View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Apollo

Sampling from binary measurements - On Reconstructions from Walsh coefficients

Anders Hansen
University of Cambridge
DAMTP
Email: a.hansen@damtp.cam.ac.uk

Laura Terhaar
University of Cambridge
DAMTP
Email: lt420@cam.ac.uk

***Abstract*—Reconstructing infinite-dimensional signals from a limited amount of linear measurements is a key problem in many applications such as medical imaging [35], single-pixel and lensless cameras [27], fluorescence microscopy [39] etc. Efficient techniques for such a problem include generalized sampling [6], [23], [31], [43] and its compressed versions [5], [27], as well as methods based on data assimilation [9], [11], [20]. All of these methods have in common that the reconstruction quality depends highly on the subspace angle between the sampling and the reconstruction space. In this paper we consider the case of binary measurements, which, after a standard subtraction trick, can be converted to a 1 and -1 setup. These measurements are modelled with Walsh functions, which form the kernel for the Hadamard transform. For the reconstruction we use wavelets. We show that the relation between the amount of data sampled and the coefficients reconstructed has to be only linear to ensure that the angle is bounded from below and hence the reconstruction is accurate and stable.**

I. Introduction

The theory of sampling has a wide history and is still a highly popular research topic. In 1948 Shannon introduced the Shannon Sampling Theorem [32], [43], [44], which is now a classical result in sampling theory. The message of the theorem is that for functions with a bounded bandwidth, one can sample pointwise at a certain rate, and the function can be represented as an infinite sum over the measurements times the sinc functions. Although a beautiful mathematical result, this approach is not always feasible in reality. The coefficients usually do not decay fast, which results in slow convergence and unpleasant artefacts [25], [40], [44].

New possibilities in sampling theory were introduced by the systems of wavelets [37], shearlets [18], [19], [33], curvelets [14]–[16] and contourlets [21], [41]. These systems allow much better representation than the sinc functions or complex exponentials. However, it is in most applications not possible to change the measurements. For example, a Magnetic Resonance Imaging (MRI) scanner does provide Fourier samples, and cannot easily be tweaked to provide wavelet, curvelet or shearlet coefficients.

The amount of research done in the last decades on utilising other bases and frames in sampling theory is comprehensive. Concepts like the finite section method [12], [29], [30], [34], consistent sampling by Unser and Aldroubi [8], [45] as well as the generalization by Eldar [22]–[24], [26] are now widely known and understood. Nevertheless, they have certain drawbacks in practice. Although these approaches may work well in certain cases [23], [45], they may also be highly unstable and even non-convergent for particular problems.

These drawbacks can be resolved by different approaches. For example, generalized sampling [1], [6], [7], [31], which allows a different amount of samples and reconstructed coefficients, leads to a stable and quasi-optimal reconstruction. Another new and promising method is the data assimilation in reduced modelling approach introduced by Binev, Cohen, Dahmen, DeVore, Petrova and Wojtaszczyk [9], [11], [20]. These methods have in common that the subspace angle

$$\cos(\omega(U,V)) = \inf_{u\in U, \|u\|=1} \|P_V u\| = \frac{1}{\mu(U,V)},$$

$\omega(U,V) \in [0,\pi/2]$, between the reconstruction and sampling space U and V respectively, is crucial in the estimation of the reconstruction error. In fact, the error and stability depends directly on this quantity.

Both methods mentioned can take any sample and reconstruction space, which allows substantial freedom. In the case of Fourier measurements, which are the basis for MRI, Helium atom scattering, radio interferometry etc., there has been quite a lot of recent developments. It is possible to show that the sampling rate, i.e. the amount of samples needed for a certain amount of stable reconstructed coefficients, is linear for Fourier measurements to wavelets [2], [4] in one and two dimensions. In addition, there are sharp bounds for the stable sampling rate in the case of Fourier measurements and wavelets [3]. The two dimensional case could even be extended to shearlets [36], which do not form a basis but rather a frame. The stable sampling rate was also investigated for Fourier samples and polynomial reconstruction, where one gets a quadratic relation [31].

Beside the research done in the field of Fourier measurements, there is also a lot of research done in the area of dyadic analysis and the Walsh functions, which we will use to model the binary measurements. A great overview is given by Gauss in [28], further information can be found in [10], [13], [38], [42]. After a change of the ordering, Walsh functions are the kernel for the well known Hadamard transform. The new ordering is obtained by applying the Gray transform to the input parameter. Due to there direct relation with the

Hadamard transform, Walsh functions are frequently used in the context of binary measurements. As binary measurements are the mainstay in many imaging applications such as lensless camera and single-pixel cameras [27] and fluorescence microscopy [39] it is natural to ask how the stable sampling rate behaves for reconstructions from Walsh coefficients with wavelets, shearlets, polynomials or splines.

The main contribution of this paper is the analysis of the subspace angle for binary measurements via Walsh functions and wavelets, where we show that the stable sampling rate (1) is linear. This means for a linear relation between the amount of samples and reconstructed coefficients, the subspace angle is bounded from below. Moreover, this result is sharp. This might be surprising at first, because of the very different structure of Walsh functions and wavelets. Walsh functions provide a lot of structure and properties, i.e. smoothness, in terms of the dyadic analysis, but not in terms of the decimal analysis. Wavelets behave the other way around. Most wavelets are continuous and have regularity properties in the decimal analysis but not in the dyadic analysis. This is one of the main differences in contrast to the Fourier-wavelet reconstruction. In the setting of Walsh functions we are restricted to the space $\mathrm{L}^2([0,1]^d)$. This means that the wavelets have to be adapted accordingly. That is the reason why we are dealing with boundary wavelets.

II. Background

A. Reconstruction Methods

In this paper we are dealing with the problem of reconstructing data from given measurements with linear functionals. In particular, we consider a Hilbert space $\mathcal{H}$ and linear functionals $(m_i)_{i\in\mathbb{N}} : \mathcal{H} \to \mathbb{R}$ that can be represented by elements $s_i \in \mathcal{H}$ as $m_i(f) = \langle f, s_i\rangle$. The space of these functions is called the *sampling space* and is denoted by $\mathcal{S} = \overline{\mathrm{span}}\{s_i : i \in \mathbb{N}\}$, meaning the closure of the span. In practice, one can only acquire a finite amount of samples, therefore we denote by $\mathcal{S}_M = \mathrm{span}\{s_i : i = 1, \dots, M\}$, the sampling space of the first M elements. A suitable choice of *reconstruction spaces* is widely investigated. The reconstruction space may typically fit better the data to be recovered than the sampling space, which is usually dictated by the physics of the sampling device (fluorescence microscope, MRI scanner, electron microscope etc.). For example, spaces spanned by X-lets (wavelets, curvelets, shearlets) may be preferable as reconstruction spaces in imaging applications, whereas polynomials may be useful when considering very smooth functions to be recovered. The reconstruction spaces are denoted by $\mathcal{R}$ and spanned by reconstruction functions $(r_i)_{i\in\mathbb{N}}$, i.e. $\mathcal{R} = \overline{\mathrm{span}}\{r_i : i \in \mathbb{N}\}$. As in the case of the sampling space, it is impossible to acquire and to save an infinite amount of reconstruction coefficients. Hence, one has to restrict to a finite reconstruction space, which is denoted by $\mathcal{R}_N = \mathrm{span}\{r_i : i = 1, \dots, N\}$.

As mentioned above the two methods based on generalized sampling and reduced modelling are similar reconstruction methods that offer an alternative to the existing literature. For the generalized sampling approach it is possible to use a different amount of samples M and coefficients that are reconstructed N. This leads to substantial improvements on stability and convergence. The method can be described as follows, for $f \in \mathcal{H}$ and $N, M \in \mathbb{N}$, we define the reconstruction method of *generalized sampling* $G_{N,M} : \mathcal{H} \to \mathcal{R}_N$ by

$$\langle P_{\mathcal{S}_M} G_{N,M}(f), r_j\rangle = \langle P_{\mathcal{S}_M} f, r_j\rangle, \quad r_j \in \mathcal{R}_N,$$

where $P_{\mathcal{S}_M}$ denotes the orthogonal projection on the subspace $\mathcal{S}_M$. Note that the stability and accuracy of this method depends on the subspace angle between the sampling and the reconstruction space, i.e.

$$\|f - G_{N,M}(f)\| \le \mu(\mathcal{R}_N, \mathcal{S}_M)\|f - P_{\mathcal{R}_N} f\|.$$

Moreover, the condition number κ of $G_{N,M}$ is also given by $\kappa(G_{N,M}) = \mu(\mathcal{R}_N, \mathcal{S}_M)$.

In the approach presented in [20] based on data assimilation in reduced modelling the proposed algorithm calculates

$$F_{N,M}(f) = \operatorname*{argmin}_{u \in P_{\mathcal{S}_M} f + \mathcal{S}_M^\perp} \|u - P_{\mathcal{R}_N} u\|,$$

and it can be shown that the accuracy then depends on subspace angle

$$\|f - F_{N,M}(f)\| \le \mu(\mathcal{R}_N, \mathcal{S}_M)\, \mathrm{dist}(f, \mathcal{R}_N \oplus (\mathcal{S}_M \cap \mathcal{R}_N)^\perp).$$

Moreover, this is sharp in the way that the constant $\mu(\mathcal{R}_N, \mathcal{S}_M)$ cannot be improved. It is clear that, in both approaches, the key to success lies in the ability to make sure that

$$\mu(\mathcal{R}_N, \mathcal{S}_M) \le \theta, \qquad \theta \in \mathbb{R}_+.$$

Thus, we need to balance the number of samples M with the number of reconstruction vectors N, and this lead to the so-called *stable sampling rate*:

$$\Theta(N, \theta) = \min\{M \in \mathbb{N} : \mu(\mathcal{R}_N, \mathcal{S}_M) \le \theta\}. \tag{1}$$

The methods above can only be used efficiently when the stable sampling rate is known. Fortunately, it is possible to obtain sharp results on describing $\Theta(N, \theta)$ for popular sampling and reconstruction spaces. In this paper we do so for sampling with Walsh functions and reconstructing with wavelets.

B. Wavelets and Walsh functions

The sampling space used in this paper is the space of Walsh functions whereas the reconstruction takes place in the space of wavelets. Due to the structure of Walsh functions, i.e. that they only span $\mathrm{L}^2([0,1]^d)$ we have to adjust the wavelet space accordingly, such that $\mathcal{S}^\perp \cap \mathcal{R} = \{0\}$. Hence, we restrict the reconstruction space to boundary wavelets with compact support in $[0,1]^d$.

Due to the fact that we are dealing with the d-dimensional case, we introduce multi indices to make the notation more readable. Let $j = \{j_1, \dots, j_d\} \in \mathbb{N}^d$, $d \in \mathbb{N}$ be a multi index. A natural number n is in the context with a multi index interpreted as a multi index with the same entry, i.e. $n = \{n, \dots, n\}$. Then we define the addition of two

multi indices for $j, r \in \mathbb{N}^d$ by the pointwise addition, i.e. $j + r = \{j_1 + r_1, \ldots, j_d + r_d\}$ and the sum $\sum_{j=k}^{r} x_j := \sum_{j_1=k_1}^{r_1} \ldots \sum_{j_d=k_d}^{r_d} x_{j_1,\ldots,j_d}$, where $k, r \in \mathbb{N}^d$. The multiplication of an multi index with a real number is understood pointwise, as well as the division by a multi index.

With this at hand we can now give a definition of boundary wavelets and the Walsh functions in higher dimensions. The approach to construct wavelets on the interval was introduced by Cohen, Daubechies and Vial [17]. It starts with the general Daubechies wavelets of order p. These and their scaling function have support size $2p - 1$. For the construction we start with the MRA structure at level J_0 to make sure that the scaling function intersect at that level only with the left or the right boundary. Then one takes the interior scaling functions and replaces those who intersect with the boundary by linear combinations of the original scaling functions, such that the amount of vanishing moments is preserved as well as smoothness properties of the original function. For the right end of the interval the original scaling function has to be reflected. This leads to the following sampling space in one dimension.

$$V_j^b = \operatorname{span}\{\ \phi_{j,k} : k = 0, \ldots 2^j - p - 1, \\ \phi_{j,k}^{\#} : k = 2^j - p, \ldots, 2^j - 1\Big\}$$

where $\phi^{\#}$ is the reflected scaling function, $j \geq J_0$ and $\phi_{j,k}$ denotes the standard wavelet notation, see [17] for further details. In higher dimensions the sampling space is obtained by the tensor product, an overview is given in [37]. We define $V_j^{b,d} = V_j^b \otimes \ldots \otimes V_j^b$ (d times), which preserves the MRA structure. The according wavelet space is denoted by $W_j^{b,d}$, such that the reconstruction space $\mathcal{R}_N$, $N = 2^{dR}$ can be represented by

$$\mathcal{R}_N = V_{J_0}^{b,d} \oplus W_{J_0}^{b,d} \oplus \ldots \oplus W_{R-1}^{b,d} = V_R^{b,d}.$$

For the representation of elements in $V_R^{b,d}$ we define the sets $K_0 = \{0, \ldots, 2^R - p - 1\}$ and $K_1 = \{2^R - p \ldots 2^R - 1\}$. Denote the original scaling function by $\phi^0 := \phi$ and the reflected scaling function by $\phi^1 := \phi^{\#}$. Then every element $\varphi \in V_R^{b,d}$ can be represented by

$$\varphi = \sum_{s \in \{0,1\}^d} \sum_{k \in K_s} \alpha_k \phi_{R,k}^s, \qquad \alpha_k \in \mathbb{C},$$

where

$$\phi_{R,k}^s = \phi_{R,k_1}^{s_1} \otimes \ldots \otimes \phi_{R,k_d}^{s_d}.$$

The Walsh functions in higher dimensions can be represented by the tensor product of one dimensional Walsh functions.

Definition 1 (Walsh function)**.** *Let $s = \sum_{i \in \mathbb{Z}} s_i 2^{i-1}$ with $s_i \in \{0, 1\}$ be the dyadic expansion of $s \in \mathbb{R}_+$. Analogously, let $x = \sum_{i \in \mathbb{Z}} x_i 2^{i-1}$ with $x_i \in \{0, 1\}$. The generalized Walsh functions in $L^2([0, 1])$ are given by*

$$\operatorname{Wal}(s, x) = (-1)^{\sum_{i \in \mathbb{Z}} (n_i + n_{i+1}) x_{-i-1}}.$$

We extend it to functions in $\mathrm{L}^2([0, 1]^d)$ by the tensor product for $s = (s_k)_{k=1,\ldots,d}, x = (x_k)_{k=1,\ldots,d}$

$$\operatorname{Wal}(s, x) = \bigotimes_{k=1}^{d} \operatorname{Wal}(s_k, x_k).$$

These functions then span the sampling space, i.e. for $M = m^d, m \in \mathbb{N}$ we have

$$\mathcal{S}_M = \operatorname{span}\{\operatorname{Wal}(s, \cdot), s_i = 1, \ldots, m, i = 1, \ldots, d\}.$$

Moreover, Walsh functions can be extended to negative inputs by $\operatorname{Wal}(-s, x) = \operatorname{Wal}(s, -x) = -\operatorname{Wal}(s, x)$.

With the help of the Walsh functions we can define the continuous Walsh transform of a function ϕ.

$$\widehat{\phi}^W(s) = \langle \phi(\cdot), \operatorname{Wal}(s, \cdot) \rangle = \int_{[0,1]^d} \phi(x) \operatorname{Wal}(s, x) dx, \ s \in \mathbb{R}^d.$$

The Walsh function have the following very useful properties. First, they obey the scaling property, i.e. $\operatorname{Wal}(2^j s, x) = \operatorname{Wal}(s, 2^j x)$ for all $j \in \mathbb{N}$ and $s, x \in \mathbb{R}$. Second, the multiplicative identity holds, this means $\operatorname{Wal}(s, x) \operatorname{Wal}(s, y) = \operatorname{Wal}(s, x \oplus y)$, where $\oplus$ is the dyadic addition. These properties are also easily transferred to the Walsh transform.

III. Main result

The main result of this paper is that the stable sampling rate, when considering Walsh samples and wavelet reconstruction, is linear. This means that when using the methods discussed above, one obtains results that are, up to a constant, as good as if one could access the wavelet coefficients directly.

Theorem 1. *Let $\mathcal{S}$ and $\mathcal{R}$ be the sampling and reconstruction spaces of Walsh functions and boundary wavelets in $\mathrm{L}^2([0, 1]^d)$. If $N = 2^{dR}$ where $R \in \mathbb{N}$, then for every $\theta \in (1, \infty)$ there exist a constant S_θ depending on θ such that $\Theta(N, \theta) \leq S_\theta N = \mathcal{O}(N)$ for any $\theta \in (1, \infty)$, i.e. the stable sampling rate is linear.*

Before showing a sketch of the proof, we will state three Lemmas which are essential. The first one gives a relation between the discrete Walsh transform and the original signal, where the normalization is used slightly different to the common literature to enable the use for Lemma 2.

Lemma 1. *Let $N = \{N_i\}_{i=1,\ldots,d}, N_i = 2^{n_i}, n_i \in \mathbb{N}, i = 1, \ldots, d$ and $x = \{x_k\}$, where $x_k \in \mathbb{R}$, $k = \{k_i\}_{i=1,\ldots,d}, k_i = 0, \ldots, N_i - 1$ and $x \in \mathbb{R}^{N_1 \times \ldots \times N_d}$, be a discrete d-dimensional signal. The discrete d-dimensional Walsh transform is given by $X = \{X_j\}$, where $X_j \in \mathbb{R}$, $j = \{j_i\}_{i=1,\ldots,d}, j_i = 0, \ldots, N_i - 1$ and $X \in \mathbb{R}^{N_1 \times \ldots \times N_d}$, with*

$$X_j = \frac{1}{\prod_{i=1}^d N_i} \sum_{k=0}^{N-1} x_k \operatorname{Wal}(j, \frac{k}{N}).$$

Then,

$$\sum_{j=0}^{N-1} |X_j|^2 = \frac{1}{\prod_{i=1}^d N_i} \sum_{k=0}^{N-1} |x_k|^2$$

holds.

In the proof of the theorem we achieve a bound by the scaling function under the Walsh transform and the sum over the Walsh polynomial. The following lemma helps to estimate the value of the Walsh polynomial and is related to a similar result for trigonometric polynomials.

Lemma 2. *Let $A, B \in \mathbb{Z}^d$ such that $A_i \leq B_i, i = 1, \dots, d$ and consider the Walsh polynomial $\Phi(z) = \sum_{j=A}^{B} \alpha_j \operatorname{Wal}(j, z)$. If $L = \{L_1, \dots, L_d\}$ with $L_i = 2^{n_i}$, $n_i \in \mathbb{N}, i = 1, \dots, d$ such that $2L_i \geq B_i - A_i + 1$, then*

$$\sum_{j=0}^{2L-1} \frac{1}{\prod_{i=1}^{d} 2L_i} \left| \Phi(\frac{j}{2L}) \right|^2 = \sum_{j=A}^{B} |\alpha_j|^2.$$

To be able to use the scaling property for wavelets, we have to adjust their support. In detail this means, that we represent the scaling function as a sum of functions that are supported in $[0, 1]$, i.e.

$$\phi(x) = \sum_{i=-p+2}^{p} \phi_i(x-i+1) \text{ with } \phi_i(x) = \phi(x+i-1)\mathcal{X}_{[0,1]}(x)$$

and

$$\phi_{R,l} = 2^{R/2} \sum_{i=-p+2}^{p} \phi_i(2^R x - i + 1 - l).$$

This can also be done accordingly for the reflected function $\phi^{\#}$. In the higher dimensional case we have

$$\begin{aligned} \phi(x) &= (\phi_1 \otimes \dots \otimes \phi_d)(x_1, \dots, x_d) \\ &= \sum_{i=-p+2}^{p} \phi_{i_1}(x_1 - i_1 + 1) \cdot \dots \cdot \phi_{i_d}(x_d - i_2 + 1) \\ &= \sum_{i=-p+2}^{p} \phi_i(x - i + 1). \end{aligned}$$

and ϕ_{i_j} defined as above. This way the multiplicative identity holds also for the decimal time shift of the wavelets.

Next, we present the decay rate of continuous wavelets under the Walsh transform. This is not straightforward as the Walsh transform is analysing smoothness in the dyadic analysis which is very different from smoothness in the decimal analysis. Nevertheless, a large class of wavelets fulfil this property as well.

Lemma 3. *Let ϕ be a Daubechies scaling function of p vanishing moments $p > 1$ and ϕ_i be the scaling function multiplied by the cut off function. Then*

$$|\widehat{\phi_i}^W(k)| \leq \frac{A}{\prod_{i=1}^{d}(1 + |k_i|)} \quad k \in \mathbb{R}_+^d, \quad i = -p+2, \dots, p$$

for some $A \in \mathbb{R}_+$.

With this information together we can give an overview of the proof.

Sketch of the proof. For the proof we try to bound the value $C_{N,M} = \cos(\omega(\mathcal{R}_N, \mathcal{S}_M)) = \inf_{f \in \mathcal{R}_N, ||f||=1} ||P_{\mathcal{S}_M} f|| \geq 1 - ||P_{\mathcal{S}_M}^{\perp} \varphi||$, where $P_{\mathcal{S}_M}$ is the orthogonal projection on the subspace $\mathcal{S}_M$. Then we use the assumptions on the wavelets to represent $\varphi \in \mathcal{R}_N$ with $||\varphi|| = 1$ by

$$\varphi = \sum_{s \in \{0,1\}^d} \sum_{k \in K_s} \alpha_k \phi_{R,k}^s.$$

Due to the orthogonal projection we need to handle the values $\langle \phi_{R,n}, \operatorname{Wal}(l, \cdot) \rangle$. For this purpose we use the functions ϕ_i defined as before. As this enables us to transfer the scaling and shift from the wavelet over to the Walsh function. Moreover, let $l = \{l_1, \dots, l_d\} \in \mathbb{Z}^d$ and define $p_R : \mathbb{Z}^d \to \mathbb{N}^d$ with $\{z_i\}_{i=1,\dots,d} = z \mapsto p_R(z) = \{p_R(z)_i\}_{i=1,\dots,d}$ and $p_R(z)_i$ be the smallest integer such that $p_R(z)_i 2^R - z_i > 0$ for $i = 1, \dots, d$. Then

$$\begin{aligned} &\langle \phi_{i,R,n}^s, \operatorname{Wal}(l, \cdot) \rangle \\ &\quad = \operatorname{Wal}(n + i - 1 + 2^R p_R(n + i - 1), \frac{l}{2^R}) \widehat{\phi_i^s}^W(\frac{l}{2^R}). \end{aligned}$$

Define now

$$\Phi_i^s(z) = \sum_{n \in K_s} \alpha_n \operatorname{Wal}(n + i - 1 + 2^R p_R(n + i - 1), \frac{z}{2^R})$$

Then, one can estimate the parts of the orthogonal projection by

$$\sum_{n \in K_s} \alpha_n \langle \phi_{i,R,n}^s, \operatorname{Wal}(l, \cdot) \rangle = (2^{-R/2})^d \Phi_i^s(\frac{l}{2^R}) \widehat{\phi_i^s}^W(\frac{l}{2^R})$$

With $M = \{M_1, \dots, M_d\}$ and the set $\mathcal{I}_M = \left\{ l \in \mathbb{N}^d, l_i \leq M_i \right\}$ this allows to see

$$||P_{\mathcal{S}_M}^{\perp} \varphi|| \leq \sum_{i=-p+2}^{p} \sqrt{\sum_{l \notin I_M} 2^{-dR} \left| \sum_{s \in \{0,1\}^d} \Phi_i^s(\frac{l}{2^R}) \widehat{\phi_i^s}^W(\frac{l}{2^R}) \right|^2}.$$

Now, let $S \in \mathbb{N}$ be given, such that the amount of samples $M = \{M_1, \dots, M_d\} \in \mathbb{N}^d$ is $M_k = S2^R$. Then, if $l = \{l_1, \dots, l_d\} \notin I_M$ at least one $l_k > M_k$. The sum is the largest, if only one l_k is fulfils this estimate. Hence, without loss of generality let $l_1 > M_1$ and $l_k \leq M_k$ for $k = 2, \dots, d$. Let now $l_k = j_k + S2^R$. Then for some constant C the sumands over $s \in \{0, 1\}^d$ can be estimated by the periodicity of the Walsh functions and Lemma 3 by

$$\begin{aligned} &\sum_{l_1 > M_1} \sum_{l_2 \leq M_2} \dots \sum_{l_d \leq M_d} 2^{-dR} \left| \Phi_i^s(\frac{l}{2^R}) \widehat{\phi_i^s}^W(\frac{l}{2^R}) \right|^2 \\ &\quad \leq \frac{C^{d-1}}{(2\alpha - 1) S^{2\alpha - 1}} \sum_{j=0}^{2^R - 1} \frac{1}{2^{dR}} \left| \Phi_i^s(\frac{j}{2^R}) \right|^2 \end{aligned}$$

The last sumand can be controlled with Lemma 2 and the fact, that $||\varphi|| = 1$

$$\sum_{j=0}^{2^R - 1} \frac{1}{2^{dR}} \left| \Phi_i(\frac{j}{2^R}) \right|^2 \leq 1.$$

All together allows to estimate

$$C_{N,M} \leq \theta$$

for all $M \geq SN$ for S large enough independent of N dependent on θ. □

IV. Conclusion

We were able to investigate a very important part of the error estimate for different reconstruction methods. Moreover, we showed that binary measurements modelled by Walsh functions are well suited to reconstruct images with wavelets. This gives together with the results in [2], [4] provide a broad knowledge about the accuracy and stability for two major applications of sampling theory, i.e. systems with Fourier samples and those with binary measurements.

References

[1] B. Adcock and A. Hansen. Stable reconstructions in hilbert spaces and the resolution of the gibbs phenomenon. *Appl. Comput. Harmon. Anal.*, 32(3):357388, 2012.

[2] B. Adcock, A. Hansen, G. Kutyniok, and J. Ma. Linear stable sampling rate: Optimality of 2d wavelet reconstructions from fourier measurements. *SIAM J. Math. Anal.*, 47(2):1196–1233, 2015.

[3] B. Adcock, A. Hansen, and C. Poon. Beyond consistent reconstructions: optimality and sharp bounds for generalized sampling, and application to the uniform resampling problem. *SIAM J. Math. Anal.*, 45(5):3132–3167, 2013.

[4] B. Adcock, A. Hansen, and C. Poon. On optimal wavelet reconstructions from fourier samples: linearity and universality. *Appl. Comput. Harmon. Anal.*, 36(3):387–415, 2014.

[5] B. Adcock, A. Hansen, C. Poon, and B. Roman. Breaking the coherence barrier: A new theory for compressed sensing. *preprint, arXiv:1302.0561v4*, 2014.

[6] B. Adcock and A. C. Hansen. A generalized sampling theorem for stable reconstructions in arbitrary bases. *J. Fourier Anal. Appl.*, 18(4):685–716, 2010.

[7] B. Adcock and A. C. Hansen. Reduced consistency sampling in hilbert spaces. *In Proceedings of the 9th International Conference on Sampling Theory and Applications*, 2011.

[8] A. Aldroubi and M. Unser. A general sampling theory for nonideal acquisition devices. *IEEE Trans. Signal Process.*, 42(11):2915–2925, 1994.

[9] M. Bachmayr, A. Cohen, R. deVore, and G. Migliorati. Sparse polynomial approximation of parametric elliptic pdes. part ii: lognormal coefficients. *arXic:1509.07050*, 2015.

[10] K. Beauchamp. *Walsh Functions and their Applications*. Academic Press, London, 1975.

[11] P. Binev, A. Cohen, W. Dahmen, R. deVore, G. Petrova, and P. Wojtaszczyk. Data assimilation in reduced modeling. *SIAM UQ, to appear*, 2015.

[12] A. Böttcher. Infinite matrices and projection methods: in lectures on operator theory and its applications, fields inst. monogr. *Amer. Math. Soc.*, (3):1–72, 1996.

[13] P. Butzer and H. Wagner. On the dyadic analysis based on pointwise dyadic derivative. *Anal. Math.*, (1):171–196, 1975.

[14] E. J. Candès and L. Demanet. Curvelets and fourier integral operators. *C. R. Acad. Sci.*, 336(1):395–398, 2003.

[15] E. J. Candès and D. Donoho. New tight frames of curvelets and optimal representations of objects with piecewise c^2 singularities. *Comm. Pure Appl. Math.*, 57(2):219–266, 2004.

[16] E. J. Candès and L. Donoho. Recovering edges in ill-posed inverse problems: optimality of curvelet frames. *Ann. Statist.*, 30(3):784–842, 2002.

[17] A. Cohen, I. Daubechies, and P. Vial. Wavelets on the interval and fast wavelet transforms. *Comput. Harmon. Anal.*, 1(1):54–81, 1993.

[18] S. Dahlke, G. Kutyniok, P. Maass, C. Sagiv, H.-G. Stark, and G. Teschke. The uncertainty principle associated with the continuous shearlet transform. *Int. J. Wavelets Multiresolut. Inf. Process.*, 2(6), 2008.

[19] S. Dahlke, G. Kutyniok, G. Steidl, and G. Teschke. Shearlet coorbit spaces and associated banach frames. *Appl. Comput. Harmon. Anal.*, 2(27), 2009.

[20] R. deVore, G. Petrova, and P. Wojtaszczyk. Data assimilation and sampling in banach spaces. *arXiv:1602.06342*, 2016.

[21] M. N. Do and M. Vetterli. The contourlet transform: An efficient directional multiresolution image representation. *IEEE Transactions on Image Processing*, 12(14), 2005.

[22] T. Dvorkind and Y. C. Eldar. Robust and consistent sampling. *IEEE Signal Process. Letters*, 16(9):739–742, 2009.

[23] Y. C. Eldar. Sampling with arbitrary sampling and reconstruction spaces and oblique dual frame vectors. *J. Fourier Anal. Appl.*, 9(1):77–96, 2003.

[24] Y. C. Eldar. Sampling without input constraints: Consistent reconstruction in arbitrary spaces. *Sampling, Wavelets and Tomography*, 2003.

[25] Y. C. Eldar and T. Michaeli. Beyond bandlimited sampling. *IEEE Signal Process. Mag.*, 3(826):48–68, 2009.

[26] Y. C. Eldar and T. Werther. General framework for consistent sampling in hilbert spaces. *Int. J. Wavelets Multiresolut. Inf. Process.*, 3(3):347, 2005.

[27] S. Foucart and H. Rauhut. *A Mathematical Introduction to Compressive Sensing*. Springer Science+Business Media, New York, 2013.

[28] E. Gauss. *Walsh Funktionen für Ingenieure und Naturwissenschaftler*. Springer Fachmedien, Wiesbaden, 1994.

[29] K. Gröchenig, Z. Rzeszotnik, and T. Strohmer. Quantitative estimates for the finite section method and banach algebras of matrices. *Integral Equations and Operator Theory*, 2(67):183–202, 2011.

[30] A. C. Hansen. On the approximation of spectra of linear operators on hilbert spaces. *J. Funct. Anal.*, 8(254):2092–2126, 2008.

[31] T. Hrycak and K. Gröchenig. Pseudospectral fourier reconstruction with the modified inverse polynomial reconstruction method. *J. Comput. Phys.*, 229(3):933–946, 2010.

[32] A. J. Jerri. The shannon sampling theorem its various extensions and applications: A tutorial review. *Proc. IEEE*, (65):1565–1596, 1977.

[33] G. Kutyniok and W.-Q. Lim. Compactly supported shearlets are optimally sparse. *J. Approx. Theory*, (163):1564–1589, 2011.

[34] M. Lindner. *Infinite matrices and their finite sections: An introduction to the limit operator method*. Birkhäuser Verlag, Basel, 2006.

[35] M. Lustig, D. Donoho, and J. M. Pauly. Sparse mri: the application of compressed sensing for rapid mr imaging. *Magnetic Resonance in Medicine*, 2007.

[36] J. Ma. Generalized sampling reconstruction from fourier measurements using compactly supported shearlets. *Appl. Comput. Harmon. Anal.*, 2015.

[37] S. Mallat. *A wavelet tour of signal processing*. Academic Press, San Diego, 1998.

[38] M. Maqusi. *Applied Walsh Analysis*. Heyden & Son Ltd, Renfrew and London, 1981.

[39] M. Müller. *Introduction to Confocal Fluorescence Microscopy*. SPIE, Bellingham, Washington, 2006.

[40] J. A. Parker, R. V. Kenyon, and D. E. Troxel. Comparison of interpolating methods for image resampling. *IEEE Trans. Med. Imaging*, 1(MI-2):31–39, 1983.

[41] D. D.-Y. Po and M. N. Do. Directional multiscale modeling of images using the contourlet transform. *Trans. Img. Proc.*, 6(15), 2006.

[42] F. Schipp, P. Simon, and W. Wade. *Walsh Series an Introduction to dyadic haromonic Analysis*. Adam Hilger, Bristol and New York, 1990.

[43] C. Shannon. A mathematical theory of communication. *Bell Syst. Tech. J.*, (27):379–423, 623–656, 1948.

[44] M. Unser. Sampling50 years after shannon. *Proc. IEEE*, 4(88):569–587, 2000.

[45] M. Unser and J. Zerubia. A generalized sampling theory without bandlimiting constraints. *IEEE Trans. Circuits Syst. II.*, 45(8):959–969, 1998.