Local coherence sampling in compressed sensing

Felix Krahmer*, Holger Rauhut†, Rachel Ward‡,
* Institute for Numerical and Applied Mathematics, University of Göttingen, Germany
† Hausdorff Center for Mathematics, Bonn, Germany
‡ Mathematics Department, University of Texas at Austin, TX, USA

***Abstract*—Sparse recovery guarantees in compressive sensing and related optimization problems often assume incoherence between the 'sensing' and 'sparsity' domains. In practice, incoherence is rarely satisfied due to physical constraints and limitations. Here we discuss the notion of local coherence, and show that by matching the sampling strategy to the local coherence at hand, sparse recovery guarantees extend to a rich new class of sensing problems beyond incoherent systems. We discuss particular applications to compressive MRI imaging and polynomial interpolation.**

I. Introduction

One of the main results in the theory of compressed sensing is that signals which allow for an approximately sparse representation in a suitable basis or dictionary can be recovered from relatively few linear measurements via convex optimization, provided these measurements are sufficiently *incoherent* with the basis in which the signal is sparse.

In practice, incoherence is rarely satisfied due to physical constraints limiting the freedom of the sensing basis. Here we recall the notion of *local coherence*, as introduced in [6] and somewhat implicitly in [5], and summarize coherence-guided sampling strategies and reconstruction guarantees that extend beyond incoherent sampling. In short, local coherence sampling implies that, if Φ is an orthonormal basis from which we can subsample to construct a sensing matrix, and if our signal class is assumed sparse in an alternative orthonormal basis Ψ, then one should sample rows from Φ proportionately to their maximal correlation to any row from Ψ.

We illustrate the power of coherence-based sampling through two examples: compressed sensing imaging and and polynomial interpolation. In compressed sensing imaging, coherence-based sampling provides a theoretical justification for empirical studies [2], [3] pointing to variable-density sampling strategies for improved MRI compressive imaging. In polynomial interpolation, coherence-based sampling implies that sampling points drawn from the Chebyshev distribution are better suited for the recovery of polynomials and smooth functions than uniformly distributed sampling points, aligning with classical results on Lagrange interpolation [4].

II. Notation

Before continuing, let us fix some notation. We will refer to the set of natural numbers $\{1, 2, \ldots, N\}$ using the shorthand notation $[N]$. For a vector $x = (x_j) \in \mathbb{C}^N$, the usual ℓ_p vector norm is $\|x\|_p$, and by an abuse of notation, the ℓ_0-"norm" is defined as $\|x\|_0 = \#\{x_j : x_j \neq 0\}$. A vector $x \in \mathbb{C}^N$ is called s-sparse if $\|x\|_0 \leq s$, and the best s-term approximation of a vector $x \in \mathbb{C}^N$ is the s-sparse vector $x_s \in \mathbb{C}^N$ satisfying $x_s = \inf_{u:\|u\|_0 \leq s} \|x - u\|_p$. Clearly, $x_s = x$ if x is s-sparse. Informally, x is called compressible if $\|x - x_s\|$ decays quickly as s increases. Finally, for two nonnegative functions $f(n)$ and $g(n)$ on the natural numbers, we write $f \gtrsim g$ (or $f \lesssim g$) if there exists a constant $C > 0$ such that $f(n) \geq Cg(n)$ (or $f(n) \leq Cg(n)$, respectively) for all $n \in \mathbb{N}$.

III. Incoherent sampling

Here we recall sparse recovery results for structured random sampling schemes corresponding to *bounded orthonormal systems*, of which the partial discrete Fourier transform is a special case. We refer the reader to [7] for an expository article including many references.

Definition 1 (Bounded orthonormal system (BOS)): Let $\mathcal{D}$ be a measurable subset of $\mathbb{R}^d$.

- A set of functions $\{\psi_j : \mathcal{D} \to \mathbb{C}, \quad j \in [N]\}$ is called an *orthonormal system* with respect to the probability measure ν if $\int_{\mathcal{D}} \bar{\psi}_j(u)\psi_k(u)d\nu(u) = \delta_{jk}$, where δ_{jk} denotes the Kronecker delta.
- Let μ be a probability measure on $\mathcal{D}$. A *random sample* of the orthonormal system $\{\psi_j\}$ is the random vector $(\psi_1(U), \ldots, \psi_N(U))$ that results from drawing a sampling point U from the measure μ.
- An orthonormal system is said to be *bounded* with bound K if $\sup_{j\in[N]} \|\psi_j\|_\infty \leq K$.

Suppose now that we have an orthonormal system $\{\psi_j\}_{j\in[N]}$ and m random sampling points $U_1, U_2, \ldots, U_m$ drawn independently from some probability measure μ. Here and throughout, we assume that the number of sampling points $m \ll N$. As shown in [7], if the system $\{\psi_j\}$ is *bounded*, and if the probability measure μ from which we sample points is the orthogonalization measure ν associated to the system, then the (underdetermined) structured random matrix $A : \mathbb{C}^N \to \mathbb{C}^m$ whose rows are the independent random samples will be well-conditioned, satisfying the so-called *restricted isometry property* [1] with nearly order-optimal restricted isometry constants with high probability. Consequently, matrices associated to random samples of bounded orthonormal systems have nice sparse recovery properties.

Proposition 2 (Sparse recovery through BOS): Consider the matrix $A \in \mathbb{C}^{m\times N}$ whose rows are independent random samples of an orthonormal system $\{\psi_j, j \in [N]\}$ with bound $\sup_{j\in[N]} \|\psi_j\|_\infty \leq K$, drawn from the orthogonalization measure ν associated to the system. If the number of random samples satisfies

$$m \gtrsim K^2 s \log^3(s) \log(N), \tag{III.1}$$

for some $s \gtrsim \log(N)$, then the following holds with probability exceeding $1 - N^{-C\log^3(s)}$.

For each $x \in \mathbb{C}^N$, given noisy measurements $y = Ax + \sqrt{m}\eta$ with $\|\eta\|_2 \leq \varepsilon$, the approximation

$$x^{\#} = \arg\min_{z\in\mathbb{C}^N} \|z\|_1 \text{ subject to } \|Az - y\|_2 \leq \sqrt{m}\epsilon$$

satisfies the error guarantee

$$\|x - x^{\#}\|_2 \lesssim \frac{1}{\sqrt{s}}\|x - x_s\|_1 + \varepsilon.$$

An important special case of such a matrix construction is the *subsampled discrete Fourier matrix*, constructed by sampling $m \ll N$ rows uniformly at random from the unitary discrete Fourier matrix $\Psi \in \mathbb{C}^{N\times N}$ with entries $\psi_{j,k} = \frac{1}{\sqrt{N}} e^{i2\pi(j-1)(k-1)}$. Indeed, the system of complex exponentials $\psi_j(u) = e^{i2\pi(j-1)u}$, $j \in [N]$, is orthonormal with respect to the uniform measure over the discrete set $\mathcal{D} = \{0, \frac{1}{N}, \ldots, \frac{N-1}{N}\}$, and is bounded with optimally small constant $K = 1$. In the discrete setting, we may speak of a more general procedure for forming matrix constructions adhering to the conditions of Proposition 2: given any two unitary matrices Φ and Ψ, the composite matrix $\Phi^*\Psi$ is also a unitary matrix, and this composite matrix will have uniformly bounded entries if the orthonormal bases (ϕ_j) and (ψ_k), corresponding to the rows of Φ and Ψ respectively, are *mutually incoherent:*

$$\mu(\Phi, \Psi) \quad := \sqrt{N} \sup_{1\leq j,k\leq N} |\langle \phi_j, \psi_k \rangle| \leq K \tag{III.2}$$

Indeed, if Φ and Ψ are mutually incoherent, then the rows of $B = \sqrt{N}\Psi^*\Phi$ constitute a bounded orthonormal system with respect to the uniform measure on $\mathcal{D} = \{0, \frac{1}{N}, \ldots, \frac{N-1}{N}\}$. Proposition 2 then implies a sampling strategy for reconstructing signals $x \in \mathbb{C}^N$ with assumed sparse representation in the basis Ψ, that is $x = \Psi b$ and $b \approx b_s$, from a few linear measurements: form a sensing matrix $A \in \mathbb{C}^{m\times N}$ by sampling rows i.i.d. uniformly from an incoherent basis Φ, collect measurements $y = Ax + \eta$, $\|\eta\|_2 \leq \epsilon$, and solve the ℓ_1 minimization program,

$$x^{\#} = \arg\min_{z\in\mathbb{C}^N} \|\Psi^* z\|_1 \text{ subject to } \|Az - y\|_2 \leq \sqrt{m}\epsilon$$

This scenario is referred to as *incoherent sampling*.

IV. Local coherence sampling

Consider more generally the setting where we aim to compressively sense signals $x \in \mathbb{C}^N$ with assumed sparse representation in the orthonormal basis $\Psi \in \mathbb{C}^{N\times N}$, but our sensing matrix $A \in \mathbb{C}^{m\times N}$ can only consist of rows from some fixed orthonormal basis $\Phi \in \mathbb{C}^{N\times N}$ that is not necessarily incoherent with Ψ. In this setting, we ask: *Given a fixed sensing basis Ψ and sparsity basis Φ, how should we sample rows of Ψ in order to make the resulting system as incoherent as possible?* We will answer this question by introducing the concept of *local coherence* between two bases as described in [5], [6], whereby in the discrete setting the coherences of individual elements of the sensing basis are calculated and used to derive the sampling strategy.

The following result says that regions of the sensing basis that are *more* coherent with the sparsity basis should be sampled with higher density. The following is essentially a generalization of Theorem 2.1 in [5], but for completeness, we include a short self-contained proof.

Theorem 3 (Sparse recovery via local coherence sampling): Consider a measurable set $\mathcal{D}$ and a system $\{\psi_j, j \in [N]\}$ that is orthonormal with respect to a measure ν on $\mathcal{D}$ which has square-integrable local coherence,

$$\sup_{j\in[N]} |\psi_j(u)| \leq \kappa(u), \qquad \int_{u\in\mathcal{D}} |\kappa(u)|^2 \nu(u) du = B. \tag{IV.1}$$

We can define the probability measure $\mu(u) = \frac{1}{B}\kappa^2(u)\nu(u)$ on $\mathcal{D}$. Draw m sampling points $u_1, u_2, \ldots, u_m$ independently from the measure μ, and consider the the matrix $A \in \mathbb{C}^{m\times N}$ whose rows are the random samples $\psi_j(u_k), j \in [N]$. Consider also the diagonal preconditioning matrix $\mathcal{P} \in \mathbb{C}^{m\times m}$ with entries $p_{k,k} = 1/\mu(u_k)$. If the number of sampling points

$$m \gtrsim B^2 s \log^3(s) \log(N), \tag{IV.2}$$

for some $s \gtrsim \log(N)$, then the following holds with probability exceeding $1 - N^{-C\log^3(s)}$.

For each $x \in \mathbb{C}^N$, given noisy measurements $y = Ax + \sqrt{m}\eta$ with $\|\mathcal{P}\eta\|_2 \leq \sqrt{m}\varepsilon$, the approximation

$$x^{\#} = \arg\min_{z\in\mathbb{C}^N} \|z\|_1 \text{ subject to } \|\mathcal{P}Az - \mathcal{P}y\|_2 \leq \sqrt{m}\epsilon$$

satisfies the error guarantee

$$\|x - x^{\#}\|_2 \lesssim \frac{1}{\sqrt{s}}\|x - x_s\|_1 + \varepsilon$$

Proof: Consider the functions $Q_j(u) = \frac{\sqrt{B}}{\kappa(u)}\psi_j(u)$. The system $\{Q_j\}$ is bounded with $\sup_{j\in[N]} \|Q_j\|_\infty \leq \sqrt{B}$, and this system is orthonormal on $\mathcal{D}$ with respect to the sampling measure μ:

$$\begin{aligned}
&\int_{u\in\mathcal{D}} \bar{Q}_j(u) Q_k(u) \mu(u) du \\
&= \int_{u\in\mathcal{D}} \left(\frac{1}{\kappa(u)}\bar{\psi}_j(u)\right)\left(\frac{1}{\kappa(u)}\psi_k(u)\right)(\kappa^2(u)\nu(u)) du \\
&= \int_{u\in\mathcal{D}} \bar{\psi}_j(u)\psi_k(u)\nu(u) du = \delta_{jk}
\end{aligned} \tag{IV.3}$$

Thus we may apply Proposition 2 to the system $\{Q_j\}$, noting that the matrix of random samples of the system $\{Q_j\}$ may be written as $\mathcal{P}A$. ■

In the discrete setting where $\{\psi_j\}_{j\in[N]}$ and $\{\phi_k\}$ are rows of unitary matrices Ψ and Φ, and ν is the uniform measure over the set $\mathcal{D} = \{0, \frac{1}{N}, \ldots, \frac{N-1}{N}\}$, the integral in condition IV.1 reduces to a sum,

$$\sup_{k\in[N]} \sqrt{N} |\langle \psi_j, \phi_k \rangle| \leq \kappa_j, \quad \frac{1}{N}\sum_{j=1}^{N} \kappa_j^2 = B. \tag{IV.4}$$

This motivates the introduction of the local coherence of an orthonormal basis $\{\phi_j\}_{j=1}^N$ of $\mathbb{C}^N$ with respect to the orthonormal basis $\{\psi_k\}_{k=1}^N$ of $\mathbb{C}^N$ is the function $\mu^{loc} = (\mu_j) \in \mathbb{R}^N$ defined coordinate-wise by

$$\mu_j = \sup_{1\leq k\leq N} \sqrt{N}|\langle \varphi_j, \psi_k\rangle|.$$

We have the following corollary of Theorem 3.

Corollary 4: Consider a pair of orthonormal basis (Φ, Ψ) with local coherences bounded by $\mu_j \leq \kappa_j$. Let $s \geq 1$, and suppose that

$$m \gtrsim s(\frac{1}{N}\sum_{j=1}^N \kappa_j^2)\log^4(N).$$

Select m (possibly not distinct) rows of Φ^* independent and identically distributed from the multinomial distribution on $\{1, 2, \ldots, N\}$ with weights $c\kappa_j^2$ to from the sensing matrix $A : \mathbb{C}^N \to \mathbb{C}^m$. Consider also the diagonal preconditioning matrix $\mathcal{P} \in \mathbb{C}^{m\times m}$ with entries $p_{k,k} = \frac{1}{\sqrt{c}\kappa_j}$.

Then the following holds with probability exceeding $1 - N^{-C\log^3(s)}$.

For each $x \in \mathbb{C}^N$, given measurements $y = Ax + \eta$, with $\|\mathcal{P}\eta\|_2 \leq \sqrt{m}\varepsilon$, the approximation

$$x^{\#} = \arg\min_{u\in\mathbb{C}^N} \|\Psi^* u\|_1 \text{ subject to } \|y - \mathcal{P}Au\|_2 \leq \sqrt{m}\varepsilon$$

satisfies the error guarantee

$$\|x - x^{\#}\|_2 \lesssim \frac{1}{\sqrt{s}}\|\Psi^* x - (\Psi^* x)_s\|_1 + \varepsilon.$$

Remark 5: Note that the local coherence not only influences the embedding dimension m, it also influences the sampling measure. Hence a priori, one cannot guarantee the optimal embedding dimension if one only has suboptimal bounds for the local coherence. That is why the sampling measure in Theorem 3 is defined via the (known) upper bounds κ and $\|\kappa\|_2$ rather than the (usually unknown) exact values μ_{loc} and $\|\mu_{loc}\|_2$, showing that local coherence sampling is *robust with respect to the sampling measure:* suboptimal bounds still lead to meaningful bounds on the embedding dimension.

We now present two applications where incoherent sampling fails, but local coherence sampling provides a sampling scheme with sparse recovery guarantees.

V. Applications

A. Variable-density sampling for compressed sensing MRI

In Magnetic Resonance Imaging, after proper discretization, the unknown image (x_{j_1,j_2}) is a two-dimensional array in $\mathbb{R}^{n\times n}$, and allowable sensing measurements are two-dimensional Fourier transform measurements:

$$\phi_{k_1,k_2} = \frac{1}{n}\sum_{j_1,j_2} x_{j_1,j_2} e^{2\pi i(k_1 j_1 + k_2 j_2)/n}, \quad -n/2+1 \leq k_1, k_2 \leq n/2$$

Natural sparsity domains for images, such as discrete spatial differences, are not incoherent to the Fourier basis.

A number of empirical studies, including the very first papers on compressed sensing MRI , observed that image reconstructions from compressive frequency measurements could be significantly improved by variable-density sampling.

Note that lower frequencies are more coherent with wavelets and step functions than higher frequencies. In [6], the local coherence between the two-dimensional Fourier basis and bivariate Haar wavelet basis was calculated:

Proposition 6: The local coherence between frequency ϕ_{k_1,k_2} and the bivariate Haar wavelet basis $\Psi = (\psi_I)$ can be bounded by

$$\mu(\phi_{k_1,k_2}, \Psi) \lesssim \frac{\sqrt{N}}{(|k_1+1|^2 + |k_2+1|^2)^{1/2}}$$

Note that this local coherence is *almost square integrable independent of discretization size* n^2, as

$$\frac{1}{N}\sum_{j=1}^N \mu_j^2 \lesssim \log(n).$$

Applying Corollary 4 to compressive MRI imaging, we then have

Corollary 7: Let $n \in \mathbb{N}$. Let Ψ be the bivariate Haar wavelet basis and let $\Phi = (\phi_{k_1,k_2})$ be the two-dimensional discrete Fourier transform. Let $s \geq 1$, and suppose that $m \gtrsim s(\frac{1}{N}\log^5(N)$. Select m (possibly not distinct) frequencies (ϕ_{k_1,k_2}) independent and identically distributed from the multinomial distribution on $\{1, 2, \ldots, N\}$ with weights proportional to the inverse squared Euclidean distance to the origin, $\frac{1}{(|k_1+1|^2+|k_2+1|^2)}$, and form the sensing matrix $A : \mathbb{C}^N \to \mathbb{C}^m$.

Then the following holds with probability exceeding $1 - N^{-C\log^3(s)}$.

For each image $x \in \mathbb{C}^{n\times n}$, given measurements $y = Ax$, the approximation

$$x^{\#} = \arg\min_{u\in\mathbb{C}^{n\times n}} \|\Psi^* u\|_1 \text{ subject to } \|\mathcal{D}y - Au\|_2 \leq \epsilon$$

satisfies the error guarantee

$$\|x - x^{\#}\|_2 \lesssim \frac{1}{\sqrt{s}}\|\Psi^* x - (\Psi^* x)_s\|_1 + \varepsilon.$$

Numerical results such as those detailed in [**?**] and illustrated below in Figure 1 confirm that variable-density sampling strategies significantly outperform uniform sampling strategies as well as deterministic sampling strategies, and Corollary 7 provides theoretical justification for such observations. Below we provide a numerical comparison of various sampling strategies, including the sampling distribution given in Corollary 7. The following images were made from total variation minimization rather than Haar wavelet minimization, but the theory for Fourier-Wavelet sampling is extended to the total variation minimization setting in [6].

B. Sparse Legendre expansions for smooth function interpolation

Here we consider the problem of recovering polynomials g from m sample values $g(x_1), g(x_2)\ldots, g(x_m)$, with sampling points $x_\ell \in [-1, 1]$ for $\ell = 1, \ldots, m$. If the number of

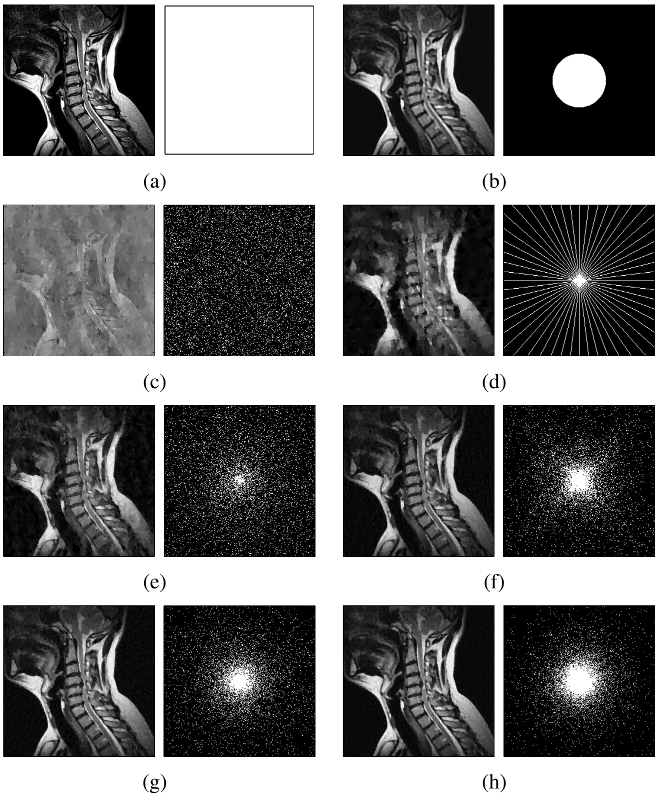


Fig. 1. Various reconstructions of an MRI image $x \in \mathbb{R}^{256\times256}$ with total variation minimization from $m = 6400$ noiseless partial DFT measurements sampled from various distributions. Beside each reconstruction is a plot of frequency space $\{(k_1, k_2) : -N/2+1 \le k_1, k_2 \le N/2\}$ and the frequencies used in its reconstruction. (a) Original image. (b) Reconstruction using only lowest frequencies: $\Omega = \{(k_1, k_2) : k_1^2 + k_2^2 \le 80\}$. (c) $\text{Prob}\big((k_1, k_2) \in \Omega\big) \sim 1$ (Uniform subsampling) (d) Ω comprised of frequencies in equispaced radial lines. (e) $\text{Prob}\big((k_1, k_2) \in \Omega\big) \propto (k_1^2+k_2^2+1)^{-1/2}$(f) $\text{Prob}\big((k_1, k_2) \in \Omega\big) \propto \big(\max(|k_1|, |k_2|)+1\big)^{-1}$ (g) $\text{Prob}\big((k_1, k_2) \in \Omega\big) \propto (k_1^2+k_2^2+1)^{-1}$. (h) $\text{Prob}\big((k_1, k_2) \in \Omega\big) \propto (k_1^2 + k_2^2 + 1)^{-3/2}$. The relative reconstruction error $\|f - f^{\#}_{TV}\|_2/\|f\|_2$ corresponding to each reconstruction is (b) .2932, (c) .8229, (d) .4074, (e) .3192, (f) .2603, (g) .2537, and (h) .2463.

sampling points is less or equal to the degree of g, then in general such reconstruction is impossible due to dimension reasons. However, the situation becomes tractable if we make a sparsity assumption. In order to introduce a suitable notion of sparsity, we consider the orthonormal basis of Legendre polynomials.

Definition 8: The (orthonormal) Legendre polynomials

$$P_0, P_1, \dots, P_n, \dots$$

are uniquely determined by the following conditions:

- $P_n(x)$ is a polynomial of precise degree n in which the coefficient of x^n is positive,
- the system $\{P_n\}_{n=0}^{\infty}$ is orthonormal with respect to the normalized Lesbegue measure on $[-1, 1]$:

$$\frac{1}{2}\int_{-1}^{1} P_n(x)P_m(x)dx = \delta_{n,m}, \quad n, m = 0, 1, 2, \dots$$

Since the interval $[-1, 1]$ is symmetric, the Legendre polynomials satisfy $P_n(x) = (-1)^n P_n(-x)$. For more information see [Szego].

An arbitrary real-valued polynomial g of degree $N-1$ can be expanded in terms of Legendre polynomials,

$$g(x) = \sum_{j=0}^{N-1} c_j P_j(x), \quad x \in [-1, 1]$$

with coefficient vector $c \in \mathbb{R}^N$. The vector is s-sparse if $\|c\|_0 \le s$. Given a set of m sampling points $(x_1, x_2, \dots, x_m)$, the samples $y_k = g(x_k)$, $k = 1, \dots, m$, may be expressed concisely in terms of the coefficient vector according to

$$y = \Phi c,$$

where $\phi_{k,j} = P_j(x_k)$. If the sampling points $x_1, \dots, x_m$ are random variables, then the matrix $\Phi \in \mathbb{R}^{m\times N}$ is exactly the sampling matrix corresponding to random samples from the Legendre system $\{P_j\}_{j=1}^N$. This is not a bounded orthonormal system, however, as the Legendre polynomials grow like

$$|P_n(x)| \le (n + 1/2)^{1/2}, \quad -1 \le x \le 1.$$

Nevertheless the Legendre system does have bounded local coherence. A classic result [szego] follows.

Proposition 9: For all $n > 0$ and for all $x \in [-1, 1]$,

$$|P_n(x)| < \kappa(x) = 2\pi^{-1/2}(1 - x^2)^{-1/4}.$$

here, the constant is $2\ \pi^{-1/2}$ cannot be replaced by a smaller one.
Indeed, $\kappa(x)$ is a square integrable function proportional to the Chebyshev measure $\pi^{-1}(1 - x^2)^{-1/2}$. We arrive at the following result for Legendre polynomial interpolation as a corollary of Theorem 3.

Corollary 10: Let $x_1, \dots, x_m$ be chosen independently at random on $[-1, 1]$ according to the Chebyshev measure $\pi^{-1}(1 - x^2)^{-1/2}dx$. Let Ψ be the matrix with entries $\Psi_{k,j} = \sqrt{\pi/2}(1 - x_k^2)^{1/4}P_n(x_k)$. Suppose that

$$m \gtrsim s \log^3$$

Consider the matrix $A \in \mathbb{C}^{m\times N}$ whose rows are independent random vectors$(\psi_j(X_k))$ drawn from the measure μ. If

$$m \gtrsim B^2 s \log^3(s) \log(N), \tag{V.1}$$

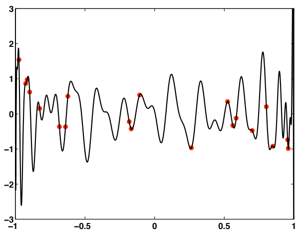
for some $s \gtrsim \log(N)$, then the following holds with probability exceeding $1 - N^{-C\log^3(s)}$. Let $\mathcal{D} \in \mathbb{C}^{m\times m}$ be the diagonal matrix with entries $d_{k,k} = \frac{1}{\mu(X_k)}$. For each $x \in \mathbb{C}^N$, given noisy measurements $y = Ax + \sqrt{m}\eta$ with $\|\mathcal{D}\eta\|_2 \le \sqrt{m}\varepsilon$, the approximation

$$x^{\#} = \arg\min_{u\in\mathbb{C}^N} \|u\|_1 \text{ subject to } \|\mathcal{D}Au - \mathcal{D}y\|_2 \le \sqrt{m}\epsilon$$

satisfies the error guarantee

$$\|x - x^{\#}\|_2 \lesssim \frac{1}{\sqrt{s}}\|x - x_s\|_1 + \varepsilon$$

where x_s is the best s-term approximation to x.



We illustrate exact recovery of a Legendre sparse polynomial from randomly sampled points from the Chebyshev measure.

In fact, more general theorems exist: the Chebyshev measure is a universal sampling strategy for interpolation with any set of orthogonal polynomials [5].

An extension to the setting of interpolation with spherical harmonics can be found in [5], [**?**].

VI. Conclusion

Here we summarize local coherence sampling, and demonstrate its power for generalized sparse recovery results in compressed sensing in two seemingly disparate settings - MRI compressive imaging and Legendre polynomial interpolation. Unlike incoherence-based results, local coherence sampling gives a sampling strategy for fixed sparsity basis and fixed sensing basis from which one can subsample; if the local coherence function is square integrable and this integral depends only mildly on the ambient dimension of the signal, then stable and robust sparse recovery results for incoherent sampling generalize to this setting. Several questions remain, such as the optimality of the local coherence sampling, extensions to frames rather than orthonormal dictionaries, and connections to designing sensing matrices via minimizing the local coherence [**?**].

Acknowledgment

The authors would like to thank Wotao Yin, Mark Tygert for helpful discussions related to this paper.

References

[1] E. J. Candes, J. Romberg, and T. Tao. *Stable signal recovery from incomplete and inaccurate measurements.* Comm. Pure Appl. Math., 59(8):1207–1223, 2006.

[2] M. Lustig, D. Donoho, and J.M. Pauly, *Sparse MRI: the application of compressed sensing for rapid MRI imaging*, Magnetic Resonance in Medicine, 58(6):1182-1195, 2007.

[3] M. Lustig, D.L. Donoho, J.M. Santos, and J.M. Pauly. *Compressed sensing MRI.* IEEE Sig. Proc. Mag., 25(2):72-82, 2008.

[4] L. Brutman. *Lebesgue functions for polynomial interpolatio - a survey.* Ann. Numer. Math., 4 (1-4): 111– 127, 1997.

[5] H. Rauhut and R. Ward, *Sparse Legendre expansions via ℓ_1-minimization*, Journal of Approximation Theory, 164:517–533, 2012.

[6] F. Krahmer and R. Ward, *Beyond incoherence: Stable and robust sampling strategies for compressive imaging*, Preprint, 2012.

[7] H. Rauhut, *Compressive sensing and structured random matrices*, In M. Fornasier, editor, *Theoretical Foundations and Numerical Methods for Sparse Recovery*, Volume 9 of *Radon Series Comp. Appl. Math.*, pages 1–92. deGruyter, 2010. 0.5em minus 0.4emHarlow, England: Addison-Wesley, 1999.