ON APPROXIMATION OF BANDLIMITED FUNCTIONS WITH COMPRESSED SENSING

Adrian E.G. Huber, Shih-Chii Liu

Institute of Neuroinformatics, University of Zurich and ETH Zurich, Zurich, Switzerland
huberad@ethz.ch, shih@ini.ethz.ch

ABSTRACT

The application of Compressed Sensing techniques to bandlimited functions is investigated in this paper. It is shown that under the assumption of sparsity, stable reconstruction of a bandlimited function is possible from finitely many samples, contrary to classical results from signal processing theory. The number of measurements that need to be taken is proportional to the sparsity of the function. In compact intervals, it is shown that the number of pointwise measurements required scales quadratically with the size of the largest expansion coefficient (in a basis in which sparsity is measured) which is sufficient for a faithful function approximation.

Index Terms— Bandlimited Functions, Irregular Sampling, Compressed Sensing

1. INTRODUCTION

The approximate reconstruction of a bandlimited function in an interval from a finite number of samples is a well-studied problem. The classical Whittaker-Shannon-Kotelnikov sampling theorem requires infinitely many samples for the reconstruction of the entire bandlimited function [1]; reconstruction within a compact interval from samples taken at the Nyquist rate can lead to major errors as the sinc-function decays only with a rate of $\frac{1}{t}$ for $t \to \infty$, i.e. samples from far outside the interval to be reconstructed can influence the values of the function in the interval to a major extent. Such truncation errors have been thoroughly studied in the literature. By introducing oversampling, local reconstruction can be achieved with far greater accuracy, as the reconstruction function can be chosen to decay much faster in time than the sinc-function associated with Nyquist-rate sampling. Errors arising from truncated sinc-expansions in the oversampling regime have been studied by Helms and Thomas [2]; the same authors also developed bounds for the truncation errors in case of self-truncating sampling expansions. Truncation error bounds for finite-energy signals were derived by Brown [3]. A related approach was followed by Knab who analyzed error bounds arising from estimating a bandlimited function in an interval from a finite number of equidistant samples using Lagrange interpolation [4]. Error bounds for Lagrange interpolation from equidistant samples of a bandlimited function depending on the sampling rate and the Nyquist rate were developed by Radzyner and Bason [5]. Klamer and Masry studied error bounds arising from Lagrange interpolation of bandlimited functions with finitely many sampling points distributed according to a point process [6]. They derived error bounds for sampling points distributed according to a Poisson point process in particular. Strohmer and Tanner considered nonuniform periodic sampling, deriving a reconstruction algorithm

using a finite number of samples [7]. Returning to Lagrange interpolation, Selva considered a weighted Lagrange interpolation scheme for the local approximation of a bandlimited function from nonuniform samples [8]. Explicit error bounds were given for nonuniform sampling schemes with a maximum deviation of individual samples from a uniform grid.

Compressed Sensing (CS) studies the solution of highly underdetermined linear systems, exploiting random measurements and the sparsity of the signal to be reconstructed [9]. Classical CS theory applies to finite-dimensional spaces. Recently, Adcock and Hansen extended CS theory to infinite-dimensional spaces, thereby enabling the application of CS to functions living in Hilbert spaces [10].

By combining infinite-dimensional CS with the theory of Prolate Spheroidal Wave Functions (PSWF), we derive approximation methods for bandlimited functions. A similar approach which can be found in [11] uses CS to recover functions sampled pointwise, assuming that the functions are sparse in a PSWF basis. The main difference of our work from [11] is that we lower the lower bounds on the number of measurements sufficient for faithful approximation; additionally, our method does not require sampling points distributed according to a Chebyshev distribution for reconstructing expansion coefficients above a certain index. Instead, uniform sampling can be used throughout in our formulation. In Section 2 we discuss the main aspects of infinite-dimensional CS and recapitulate the basics of PSWF and Reproducing Kernel Hilbert Spaces (RKHS). Thereafter we derive our main results.

1.1. Notation

The version of the Fourier transform used in this paper is $\hat{f}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t)e^{-i\omega t}dt$. All considered functions live in the Paley-Wiener space $PW_{\pi w} := \{f : f \in L^2(\mathbb{R}) \bigcap C(\mathbb{R}), \text{supp}\hat{f} \subseteq [-\pi w, \pi w], w > 0\}$, where $L^2(\mathbb{R})$ is the space of all square-integrable (Lebesgue) functions over $\mathbb{R}$, $C(\mathbb{R})$ is the space of all continuous functions on the real line, supp is the support of a function, and $\hat{f}$ is the Fourier transform of the function f. The reproducing kernel of $PW_{\pi w}$ is $k_{\pi w}(t,s) = w \cdot \text{sinc}(w(t-s))$, where $\text{sinc}(t) = \frac{\sin(\pi t)}{\pi t}$. The coherence $v(U)$ of an infinite matrix U is defined as $v(U) = \sup_{i,j \in \mathbb{N}} |u_{ij}|$ with u_{ij} the entries of matrix U. $\langle \cdot, \cdot \rangle$ denotes the inner product in a generic Hilbert space H over $\mathbb{C}$. The effective interval in which most of the function energy is concentrated has length $T > 0$.

2. INFINITE-DIMENSIONAL COMPRESSED SENSING

Let H be a separable Hilbert space with an orthonormal basis $\{\phi_j\}_{j \in \mathbb{N}}$. Then every function in H can be expanded as $f = \sum_{j=1}^{\infty} \alpha_j \phi_j$ with $\alpha_j = \langle f, \phi_j \rangle$. Let $\Delta = \text{supp}(f) \subset \{1, \ldots, M\}$ with $M \in \mathbb{N}$ and $\text{supp}(f) = \{j \in \mathbb{N} : \alpha_j \neq 0\}$. If $|\Delta| = r$, f is

(r, M)-sparse in the basis $\{\phi_j\}_{j \in \mathbb{N}}$. The best approximation error for compressible signals (see [9]) can then be defined as

$$\sigma_{r,M}(\alpha) = \min\{\|\alpha - \eta\|_1 : \eta \text{ is } (r, M)\text{-sparse}\}. \tag{1}$$

Let $\zeta_1(f)$, $\zeta_2(f)$, ..., be a countable collection of samples with $\zeta_j(f) = \langle f, \psi_j \rangle$ and $\{\psi_j\}_{j \in \mathbb{N}}$ an orthonormal basis for H. The infinite matrix

$$U = \begin{pmatrix} \langle \phi_1, \psi_1 \rangle & \langle \phi_2, \psi_1 \rangle & \langle \phi_3, \psi_1 \rangle & \dots \\ \langle \phi_1, \psi_2 \rangle & \langle \phi_2, \psi_2 \rangle & \langle \phi_3, \psi_2 \rangle & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \tag{2}$$

is an isometry. Suppose that m of the first N measurements $\zeta_j(f)$ are chosen uniformly at random with the position of the chosen measurements being indicated by $\Omega \subset \{1, \ldots, N\}$ and $|\Omega| = m$. Define P_Ω to be the orthogonal projection from $l^2(\mathbb{N})$ to $\text{span}\{e_j : j \in \Omega\}$ with $\{e_j : j \in \mathbb{N}\}$ the canonical basis of $l^2(\mathbb{N})$ and P_M the orthogonal projection to $\text{span}\{e_j : j = 1, \ldots, M\}$. One result from [10] reads then as follows: Provided certain technical requirements on N and m are met, then by solving the finite-dimensional problem

$$\inf_{\eta \in l^1(\mathbb{N})} \|\eta\|_1 \text{ subject to } P_\Omega U P_M \eta = P_\Omega \zeta, \tag{3}$$

a solution ξ can be found with probability $1 - \epsilon$, $\epsilon > 0$, which is close in norm to the true solution α:

$$\|\xi - \alpha\| \leq 8 \left(1 + \frac{2N}{m}\right) \sigma_{|\Delta|,M}(\alpha). \tag{4}$$

One requirement on m is

$$m \geq C \cdot N \cdot v^2(U) \cdot |\Delta| \cdot \left(\log\left(\epsilon^{-1}\right) + 1\right) \cdot \log\left(\frac{MN\sqrt{|\Delta|}}{m}\right), \tag{5}$$

i.e. the number of necessary measurements m in order to obtain a faithful reconstruction with sufficiently high probability is bounded from below. C is a fixed constant in Eq. (5).

2.1. Prolate Spheroidal Wave Functions

We recapitulate the basics of Prolate Spheroidal Wave Functions (PSWF) briefly. The PSWF were introduced into signal analysis in a series of papers [12–14]. They are solutions to both a differential equation and to an integral equation, forming an orthonormal basis for the Paley-Wiener space on the real line. The PSWF satisfy the integral equation

$$\int_{-1}^{1} \frac{\sin(c(x-y))}{\pi(x-y)} \phi(y) dy = \lambda \phi(x), \ |x| \leq 1, \tag{6}$$

where $c = \frac{\pi w T}{2}$. The differential equation which the PSWF satisfy is

$$\frac{d}{dx}(1-x^2)\frac{d\phi}{dx} + (\chi - c^2 x^2)\phi = 0. \tag{7}$$

Equation (6) has solutions for discrete values of λ which can be sorted in a descending order:

$$\lambda_0 > \lambda_1 > \lambda_2 > \cdots > 0. \tag{8}$$

Possible eigenvalues λ are functions of c, i.e. $\lambda_i = \lambda_i(c), i \in \mathbb{N}_0$. The PSWF basis numbering starts with 0, contrary to the generic basis $\{\phi\}_{i \in \mathbb{N}}$. The PSWF constitute an optimal basis for the space

of bandlimited functions [15]. We consider a subclass of the Paley-Wiener space of functions with maximum energy E. Then the Kolmogorov n-width d_n in $L_2(-\frac{T}{2}, \frac{T}{2})$ of the energy-bounded subclass of the Paley-Wiener space equipped with the L_2-norm in $\left(-\frac{T}{2}, \frac{T}{2}\right)$ is equal to $d_n = \sqrt{E \lambda_n}$ and the subspace which leads to this infinum is $S_n = \text{span}(\phi_0, \phi_1, \ldots, \phi_{n-1})$, ϕ_j being the PSWF. The best approximation to a function f in an interval living in the energy-bounded Paley-Wiener space in any n-dimensional subspace is then $\sum_{j=0}^{n-1} \langle f, \phi_j \rangle \phi_j$. The worst case error arising from this approximation is equal to $\sqrt{E \lambda_n}$.

2.2. Reproducing Kernel Hilbert Space

$PW_{\pi w}$ is a Reproducing Kernel Hilbert Space (RKHS). The reproducing kernel is given by $k_{\pi w}(t, s) = w \cdot \text{sinc}(w(t-s))$. If s is chosen to correspond to $\{\frac{n}{w}\}_{n \in \mathbb{Z}}$, then the set $\frac{1}{\sqrt{w}}\{k_{\pi w}(t, \frac{n}{w})\}_{n \in \mathbb{Z}}$ is an orthonormal basis for $PW_{\pi w}$. Sampling at the Nyquist rate corresponds therefore implicitly (after normalization of sample values by $\frac{1}{\sqrt{w}}$) to inner products with an orthonormal basis. In the case of oversampling (samples are taken at the rate $\{\frac{n}{w'}\}_{n \in \mathbb{Z}}$, $w' > w$) the induced set of functions

$$\frac{1}{\sqrt{w'}}\{k_{\pi w'}(t, \frac{n}{w'})\}_{n \in \mathbb{Z}}, \tag{9}$$

forms a tight frame with unit frame bound in $PW_{\pi w}$. Consider now the infinite matrix U from Eq. (2) with $\{\psi_i\}_{i \in \mathbb{N}}$ a tight frame with unit frame bound and $\{\phi_i\}_{i \in \mathbb{N}}$ an orthonormal basis for $PW_{\pi w}$. U is an isometry. Indeed,

$$U\alpha = U \begin{pmatrix} \langle f, \phi_1 \rangle \\ \langle f, \phi_2 \rangle \\ \langle f, \phi_3 \rangle \\ \vdots \end{pmatrix} = \begin{pmatrix} \langle f, \psi_1 \rangle \\ \langle f, \psi_2 \rangle \\ \langle f, \psi_3 \rangle \\ \vdots \end{pmatrix} = \zeta(f), \tag{10}$$

with $f \in PW_{\pi w}$. By Parseval's identity it holds that $\|\alpha\|^2 = \|f\|^2$. $\|\zeta(f)\|^2$ is furthermore equal to $\|f\|^2$ as the set $\{\psi_i\}_{i \in \mathbb{N}}$ is a tight frame with unit frame bound. Hence $\|U\alpha\|^2 = \|\alpha\|^2$ and U is an isometry.

2.3. Infinite Dimensional Compressed Sensing with PSWF basis

2.3.1. Approximation on the real line

We estimate an upper bound for the coherence of matrix U defined in Eq. (2) with the PSWF basis $\{\phi_i\}$ and the tight frame $\{\psi_i\}$ from Eq. (9). All PSWF have unit energy on the real line. A standard result from the theory of RKHS (see [16]) yields that the maximum value attainable by a function from the RKHS at an arbitrary point $t_0 \in \mathbb{R}$, $\mathbb{R}$ being the domain, assuming that the function has energy E, is given by $\max_{\|f\|^2 \leq E} |f(t_0)|^2 = Ek(t_0, t_0)$, $k(t, s)$ being the reproducing kernel. It follows that the coherence of the infinite matrix U is bounded from above by $v(U) \leq \sqrt{\frac{w}{w'}}$. We conclude that the coherence of the infinite matrix U in Eq. (2) depends on the bandwidth πw and on the sampling rate $w' > w$. By oversampling suitably, the coherence of U can be made as small as required.

Let us now study the term $N \cdot v^2(U)$ from Eq. (5). As discussed above, an upper bound for the squared coherence $v^2(U)$ of matrix U in Eq. (2) is equal to w divided by w'. Furthermore, the following is valid:

$$f \in PW_{\pi w} \Rightarrow f \to 0 \quad \text{as} \quad t \to \pm \infty. \tag{11}$$

Therefore, for every function in $PW_{\pi w}$ there exists a $T > 0$ such that almost all of the energy of the function is located in $\left[-\frac{T}{2}, \frac{T}{2}\right]$.

We consider equispaced sampling points in $\left[-\frac{T}{2}, \frac{T}{2}\right]$, spaced apart by $\frac{1}{w'}$. N, the number of sampling points in $\left[-\frac{T}{2}, \frac{T}{2}\right]$, is therefore linear in w'. The matrix $P_N U P_M$ from which m rows are drawn with uniform distribution is therefore close to an isometry, implying that the theorems from [10] apply. The product of N and $v^2(U)$ reduces to a constant value which depends on w. Disregarding log-terms, the number of measurements needed scales then linearly with $|\Delta|$ as can be seen from Eq. (5):

$$m \geq C \cdot T \cdot w \cdot |\Delta| \cdot \left(\log\left(\epsilon^{-1}\right) + 1\right) \cdot \log\left(\frac{MN\sqrt{|\Delta|}}{m}\right). \quad (12)$$

In the case of sampling the function *globally*, the number of measurements would therefore be proportional to the support $|\Delta|$ and the time-bandwidth product $T \cdot w$. For a general bandlimited function, one cannot assume a priori knowledge on Δ. By assuming sparsity, however, a bandlimited function can be fully reconstructed from finitely many samples in a stable way. Without the assumption of sparsity, infinitely many samples are necessary for a sampling set to be stable [1]. A stable sampling scheme in this sense is then also a set of uniqueness, i.e. there is only one bandlimited function whose values at the sampling set correspond to the sampled values.

Remark 1. In principle, one could choose a different orthonormal basis for $PW_{\pi w}$ than the one set up by the PSWF. Since bandlimited functions are contained in $L^2(\mathbb{R})$ on the real line, any orthonormal basis for $L^2(\mathbb{R})$ would suffice. The reproducing kernel for $PW_{\pi w}$, however, acts as a lowpass filter, implying that any non-bandlimited basis element from such a hypothetical orthonormal basis would first have to be ideally lowpass-filtered before being evaluated at a specific point. Since the PSWF are bandlimited themselves, pointwise evaluation suffices.

Remark 2. The size of N (or equivalently of w') influences the irregularity of the sampling pattern. The larger N becomes, the more the sampling process resembles truly uniform sampling on the interval $\left[-\frac{T}{2}, \frac{T}{2}\right]$. In fact, a larger N implies a greater possible sampling pattern irregularity in $\left[-\frac{T}{2}, \frac{T}{2}\right]$. This can be seen as follows: Set $C' = C \cdot N \cdot v^2(U) \cdot \left(\log\left(\epsilon^{-1}\right) + 1\right)$ (C' is a constant for fixed ϵ), and $|\Delta| = M$. We obtain the following inequality from Eq. (5):

$$e^m m^{C'M} \geq C'' N^{C'M}, \quad (13)$$

with $C'' = \left(M\sqrt{M}\right)^{C'M}$. In the case of a growing N, the required m in order to fulfil Eq. (13) (and hence Eq. (5)) will grow slower than N. This can be seen by comparing the derivative of the left hand side of Eq. (13) with respect to m with the derivative of the right hand side with respect to N. For the left hand side one obtains $e^m m^{C'M} e^{-1} \left(C'M + m\right)$ and for the right hand side $C'' N^{C'M} N^{-1} C'M$. If N and m are chosen in such a way as to fulfil Eq. (13) it follows that

$$e^m m^{C'M} e^{-1} \left(C'M + m\right) \geq C'' N^{C'M} e^{-1} \left(C'M + m\right)$$
$$\geq C'' N^{C'M} N^{-1} C'M, \quad (14)$$

provided that $C'M + m \geq eN^{-1}C'M$. The latter inequality is certainly fulfilled as N is always larger than e.

2.3.2. Approximation in intervals

Assume now that we are only interested in the approximation of a bandlimited function in an interval. In general, finitely many samples in an interval do not determine a bandlimited function uniquely.

By restricting a bandlimited function to an interval, the resulting function space ceases to be a RKHS, that is, pointwise sampling is no longer continuous. After normalization, the PSWF form an orthonormal basis for the interval of interest. Within the interval, an upper bound for the values of the PSWF cannot be obtained from RKHS techniques as in Section 2.3.1. It is known, however, that for large enough integers $n \geq 0$ the largest absolute value of the normalized PSWF can be found at $-\frac{T}{2}$ and $\frac{T}{2}$ [17]. Furthermore, an upper bound for this largest value is proportional to $\sqrt{n}$ [17]. Hence the approach from Section 2.3.1 cannot be used, as no upper bound for the coherence can be given. Following Corollary 7.1 in [18], we use a weighted minimization scheme to solve the interpolation problem in the interval by introducing weights $\{w_i\}$ which grow as fast as the maximum value of the PSWF in our interval of interest, i.e. with rate $\sqrt{i}$. Following the line of argument given in [18], it transpires that the number of measurements needed in the interval of interest is proportional to M^2, with M the largest integer for which a coefficient that is nonzero is expected. As discussed above in Section 2.1, the worst case error of functions from the energy-bounded Paley-Wiener space expanded in a subspace spanned by the first n PSWF basis elements is equal to $\sqrt{E\lambda_n}$. Given that λ decays rapidly for $n > \frac{2c}{\pi}$, in general one will need more than $\frac{2c}{\pi}$ PSWF basis elements for an acceptable worst case approximation error. Equating n with M and setting $\Delta = \{1, \ldots, M\}$ (in general, all coefficients $\langle f, \phi_j \rangle$ for $j \in \{1, \ldots, M\}$ will be nonzero), one can conclude that one has to oversample locally, as $\frac{2c}{\pi}$ corresponds to the number of Nyquist-rate samples in an interval $\left[-\frac{T}{2}, \frac{T}{2}\right]$.

It is interesting to observe the qualitative difference between the sampling of bandlimited functions on the real line and on intervals assuming sparsity. In the former case, the number of sampling points sufficient scales linearly with $|\Delta|$, in the latter case quadratically. As mentioned above, one of the reasons for this behavior is the lack of continuity in the sampling process in the time-limited case since the space is no longer a RKHS.

Remark 3. A different approach to approximate a bandlimited function locally from uniformly distributed samples can be based on previous work showing how to determine the Restricted Isometry Property (RIP) for finitely many measurements in potentially infinite-dimensional Hilbert spaces [19]. It is known (see [20]) that for sampling times $\{t_n\}_{n \in \mathbb{Z}}$ with a maximum separation $\delta = \sup_{n \in \mathbb{Z}} (t_{n+1} - t_n) < \frac{1}{w}$ the following is true for any $f \in PW_{\pi w}$:

$$(1 - \delta w)^2 \|f\|^2 \leq \sum_{n \in \mathbb{Z}} \omega_n |f(t_n)|^2 \leq (1 + \delta w)^2 \|f\|^2 \quad (15)$$

with $\omega_n = \frac{t_{n+1} - t_{n-1}}{2}$. We assume that nearly all of the energy of f is located in the interval of interest. This assumption is in contrast to the first part of Section 2.3.2 in which it was not assumed that the interval of interest contains most of the signal energy. By suitable time-windowing that leaves the function bandlimited (albeit with a potentially different bandlimit), however, one can enforce this energy condition. By sampling only in the interval of interest, Eq. (15) will not be strictly fulfilled; we disregard this error from now on as it can be made arbitrarily small by increasing our interval. Using Theorem II.2 from [19] and Eq. (15), we now show that uniformly distributed sampling points satisfy the RIP with a large probability. Define the continuous linear map L that operates on the coefficients η of $f \in PW_{\pi w}$ in the PSWF basis and that returns m pointwise evaluations of f distributed uniformly in the interval $\left[-\frac{T}{2}, \frac{T}{2}\right]$, each sample being rescaled by $\sqrt{\omega_n}$. Let μ be the probability measure which leads to uniformly distributed rescaled sampling

points in $\left[-\frac{T}{2}, \frac{T}{2}\right]$. Then $\mathbb{E}_\mu \|L(\eta)\|^2 = \omega \sum_{i=1}^{m} \left| f\left(-\frac{T}{2} + i\Lambda\right)\right|^2$ with $\Lambda = \frac{T}{m+1}$ and $\omega = \Lambda = \frac{1}{w'}$. Hence, if m is chosen such that the induced $w' \geq w$, it follows that $\mathbb{E}_\mu \|L(\eta)\|^2 = \|f\|^2$. As in [19], define $\delta_{S,\mu,2} = \sup_{x \in S} \left| \|L(\eta)\|^2 - \mathbb{E}_\mu \|L(\eta)\|^2 \right| = \left(2\delta w + \delta^2 w^2\right)\|f\|^2$ using Eq. (15), with S being the set of $2|\Delta|$-sparse coefficient vectors η with unit energy $\|f\|^2 = 1$. Since the RIP constant $\delta_{\text{RIP}} \geq \delta_{S,\mu,2}$, it is necessary to oversample for a RIP constant below one to be feasible: $(2\delta w + \delta^2 w^2) < 1 \Leftrightarrow \delta < \frac{\sqrt{2}-1}{w}$. If, as above, $\Delta \subset \{1, \ldots, M\}$, then S has a finite upper box-counting dimension (cf. [19] for definitions). Additionally, for Theorem II.2 from [19] to work, the following probability must be bounded from above:

$$\mathbb{P}\left\{\left| \|L(\eta_1)\|^2 - \mathbb{E}_\mu \|L(\eta_1)\|^2 - \|L(\eta_2)\|^2 + \mathbb{E}_\mu \|L(\eta_2)\|^2 \right| \geq \right.$$
$$\left. \lambda\|\eta_1 - \eta_2\| \right\}, \tag{16}$$

with η_1 and $\eta_2 \in S$ or zero in all its entries and $\lambda \geq 0$. For η_1 and η_2 either all zero or identical, a bound is trivial, i.e. statement (16) has probability one. For η_1 and η_2 distinct, we can use Hoeffding's inequality to obtain an upper bound. Using Eq. (15), we derive:

$$\mathbb{P}\left\{\left| \|L(\eta_1)\|^2 - \mathbb{E}_\mu \|L(\eta_1)\|^2 - \|L(\eta_2)\|^2 + \mathbb{E}_\mu \|L(\eta_2)\|^2 \right| \geq \right.$$
$$\left. \lambda\|\eta_1 - \eta_2\| \right\} \leq 2\exp\left(-\frac{\lambda^2}{8\delta^2 w^2}\right). \tag{17}$$

Equation (17) is valid for all λ. It is worth emphasizing that Eq. (17) is only a correct statement insofar as Eq. (15) is true which necessitates oversampling. In case m points are chosen with uniform probability independently in an interval $\left[-\frac{T}{2}, \frac{T}{2}\right]$ and sorted in ascending order, the distribution for the greatest distance between any two consecutive points is given by $F_\delta(z) = \left(1 - \frac{(T-z)^m}{T^m}\right)^{m-1}$ with z ranging from zero to T. $F_{\delta^2}(z)$ is then given by $F_{\delta^2}(z) = \left(1 - \frac{\left(T - z^{\frac{1}{2}}\right)^m}{T^m}\right)^{m-1}$. Hence it follows that $\delta^2 < \Lambda$ (assuming $\delta < 1$) with overwhelming probability for even slight oversampling (factor two); Eq. (17) can then be manipulated to yield:

$$\mathbb{P}\left\{\left| \|L(\eta_1)\|^2 - \mathbb{E}_\mu \|L(\eta_1)\|^2 - \|L(\eta_2)\|^2 + \mathbb{E}_\mu \|L(\eta_2)\|^2 \right| \geq \right.$$
$$\left. \lambda\|\eta_1 - \eta_2\| \right\} \leq 2\exp\left(-\frac{\lambda^2 m}{8Tw^2}\right). \tag{18}$$

Given that the continuous linear map L has a finite upper box-counting dimension and that inequality (18) holds, we conclude by invoking Theorem II.2 from [19]: The RIP holds with probability $1 - \xi$, i.e. $\delta_{S,\mu,2} \leq \delta_{\text{RIP}}$ for any $\xi, \delta_{\text{RIP}} \in (0, 1)$ if

$$m \geq \frac{8CTw^2}{\delta_{\text{RIP}}^2} max\left\{ (2|\Delta| + 1)\log\left(\frac{1}{\epsilon_S}\right), \log\left(\frac{6}{\xi}\right) \right\}, \tag{19}$$

with $C > 0$ a constant independent of all other parameters and ϵ_S given in [19]. The bound shown in Eq. (19) is structurally similar to the one derived in Eq. (12).

3. IMPLEMENTATION AND RESULTS

The PSWF are generated by using a freely available numerical software package [21]. Since the generation of PSWF with a large parameter c is numerically difficult, we have restricted our practical investigation to $c = 30$. To test the algorithm, a speech signal is

bandpass-filtered and the resulting bandpass-filtered signal is represented in its equivalent baseband form. This equivalent baseband form is then sampled uniformly in an interval. An illustrative example is shown in Fig. 1. In this example, T is set to 2, while $c = 30$. 90 sampling points are uniformly distributed within the interval $\left[-\frac{T}{2}, \frac{T}{2}\right]$. The approximation via weighted l_1-minimization in the interval is essentially perfect. It is worth pointing out that the actual number of measurements needed is smaller than the number required by theory. The upper bound discussed in Section 2.3.2 for the normalized PSWF element i (proportional to $\sqrt{i}$) seems to be a conservative estimate in practice. The homotopy method is used in our implementation for the minimization of the CS problem. In the case of noisy measurements, a weighted l_1-minimization is able to recover a solution with an error proportional to the noise term. Fig. 2 depicts the same signal as Fig. 1; contrary to the case in Fig. 1, however, uniform i.i.d. noise (SNR ≈ 4) is added to the perfect sampling values.
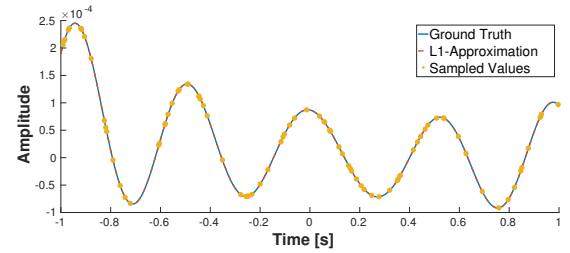

Fig. 1. Example of a reconstructed equivalent baseband signal (real part): M = 30; 90 uniformly distributed sampling points; $c = 30$
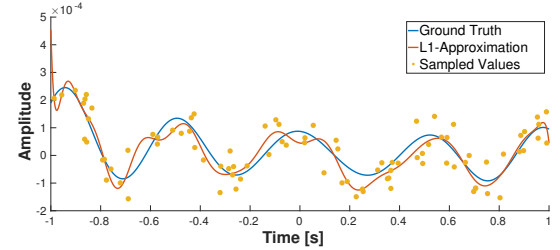

Fig. 2. Example of a reconstructed equivalent baseband signal (real part): M = 30; 90 uniformly distributed sampling points; $c = 30$; SNR ≈ 4, i.i.d. uniformly distributed noise on sampling values

4. CONCLUSION

We have discussed and derived methods for approximating bandlimited functions on the real line or in finite intervals from finitely many samples generated by a uniformly distributed sampling process, assuming sparsity in the PSWF basis. In the case that nearly all of the signal energy is concentrated in the interval of interest, the number of sampling points necessary is proportional to the sparsity in the PSWF basis. In the case that a significant portion of the signal energy is outside the interval of interest, we have derived results showing that a lower bound for the number of sampling points sufficient is proportional to M^2, assuming all coefficients up to M are to be recovered. Future work includes investigating the reason for this qualitative change in behavior when moving from the real line to finite intervals.

5. REFERENCES

[1] J. R. Higgins, *Sampling Theory in Fourier and Signal Analysis: Foundations*, Oxford University Press on Demand, 1996.

[2] H. D. Helms and J. B. Thomas, "Truncation error of sampling-theorem expansions," *Proceedings of the IRE*, vol. 50, no. 2, pp. 179–184, 1962.

[3] J. Brown, "Bounds for truncation error in sampling expansions of band-limited signals," *IEEE Transactions on Information Theory*, vol. 15, no. 4, pp. 440–444, 1969.

[4] J. Knab, "System error bounds for Lagrange polynomial estimation of band-limited functions," *IEEE Transactions on Information Theory*, vol. 21, no. 4, pp. 474–476, 1975.

[5] R. Radzyner and P. T. Bason, "An error bound for Lagrange interpolation of low-pass functions," *IEEE Transactions on Information Theory*, vol. 18, no. 5, pp. 669–671, 1972.

[6] D.M. Klamer and E. Masry, "Polynomial interpolation of randomly sampled bandlimited functions and processes," *SIAM Journal on Applied Mathematics*, vol. 42, no. 5, pp. 1004–1019, 1982.

[7] T. Strohmer and J. Tanner, "Fast reconstruction methods for bandlimited functions from periodic nonuniform sampling," *SIAM J. Numer. Anal.*, vol. 44, no. 3, pp. 1073–1094, 2006.

[8] J. Selva, "Functionally weighted Lagrange interpolation of band-limited signals from nonuniform samples," *IEEE Transactions on Signal Processing*, vol. 57, no. 1, pp. 168–181, 2009.

[9] S. Foucart and H. Rauhut, *A Mathematical Introduction to Compressive Sensing*, Springer New York, 2013.

[10] B. Adcock and A.C. Hansen, "Generalized sampling and infinite-dimensional compressed sensing," *Foundations of Computational Mathematics*, vol. 16, no. 5, pp. 1263–1323, 2016.

[11] L. Gosse, "Compressed sensing with preconditioning for sparse recovery with subsampled matrices of Slepian prolate functions," *Annali dell'Universita di Ferrara*, vol. 59, no. 1, pp. 81–116, 2013.

[12] D. Slepian and H.O. Pollak, "Prolate spheroidal wave functions, Fourier analysis and uncertainty - i," *Bell System Technical Journal, The*, vol. 40, no. 1, pp. 43–63, 1961.

[13] H. J. Landau and H. O. Pollak, "Prolate spheroidal wave functions, Fourier analysis and uncertainty - ii," 1961.

[14] H. J. Landau and H. O. Pollak, "Prolate spheroidal wave functions, Fourier analysis and uncertainty - iii: The dimension of the space of essentially time- and band-limited signals," *Bell System Technical Journal*, vol. 41, no. 4, pp. 1295–1336, 1962.

[15] M.A. Kowalski, "On approximation of band-limited signals," *Journal of Complexity*, vol. 5, no. 3, pp. 283–302, 1989.

[16] K. Yao, "Applications of reproducing kernel Hilbert spaces - bandlimited signal models," *Information and Control*, vol. 11, no. 4, pp. 429–444, 1967.

[17] A. Osipov, V. Rokhlin, and H. Xiao, *Prolate Spheroidal Wave Functions of Order Zero: Mathematical Tools for Bandlimited Approximation*, Springer US, 2013.

[18] B. Adcock, "Infinite-dimensional compressed sensing and function interpolation," *Foundations of Computational Mathematics*, Apr 2017.

[19] G. Puy, M.E. Davies, and R. Gribonval, "Recipes for stable linear embeddings from Hilbert spaces to $\mathbb{R}^m$," *IEEE Transactions on Information Theory*, vol. 63, no. 4, pp. 2171–2187, 2017.

[20] H.G. Feichtinger and K. Gröchenig, "Theory and practice of irregular sampling," *Wavelets: mathematics and applications*, pp. 305–363, 1994.

[21] R. Adelman, N. A. Gumerov, and R. Duraiswami, "Software for computing the spheroidal wave functions using arbitrary precision arithmetic," *ArXiv e-prints*, Aug. 2014.