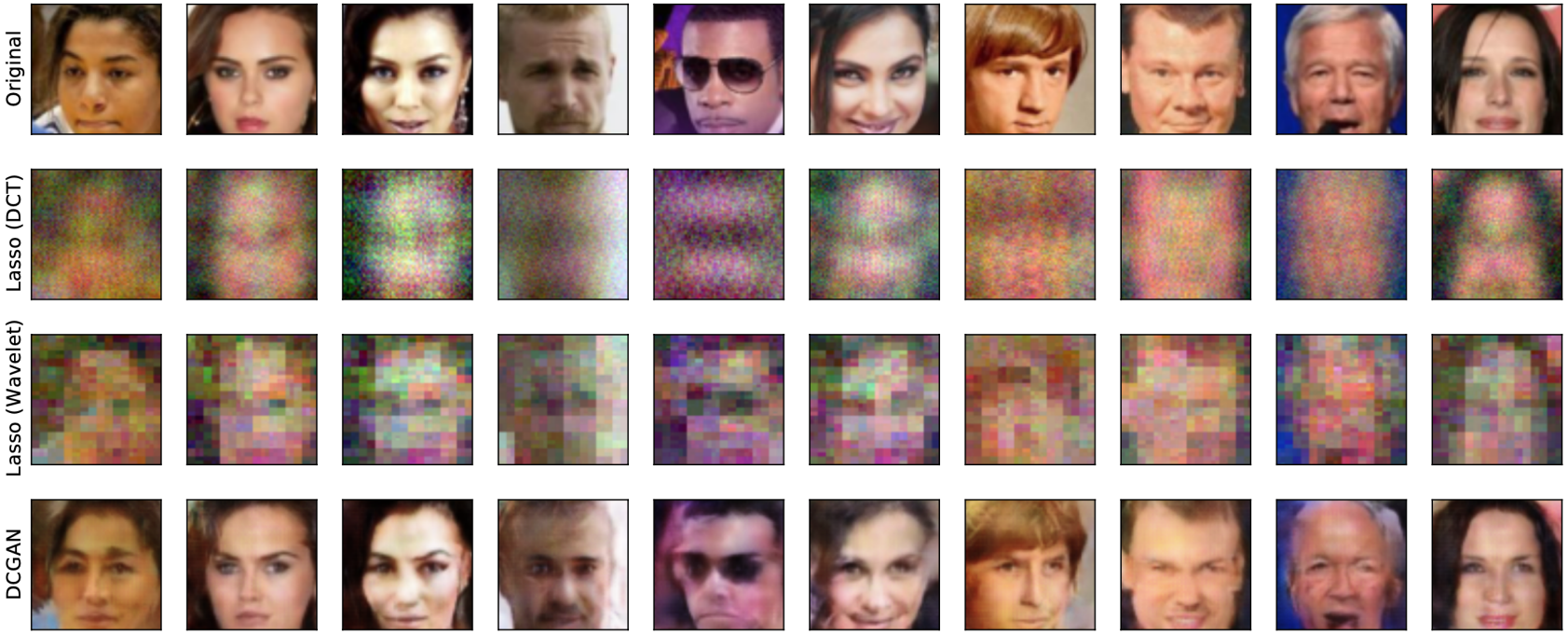


Figure: Bora et al. [3, Figure 3]: $m = 500, n = 12288$

Why use GNNs?

Applications to:

- inpainting
- compressed sensing
- super-resolution

Promising empirical success in:

- Image modeling applications
  (Goodfellow et al. [7], Kingma and Welling [9])
- MRI applications
  (Jalal et al. [8])
- parametric PDE
  (Wentz and Doostan [15])

Why use GNNs?

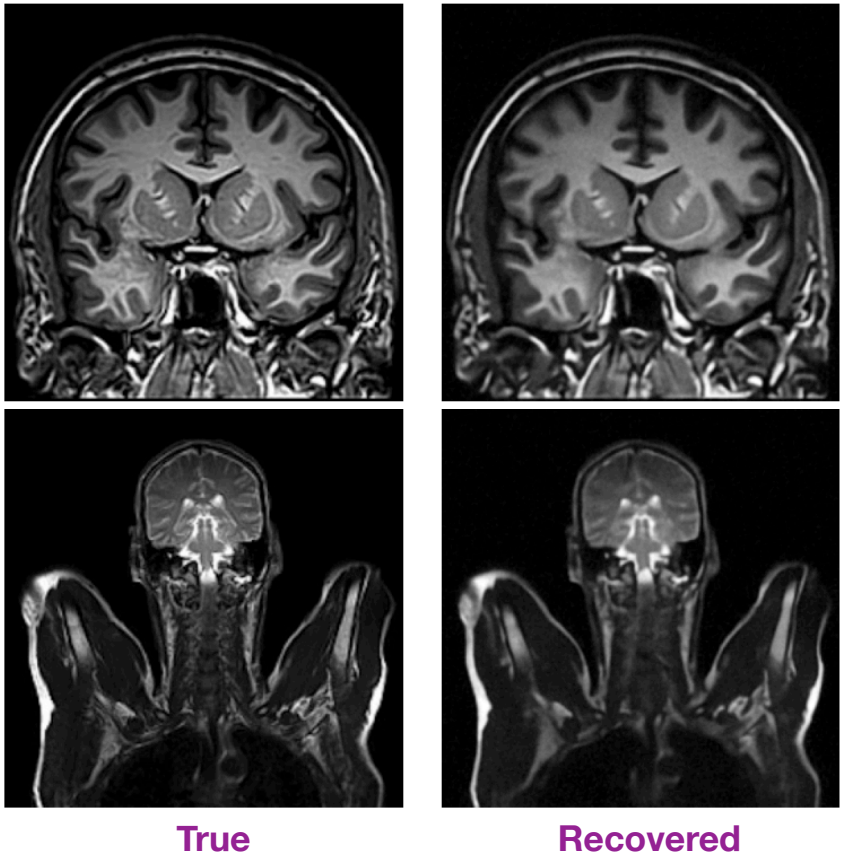


Figure: Created via Ryu et al. [13, Section 6]

Compressed sensing with generative models

An ingenuity of Bora et al. [3] replaced the **structural proxy** of sparsity/ℓ_1 norm with a **generative neural network** (GNN).

Compressed Sensing using Generative Models

Ashish Bora [1] **Ajil Jalal** [2] **Eric Price** [1] **Alexandros G. Dimakis** [2]

Abstract

The goal of compressed sensing is to estimate a vector from an underdetermined system of noisy linear measurements, by making use of prior knowledge on the structure of vectors in the relevant domain. For almost all results in this literature, the structure is represented by sparsity

of the unknown vector x^*. We need to assume that the unknown vector is "natural," or "simple," in some application-dependent way.

The most common structural assumption is that the vector x^* is k-sparse in some known basis (or approximately k-sparse). Finding the sparsest solution to an underdetermined system of linear equations is NP-hard, but still con-

Compressed sensing with generative models

However, their [3] sample complexity guarantee for CS with GNNs effectively requires A to be Gaussian.

Theorem 1.1. *Let $G : \mathbb{R}^k \to \mathbb{R}^n$ be a generative model from a d-layer neural network using ReLU activations. Let $A \in \mathbb{R}^{m \times n}$ be a random Gaussian matrix for $m = \mathcal{O}(kd \log n)$, scaled so $A_{i,j} \sim N(0, 1/m)$. For any $x^* \in \mathbb{R}^n$ and any observation $y = Ax^* + \eta$, let $\widehat{z}$ minimize $\|y - AG(z)\|_2$ to within additive ϵ of the optimum. Then with $1 - e^{-\Omega(m)}$ probability,*

$$\|G(\widehat{z}) - x^*\|_2 \leq 6 \min_{z^* \in \mathbb{R}^k} \|G(z^*) - x^*\|_2 + 3\|\eta\|_2 + 2\epsilon.$$

The goal of co
vector from an
linear measur
knowledge on
evant domain.
erature, the structure is represented by sparsity

assume that the un-
n some application-
ion is that the vec-
s (or approximately
n to an underdeter-
mined system of linear equations is NP-hard, but still con-

CS with GNNs and structured random matrices

Setup: For $b = Ax_0 + \eta$ and $G : \mathbb{R}^k \to \mathbb{R}^n$ a ReLU GNN, find $\hat{x} \in \mathcal{R}(G)$ approximately solving

$$\min_{x \in \mathcal{R}(G)} \|Ax - b\|_2 \quad \leftrightarrow \quad \min_{z \in \mathbb{R}^k} \|AG(z) - b\|_2.$$

Goal:

Sample complexity for theoretical recovery guarantee with realistic sampling matrix

A coherence parameter characterizing generative compressed sensing with Fourier measurements

Aaron Berk, Simone Brugiapaglia, Babhru Joshi, Yaniv Plan, Matthew Scott and Özgür Yilmaz

***Abstract*—In [1], a mathematical framework was developed for compressed sensing guarantees in the setting where the measurement matrix is Gaussian and the signal structure is the range of a generative neural network (GNN). The problem of compressed sensing with GNNs has since been extensively analyzed when the measurement matrix and/or network weights follow a subgaussian distribution. We move beyond the subgaussian assumption.**

(*e.g.,* sparsity) [5], [7]. Moreover, this recovery is effected using an order-optimal number of random measurements [7]. In applications like medical imaging [5], the measurement matrices under consideration are derived from a bounded orthonormal system (a unitary matrix with bounded entries), which complicates the theoretical analysis. Furthermore, for

(AB, Brugiapaglia, Joshi, Plan, Scott, and Yilmaz, 2022 [2])

Theorem 1: Subsampled isometry GCS

Let $G : \mathbb{R}^k \to \mathbb{R}^n$ be a (k, d, n)-generative network with layer widths $k = k_0 \leq k_1, \ldots, k_d$ where $k_d := n$, $\varepsilon, \hat{\varepsilon} > 0$, $\mathcal{G} := \mathcal{R}(G) - \mathcal{R}(G)$ and $A \in \mathbb{C}^{\tilde{m} \times n}$ a subsampled isometry associated with a unitary matrix $U \in \mathbb{C}^{n \times n}$. If $\Delta(\mathcal{G})$ is α-coherent with respect to $\| \cdot \|_U$, and

$$m \gtrsim \alpha^2 n \left(2k \sum_{i=1}^{d-1} \log \left(\frac{2ek_i}{k} \right) + \log \frac{4k}{\varepsilon} \right),$$

then, with probability at least $1 - \varepsilon$ on the realization of A:
For any $x_0 \in \mathbb{R}^n$, let $b := Ax_0 + \eta$ where $\eta \in \mathbb{C}^{\tilde{m}}$. Let $\hat{x} \in \mathbb{R}^n$ satisfy $\|A\hat{x} - b\|_2 \leq \min_{x \in \mathcal{R}(G)} \|Ax - b\|_2 + \hat{\varepsilon}$. Then,

$$\|\hat{x} - x_0\|_2 \leq \|x^\perp\|_2 + 3\|Ax^\perp\|_2 + 3\|\eta\|_2 + \frac{3}{2}\hat{\varepsilon}.$$

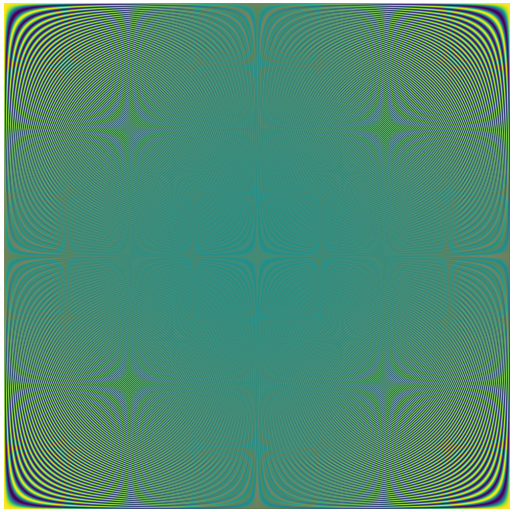
(where $x^\perp := x_0 - \Pi_{\mathcal{R}(G)}(x_0)$)

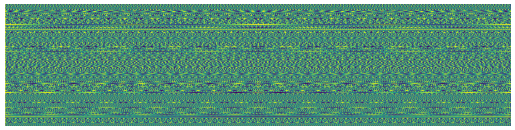
(AB, Brugiapaglia, Joshi, Plan, Scott, and Yilmaz, 2022 [2])

Subsampled isometries

Definition (subsampled isometry)

Let $n \geq m \geq 2$ be integers and $U \in \mathbb{C}^{n \times n}$ a unitary matrix. Let $\theta := (\theta_i)_{i \in [n]}$, $\theta_i \overset{\text{iid}}{\sim} \mathrm{Ber}(m/n)$. Define a subsampled isometry as $A := \frac{\sqrt{n}}{\sqrt{m}} \operatorname{rowstack}[U_i : \theta_i = 1]$.





Generative networks

Definition (Generative network)

Let $G : \mathbb{R}^k \to \mathbb{R}^n$ with $k < n$. Call G a generative network (GNN) with depth $d \in \mathbb{N}$ if, for $u \in \mathbb{R}^k$,

$$G(u) := W^{(d)}\sigma\left(W^{(d-1)}\sigma\left(\cdots W^{(1)}u\right)\right).$$

Here, $W^{(i)} \in \mathbb{R}^{k_i \times k_{i-1}}$ with $k = k_0 \leq k_1, \ldots, k_d = n$ and $\sigma(u) := \max(u, 0)$ element-wise.

Generative networks

Definition (Generative network)

Let $G : \mathbb{R}^k \to \mathbb{R}^n$ with $k < n$. Call G a generative network (GNN) with depth $d \in \mathbb{N}$ if, for $u \in \mathbb{R}^k$,

$$G(u) := W^{(d)}\sigma\left(W^{(d-1)}\sigma\left(\cdots W^{(1)}u\right)\right).$$

Here, $W^{(i)} \in \mathbb{R}^{k_i \times k_{i-1}}$ with $k = k_0 \leq k_1, \ldots, k_d = n$ and $\sigma(u) := \max(u, 0)$ element-wise.

- $W^{(i)}$ are *weight matrices* and σ is the *activation function*.
- It's mild to consider no biases — one can write $\tilde{W}^{(i)}\tilde{z}^{(i)} := \begin{bmatrix} W^{(i)} & b^{(i)} \end{bmatrix} \begin{bmatrix} z^{(i)} \\ 1 \end{bmatrix} = W^{(i)}z^{(i)} + b^{(i)}$
- The approach requires piece-wise linear σ

Assumption: (in)coherence

Definition (measurement norm)

Let $U \in \mathbb{C}^{n \times n}$ be a unitary matrix. Define the norm $\|\cdot\|_U : \mathbb{C}^n \to [0, \infty)$ by

$$\|x\|_U := \|Ux\|_\infty = \max_{i \in [n]} |\langle U_i, x \rangle|$$

Definition (coherence)

Let $T \subseteq \mathbb{R}^n$ be a set and $U \in \mathbb{C}^{n \times n}$ a unitary matrix. For $\alpha > 0$ say that T is α-coherent with respect to $\|\cdot\|_U$ if

$$\sup_{x \in T \cap \mathbb{S}^{n-1}} \|x\|_U \leq \alpha.$$

Assumption: (in)coherence

Definition (coherence)

Let $T \subseteq \mathbb{R}^n$ be a set and $U \in \mathbb{C}^{n \times n}$ a unitary matrix. For $\alpha > 0$ say that T is α-coherent with respect to $\|\cdot\|_U$ if

$$\sup_{x \in T \cap \mathbb{S}^{n-1}} \|x\|_U \leq \alpha.$$

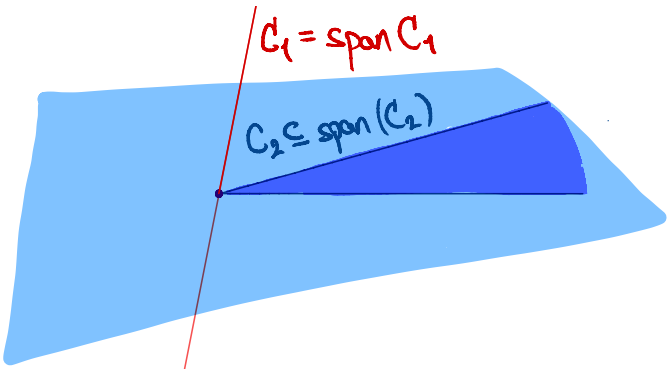
- Bounds for α should depend on:
  - Complexity of low-dimensional objects comprising T
  - Geometry between $\|\cdot\|_U$ and T.

Piecewise linear expansion

Definition

Let $\mathcal{C} = \bigcup_{i=1}^{N} \mathcal{C}_i$ be the union of N convex cones. Define

$$\Delta(\mathcal{C}) := \bigcup_{i=1}^{N} \operatorname{span}(\mathcal{C}_i) = \bigcup_{i=1}^{N} \mathcal{C}_i - \mathcal{C}_i.$$



- If $\mathcal{C}$ is a union of N cones then $\Delta\mathcal{C}$ is a union of at most N at-most k-dimensional subspaces.
- $\Delta(\mathcal{C})$ satisfies

$$\mathcal{C} \subseteq \Delta\mathcal{C} \subseteq \mathcal{C} - \mathcal{C}.$$

- Note $\Delta(\mathcal{C})$ is uniquely defined.
- If G is a generative network, then $\mathcal{R}(G)$ is a union of polyhedral cones.

Theorem 1: Subsampled isometry GCS

Let $G : \mathbb{R}^k \to \mathbb{R}^n$ be a (k, d, n)-generative network with layer widths $k = k_0 \leq k_1, \ldots, k_d$ where $k_d := n$, $\varepsilon, \hat{\varepsilon} > 0$, $\mathcal{G} := \mathcal{R}(G) - \mathcal{R}(G)$ and $A \in \mathbb{C}^{\tilde{m} \times n}$ a subsampled isometry associated with a unitary matrix $U \in \mathbb{C}^{n \times n}$. If $\Delta(\mathcal{G})$ is α-coherent with respect to $\| \cdot \|_U$, and

$$m \gtrsim \alpha^2 n \left(2k \sum_{i=1}^{d-1} \log \left(\frac{2ek_i}{k} \right) + \log \frac{4k}{\varepsilon} \right),$$

then, with probability at least $1 - \varepsilon$ on the realization of A:
For any $x_0 \in \mathbb{R}^n$, let $b := Ax_0 + \eta$ where $\eta \in \mathbb{C}^{\tilde{m}}$. Let $\hat{x} \in \mathbb{R}^n$ satisfy $\|A\hat{x} - b\|_2 \leq \min_{x \in \mathcal{R}(G)} \|Ax - b\|_2 + \hat{\varepsilon}$. Then,

$$\|\hat{x} - x_0\|_2 \leq \|x^\perp\|_2 + 3\|Ax^\perp\|_2 + 3\|\eta\|_2 + \frac{3}{2}\hat{\varepsilon}.$$

(where $x^\perp := x_0 - \Pi_{\mathcal{R}(G)}(x_0)$)

(AB, Brugiapaglia, Joshi, Plan, Scott, and Yilmaz, 2022 [2])

Proof outline I

Restricted **I**sometry **P**roperty for GNNs: when $\Delta(\mathcal{G})$ is α-coherent wrt $\|\cdot\|_U$ and $m = \tilde{\Omega}(\alpha^2 n \delta^{-2} kd)$, then with high probability on A,

$$\sup_{x \in \Delta(\mathcal{G}) \cap \mathbb{S}^{n-1}} |\|Ax\|_2 - 1| \leq \delta$$

RIP proof strategy

$\mathcal{G} = \mathcal{R}(G) - \mathcal{R}(G)$ is a union of cones. Apply **matrix Bernstein inequality** to control deviation on each minimal containing subspace. Augment to a **subgaussian tail bound** as in Vershynin [14, Theorem 3.1.1]. **Union bound** over components of $\Delta(\mathcal{G})$.

Proof outline II

The number of components N in $\Delta(\mathcal{G})$ is nicely controlled (*e.g.*, see Naderi and Plan [10]).

Lemma

Let G be a generative network. Then $\mathcal{R}(G)$ is a union of no more than N at-most k-dimensional polyhedral cones where

$$N := \prod_{\ell=1}^{d-1} I(k_\ell, k) \leq \left(\frac{2e\bar{k}}{k}\right)^{k(d-1)}, \quad \bar{k} := \left(\prod_{\ell=1}^{d-1} k_\ell\right)^{1/(d-1)}.$$

Typical coherence I

Proposition

Let $U \in \mathbb{C}^{n \times n}$ be a unitary matrix. Any k-dimensional subspace $T \subseteq \mathbb{R}^n$ is at least $\sqrt{\frac{k}{n}}$-coherent with respect to $\| \cdot \|_U$. Furthermore, this lower bound is tight.

Remark

The sample complexity for the main result is $m = \tilde{\Omega}(k^2)$.

Typical coherence II

Theorem

Let $U \in \mathbb{C}^{n\times n}$ be a unitary matrix and G be a (k,d,n)-generative network. Let the last weight matrix of G, $W^{(d)}$, be iid Gaussian: $W_{ij}^{(d)} \overset{iid}{\sim} \mathcal{N}(0,1)$, $i \in [k_d], j \in [k_{d-1}]$. Let all other weights be arbitrary and fixed. Then, with probability at least $1 - 2\exp(-\gamma^2)$, $\Delta(\mathcal{R}(G) - \mathcal{R}(G))$ is α-coherent with respect to $\|\cdot\|_U$, where

$$\alpha \lesssim \sqrt{\frac{k}{n}} + \sqrt{\frac{\log n}{n}} + \sqrt{\frac{k}{n}\sum_{i=1}^{d-1}\log\frac{2ek_i}{k}} + \frac{\gamma}{\sqrt{n}}.$$

Typical coherence III

Remark

Under mild simplifying assumptions this expression gives

$$\alpha = \mathcal{O}\left(\sqrt{\frac{kd}{n}}\right)$$

Thus, the sample complexity for the main result for a GNN with Gaussian weights may be bounded as

$$m \gtrsim \frac{2k^2 d}{\delta^2} \sum_{i=1}^{d-1} \log\left(\frac{2ek_i}{k}\right) + \frac{kd}{\delta^2} \log \frac{4k}{\varepsilon}.$$

Method: coherence upper bound I

Proposition

Let $U \in \mathbb{R}^{n \times n}$ be an orthogonal matrix. Let $W = W^{(d)}$ be the final weight matrix of a generative network $G : \mathbb{R}^k \to \mathbb{R}^n$ and let

$$W = QR; \qquad Q := \begin{bmatrix} Q_1 & Q_2 \end{bmatrix}, \qquad R := \begin{bmatrix} R_1 \\ 0 \end{bmatrix}$$

be its QR decomposition. If $\mathcal{G} := \mathcal{R}(G) - \mathcal{R}(G)$ and $\mathcal{W} := \mathcal{R}(W) \cap \mathbb{S}^{n-1}$, then

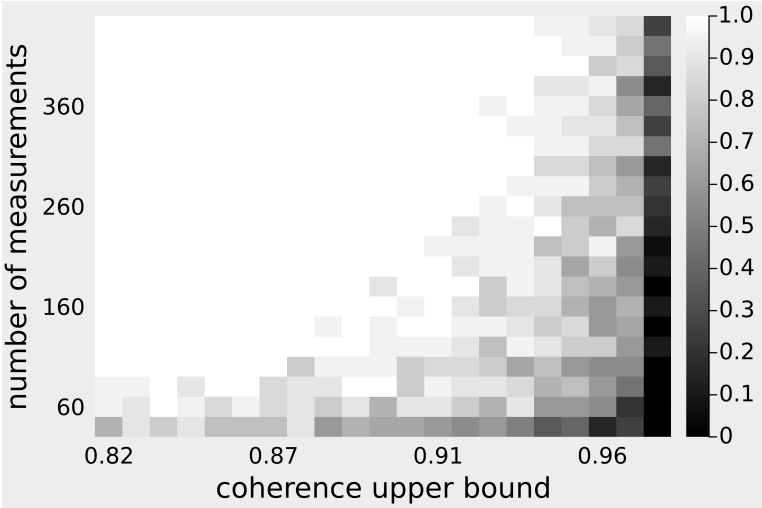
$$\sup_{x \in \Delta(\mathcal{G}) \cap \mathbb{S}^{n-1}} \|Ux\|_\infty \leq \sup_{x \in \mathcal{W}} \|Ux\|_\infty = \|UQ_1\|_{2 \to \infty}$$

Method: coherence upper bound II

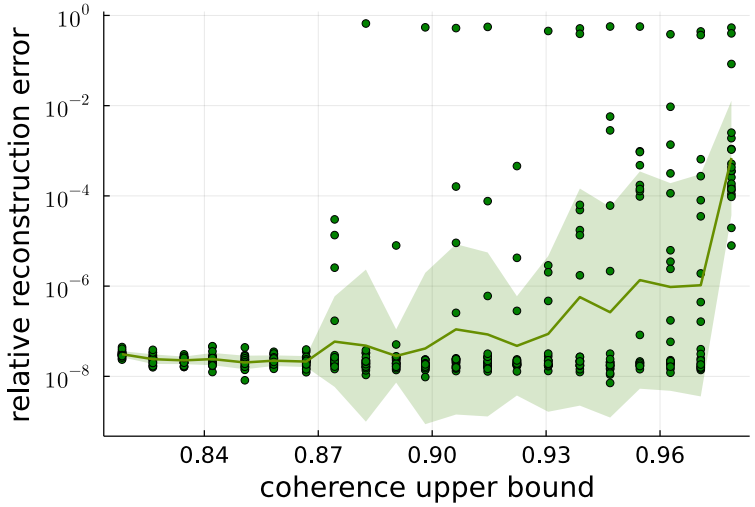
Define the following regularizer used to promote low coherence during training:

$$\rho(W) := \|UW\|_{2\to\infty} + \lambda\|W^\top W - I\|_F.$$

Numerical results I



(a) Empirical recovery probability as a function of coherence and m. Each block corresponds to the average from 20 independent trials. White corresponds with 20 successful recoveries (rre $\leq 10^{-5}$); black with no successful recoveries.



(b) Empirical rre as a function of coherence for $m = 100$. Each dot corresponds to one of 20 trials at each coherence level. The solid line shows the empirical geometric mean rre *vs.* coherence upper bound. The envelope shows 1 geometric standard deviation.

Numerical results II

Figure: Recovery comparison of MNIST images for various measurement sizes m (denoted by column heading) for a low coherence network, high-coherence network and network with final sigmoid activation. In each block: the top row corresponds to $G^{(3)}$ ($\alpha = 0.82$); middle row $G^{(2)}$ ($\alpha = 0.96$); bottom row $G^{(1)}$ (labelled Sig). The leftmost column, signal, corresponds to the target image $x_0^{(i)} \in \mathcal{R}(G^{(i)})$.

Numerical results III

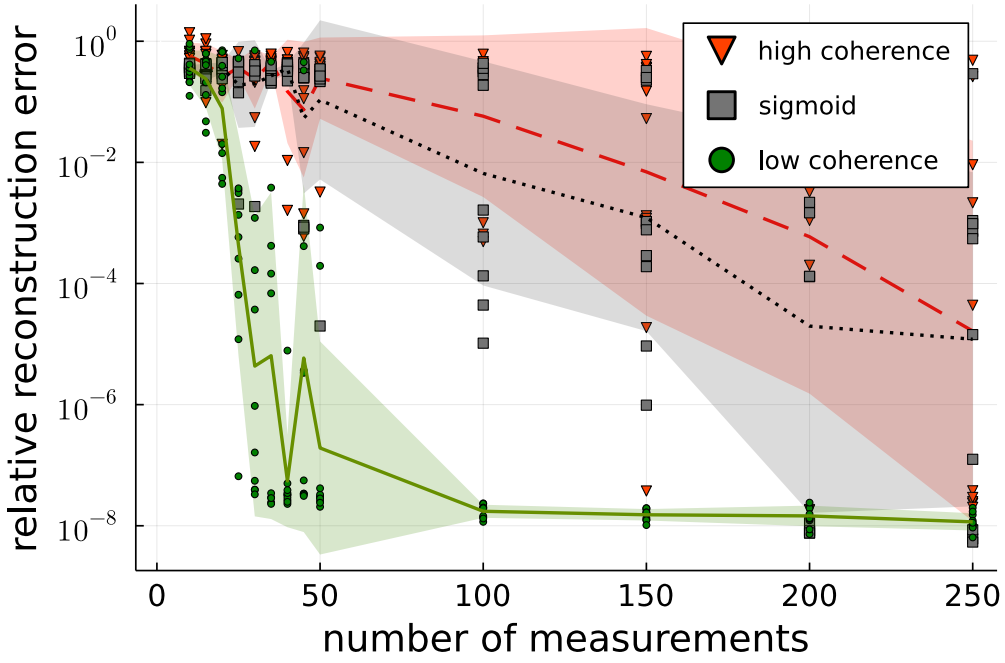


Figure: Performance comparison for three GNNs trained on the MNIST dataset — one with low coherence, another with high coherence, the last with sigmoid activation on the last layer. Plotting m *vs.* rre. For each value of m, each dot corresponds to one of 10 trials.

Contributions

- **First theory for GCS with subsampled isometries and non-random weights**
- Introduce *coherence* for characterizing recovery efficacy via alignment of the network's range and the measurement matrix
- Establish a restricted isometry property for ReLU GNNs with subsampled isometries
- Prove sample complexity and recovery bounds in this setting
- Propose regularization strategy for training GNNs with low coherence; demonstrate improved sample complexity for recovery
- Our theory and compelling numerical simulations support coherence as a natural quantity of interest linked to favourable deep generative recovery.

Future directions

- Optimal sample complexity *
  - Coherence-free argument
- Realism:
  - Guarantees for non-ReLU networks
  - Yet more realistic measurement matrices *
  - Optimal sampling strategy for given GNN
  - Incorporate knowledge of optimization strategy *

References I

[1] Ben Adcock and Anders C Hansen. *Compressive Imaging: Structure, Sampling, Learning*. Cambridge University Press, 2021.

[2] Aaron Berk, Simone Brugiapaglia, Babhru Joshi, Yaniv Plan, Matthew Scott, and Özgür Yilmaz. A coherence parameter characterizing generative compressed sensing with Fourier measurements. *IEEE Journal on Selected Areas in Information Theory (JSAIT)*, July 2022. doi:10.1109/JSAIT.2022.3220196. arXiv:2207.09340.

[3] Ashish Bora, Ajil Jalal, Eric Price, and Alexandros G Dimakis. Compressed sensing using generative models. In *International Conference on Machine Learning*, pages 537–546, 2017.

[4] Jean Bourgain. An improved estimate in the restricted isometry problem. In *Geometric aspects of functional analysis*, pages 65–70. Springer, 2014.

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

References II

[6] David L Donoho. Compressed sensing. *IEEE Transactions on information theory*, 52(4):1289–1306, 2006.

[7] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In *Advances in Neural Information Processing Systems*, pages 2672–2680, 2014.

[8] Ajil Jalal, Marius Arvinte, Giannis Daras, Eric Price, Alexandros G Dimakis, and Jonathan I Tamir. Robust compressed sensing MRI with deep generative priors. *arXiv preprint arXiv:2108.01368*, 2021.

[9] Diederik P Kingma and Max Welling. Auto-encoding variational Bayes. *arXiv preprint arXiv:1312.6114*, 2013.

[10] Alireza Naderi and Yaniv Plan. Beyond independent measurements: General compressed sensing with gnn application. *arXiv preprint arXiv:2111.00327*, 2021.

References III

[11] Holger Rauhut. Compressive sensing and structured random matrices. In *Theoretical foundations and numerical methods for sparse recovery*, pages 1–92. de Gruyter, 2010.

[12] Mark Rudelson and Roman Vershynin. On sparse reconstruction from fourier and gaussian measurements. *Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences*, 61(8):1025–1045, 2008.

[13] Ernest Ryu, Jialin Liu, Sicheng Wang, Xiaohan Chen, Zhangyang Wang, and Wotao Yin. Plug-and-play methods provably converge with properly trained denoisers. In *International Conference on Machine Learning*, pages 5546–5557. PMLR, 2019.

[14] Roman Vershynin. *High-dimensional probability: An introduction with applications in data science*, volume 47. Cambridge University Press, 2018.

[15] Jacqueline Wentz and Alireza Doostan. Genmod: A generative modeling approach for spectral representation of pdes with random inputs. *arXiv preprint arXiv:2201.12973*, 2022.