ADAPTIVE SIGNAL AND SYSTEM APPROXIMATION AND STRONG DIVERGENCE

*Holger Boche**

Technische Universität München
Lehrstuhl für Theoretische Informationstechnik

Ullrich J. Mönich†

Massachusetts Institute of Technology
Research Laboratory of Electronics

ABSTRACT

Many divergence results for sampling series are in terms of the limit superior and not the limit. This leaves the possibility of a convergent subsequence. If there exists a convergent subsequence, adaptive signal processing techniques can be used. In this paper we study sampling-based signal reconstruction and system approximation processes for the space $\mathcal{PW}_\pi^1$ of bandlimited signals with absolutely integrable Fourier transform. For all analyzed examples, which include the peak value of the Shannon and the conjugated Shannon sampling series, we prove strong divergence, i.e., divergence for all subsequences. Hence, adaptive signal processing techniques do not help in these cases. We further analyze whether an adaptive choice of the reconstruction functions in the oversampling case can improve the behavior.

Index Terms— strong divergence, Paley–Wiener space, linear time-invariant system, reconstruction, Hilbert transform

1. INTRODUCTION

Sampling theory studies the reconstruction of a signal in terms of its samples. In addition to its mathematical significance, sampling theory plays a fundamental role in modern signal and information processing because it is the basis for today's digital world [1].

The fundamental initial result of the theory states that the Shannon sampling series can be used to reconstruct bandlimited signals f with finite energy from their samples $\{f(k)\}_{k\in\mathbb{Z}}$. Since this initial result, many different sampling theorems have been developed, and determining the signal classes for which the theorems hold and the mode of convergence now constitute an entire area of research [2–5].

In this paper we study the convergence behavior of different sampling series for the Paley–Wiener space $\mathcal{PW}_\pi^1$ consisting of absolutely integrable bandlimited signals. Analyzing sampling series and finding sampling theorems for the Paley–Wiener space $\mathcal{PW}_\pi^1$ has a long tradition [4, 6, 7]. Since Shannon's initial result, efforts have been made to extend it to larger signal spaces [6, 8, 9].

Before we state our main results, we introduce and motivate the problems. Let

$$(S_N f)(t) = \sum_{k=-N}^{N} f(k) \frac{\sin(\pi(t-k))}{\pi(t-k)} \tag{1}$$

denote the finite Shannon sampling series. It is well known that $S_N f$ converges locally uniformly to f for all signals $f \in \mathcal{PW}_\pi^1$ as N tends to infinity [4, 6, 7]. However, the series is not globally uniformly convergent. The quantity

$$P_N f = \max_{t\in\mathbb{R}} |f(t) - (S_N f)(t)|,$$

i.e., the peak value of the reconstruction error, diverges for certain $f \in \mathcal{PW}_\pi^1$ as N tends to infinity: in [10] it has been shown that there exists a signal $f \in \mathcal{PW}_\pi^1$ such that $\limsup_{N\to\infty} P_N f = \infty$. The divergence is only given in terms of the $\limsup$. In a sense this is a weak notion of divergence, because it merely states the existence of a subsequence $\{N_n\}_{n\in\mathbb{N}}$ of the natural numbers such that $\lim_{n\to\infty} P_{N_n} f = \infty$. This leaves the possibility that there is a different subsequence $\{N_n^*\}_{n\in\mathbb{N}}$ such that $\lim_{n\to\infty} P_{N_n^*} f = 0$.

This was discussed in [11], and two conceivable situations were phrased in two questions.

Question Q1: Does there for every $f \in \mathcal{PW}_\pi^1$, exist a subsequence $\{N_n\}_{n\in\mathbb{N}} = \{N_n(f)\}_{n\in\mathbb{N}}$ of the natural numbers such that $\sup_{n\in\mathbb{N}} P_{N_n} f < \infty$?

Question Q2: Does there exist a subsequence $\{N_n\}_{n\in\mathbb{N}}$ of the natural numbers such that $\sup_{n\in\mathbb{N}} P_{N_n} f < \infty$ for all $f \in \mathcal{PW}_\pi^1$?

Note that the subsequence $\{N_n(f)\}_{n\in\mathbb{N}}$ in question Q1 can depend on the signal f that shall be reconstructed. Thus, the reconstruction process $S_{N_n(f)}$ is adapted to the signal f. The problem of finding an index sequence, depending on the signal f, that is suitable for achieving the desired goal, is a task of adaptive signal processing. In the above example, the goal is the adaptive reconstruction of f from the samples $\{f(k)\}_{k\in\mathbb{Z}}$. Adaptive signal processing covers most of the practical important applications [12–14].

In contrast, the subsequence $\{N_n\}_{n\in\mathbb{N}}$ in question Q2 is universal in the sense that it does not depend on f. Obviously, a positive answer to question Q2 implies a positive answer to question Q1.

This brings us to the notion of strong divergence. We say that a sequence $\{a_n\}_{n\in\mathbb{N}} \subset \mathbb{R}$ diverges strongly if $\lim_{n\to\infty}|a_n| = \infty$. Clearly this is a stronger statement than $\limsup_{n\to\infty}|a_n| = \infty$, because in case of strong divergence we have $\lim_{n\to\infty}|a_{N_n}| = \infty$ for all subsequences $\{N_n\}_{n\in\mathbb{N}}$ of the natural numbers. So, if $P_N f$ diverges strongly for all $f \in \mathcal{PW}_\pi^1$, then question Q1 and consequently question Q2 have to be answered negatively.

In [15] it has been proved that there exists a function $f \in \mathcal{PW}_\pi^1$ such that $P_N f$ diverges strongly, i.e., that $\lim_{N\to\infty} P_N f = \infty$. Hence, neither question Q1 nor question Q2 can be answered in the affirmative for the Shannon sampling series.

In this paper we analyze whether adaptivity can improve the convergence behavior of sampling series. We study adaptivity in the subsequence $\{N_n\}_{n\in\mathbb{N}}$ and adaptivity in the kernel used for the reconstruction. We prove strong divergence, i.e., divergence for all subsequences, for different sampling series, where only weak divergence, i.e., divergence for certain subsequences, was known before. We further give the order of divergence. We also study the approximation of linear time-invariant (LTI) systems and show that we have

*This work was partly supported by the German Research Foundation (DFG) under grant BO 1734/20-1.

†U. Mönich was supported by the German Research Foundation (DFG) under grant MO 2572/1-1.

strong divergence there, even in the case of oversampling. Interestingly, it is possible to show strong divergence if the system is the Hilbert transform, which is a stable LTI system for $\mathcal{PW}_\pi^1$, i.e. the space under consideration.

2. NOTATION

Let $\hat{f}$ denote the Fourier transform of a function f, where $\hat{f}$ is to be understood in the distributional sense. By $L^p(\mathbb{R})$, $1 \le p \le \infty$, we denote the usual L^p-spaces, equipped with the norm $\|\cdot\|_p$. For $\sigma > 0$ let $\mathcal{B}_\sigma$ be the set of all entire functions f with the property that for all $\epsilon > 0$ there exists a constant $C(\epsilon)$ with $|f(z)| \le C(\epsilon)\exp\big((\sigma+\epsilon)|z|\big)$ for all $z \in \mathbb{C}$. The Bernstein space $\mathcal{B}_\sigma^p$ consists of all functions in $\mathcal{B}_\sigma$ whose restriction to the real line is in $L^p(\mathbb{R})$, $1 \le p \le \infty$. A function in $\mathcal{B}_\sigma^p$ is called bandlimited to σ. For $\sigma > 0$ and $1 \le p \le \infty$, we denote by $\mathcal{PW}_\sigma^p$ the Paley-Wiener space of functions f with a representation $f(z) = 1/(2\pi)\int_{-\sigma}^{\sigma} g(\omega)\,\mathrm{e}^{iz\omega}\ \mathrm{d}\omega$, $z \in \mathbb{C}$, for some $g \in L^p[-\sigma,\sigma]$. The norm for $\mathcal{PW}_\sigma^p$, $1 \le p < \infty$, is given by $\|f\|_{\mathcal{PW}_\sigma^p} = (1/(2\pi)\int_{-\sigma}^{\sigma}|\hat{f}(\omega)|^p\ \mathrm{d}\omega)^{1/p}$. Note that $\mathcal{PW}_\pi^1 \supset \mathcal{PW}_\pi^2 = \mathcal{B}_\pi^2$.

3. SYSTEM APPROXIMATION

A more general problem than the reconstruction problem, where the goal is to reconstruct a bandlimited signals f from its equidistant samples $\{f(k)\}_{k\in\mathbb{Z}}$, is the system approximation problem, where the goal is to approximate the output Tf of a stable LTI system T from the samples $\{f(k)\}_{k\in\mathbb{Z}}$ of the input signal f. This is the situation that is encountered in digital signal processing applications, where the interest is not in the reconstruction of a signal, but rather in the implementation of a system, i.e, the interest is in some transformation Tf of the sampled input signal f.

We briefly review some basic definitions and facts about stable linear time-invariant (LTI) systems. A linear system $T : \mathcal{PW}_\pi^p \to \mathcal{PW}_\pi^p$, $1 \le p \le \infty$, is called stable if the operator T is bounded, i.e., if $\|T\| := \sup_{\|f\|_{\mathcal{PW}_\pi^p}\le 1}\|Tf\|_{\mathcal{PW}_\pi^p} < \infty$. Furthermore, it is called time-invariant if $(Tf(\,\cdot\,-a))(t) = (Tf)(t-a)$ for all $f \in \mathcal{PW}_\pi^p$ and $t, a \in \mathbb{R}$. For every stable LTI system $T : \mathcal{PW}_\pi^1 \to \mathcal{PW}_\pi^1$, there exists exactly one function $\hat{h}_T \in L^\infty[-\pi,\pi]$ such that

$$(Tf)(t) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \hat{f}(\omega)\hat{h}_T(\omega)\,\mathrm{e}^{i\omega t}\ \mathrm{d}\omega, \quad t \in \mathbb{R},$$

for all $f \in \mathcal{PW}_\pi^1$ [16]. Conversely, every function $\hat{h}_T \in L^\infty[-\pi,\pi]$ defines a stable LTI system $T : \mathcal{PW}_\pi^1 \to \mathcal{PW}_\pi^1$. The operator norm of a stable LTI system T is given by $\|T\| = \|\hat{h}\|_{L^\infty[-\pi,\pi]}$. Note that $\hat{h}_T \in L^\infty[-\pi,\pi] \subset L^2[-\pi,\pi]$, and consequently $h_T \in \mathcal{PW}_\pi^2$.

Similar to the Shannon sampling series in the signal reconstruction problem, we can use the approximation process

$$\sum_{k=-\infty}^{\infty} f(k)h_T(t-k) \tag{2}$$

in the system approximation problem. In order to analyze the convergence behavior of (2), we introduce the abbreviation

$$(T_N f)(t) = \sum_{k=-N}^{N} f(k)h_T(t-k).$$

As already mentioned before, for certain signals in $f \in \mathcal{PW}_\pi^1$, the peak value of the reconstruction process $\|S_N f\|_\infty$ diverges strongly as N tends to infinity. However, in the case of oversampling, i.e., the case where the sampling rate is higher than Nyquist rate, the signal reconstruction process $S_N f$ converges globally uniformly [17]. This is a situation where oversampling helps improve the convergence behavior, consistent with engineering intuition. In contrast, the convergence behavior of the system approximation process (2) does not improve with oversampling [16]: for every $t \in \mathbb{R}$ and every $\sigma \in (0,\pi]$ there exist stable LTI systems T and signals $f \in \mathcal{PW}_\sigma^1$ such that

$$\limsup_{N\to\infty}|(Tf)(t) - (T_N f)(t)| = \infty.$$

In this paper we refine the questions Q1 and Q2 and analyze the following three questions:

1. Do we have the same strong divergence, which was proved in [15] for S_N, for the system approximation process $T_N f$?
2. Is it possible to obtain quantitative results about the divergence speed?
3. What happens in the case of oversampling?

4. CONJUGATED SHANNON AND SHANNON SERIES

In this section we analyze the behavior of conjugated Shannon sampling series and the Shannon sampling series. We first study the conjugated Shannon sampling series with critical sampling at Nyquist rate, i.e., the case without oversampling, and show that the answer to question Q1 is negative in this case. To this end, let $S_N f$ denote the finite Shannon sampling series as defined in (1), and

$$(H_N f)(t) := (HS_N f)(t) = \sum_{k=-N}^{N} f(k)\frac{1-\cos(\pi(t-k))}{\pi(t-k)} \tag{3}$$

the conjugated finite Shannon sampling series. H denotes the Hilbert transform which is defined as the principal value integral

$$(Hf)(t) = \frac{1}{\pi}\,\text{V.P.}\!\int_{-\infty}^{\infty}\frac{f(\tau)}{t-\tau}\,\mathrm{d}\tau = \frac{1}{\pi}\lim_{\epsilon\to 0}\int_{\epsilon\le|t-\tau|\le\frac{1}{\epsilon}}\frac{f(\tau)}{t-\tau}\,\mathrm{d}\tau.$$

The Hilbert transform is of enormous practical significance and plays a central role in the analysis of signal properties [18–22]. For further applications, see for example [23] and references therein.

It is well-known that $H_N f$ converges locally uniformly to Hf as N tends to infinity, that is, for $\tau > 0$ we have

$$\lim_{N\to\infty}\left(\max_{|t|\le\tau}|(Hf)(t) - (H_N f)(t)|\right) = 0.$$

The next theorem gives an answer about the global behavior of (3) and to the question Q1.

Theorem 1. *Let $\{\epsilon_N\}_{N\in\mathbb{N}}$ be an arbitrary sequence of positive numbers converging to zero. There exists a signal $f_1 \in \mathcal{PW}_\pi^1$ such that*

$$\lim_{N\to\infty}\frac{1}{\epsilon_N\log(N)}\left(\max_{t\in\mathbb{R}}\left(\sum_{k=-N}^{N} f_1(k)\frac{1-\cos(\pi(t-k))}{\pi(t-k)}\right)\right) = \infty. \tag{4}$$

Remark 1. In Theorem 1 we have divergence to ∞. If we replace the max-operator in (4) by the min-operator, the resulting expression converges to $-\infty$. The same is true for Theorems 2– 4.

Proof. Let $\{\epsilon_N\}_{N\in\mathbb{N}}$ be an arbitrary sequence of positive numbers converging to zero, and $\bar{\epsilon}_N = \max_{M\geq N}\epsilon_M$, $N\in\mathbb{N}$. Note that $\bar{\epsilon}_N \geq \epsilon_N$ for all $N\in\mathbb{N}$. Further, let $\{N_k\}_{k\in\mathbb{N}}$ be a strictly monotonically increasing sequence of natural numbers, such that $\bar{\epsilon}_{N_k} > \bar{\epsilon}_{N_{k+1}}$, $k\in\mathbb{N}$. We set $\delta_k = \sqrt{\bar{\epsilon}_{N_k}} - \sqrt{\bar{\epsilon}_{N_{k+1}}}$, $k\in\mathbb{N}$. It follows that $\delta_k > 0$ for all $k\in\mathbb{N}$ and that

$$\sum_{k=1}^{\infty}\delta_k = \sqrt{\bar{\epsilon}_{N_1}} < \infty. \tag{5}$$

For $N\in\mathbb{N}$ we define the functions

$$w_N(t) = \sum_{k=-\infty}^{\infty} w_N(k)\frac{\sin(\pi(t-k))}{\pi(t-k)}, \quad t\in\mathbb{R},$$

where $w_N(k)$ is given by

$$w_N(k) = \begin{cases} 1, & |k|\leq N, \\ 1-\frac{|k|-N}{N}, & N<|k|<2N, \\ 0, & |k|\geq 2N. \end{cases}$$

Note that we have $w_N\in\mathcal{PW}_\pi^1$ and $\|w_N\|_{\mathcal{PW}_\pi^1} < 3$ for all $N\in\mathbb{N}$ [24]. Based on w_N we define function

$$f_1 = \sum_{k=1}^{\infty}\delta_k w_{N_{k+1}}. \tag{6}$$

Since $\|\delta_k w_{N_{k+1}}\|_{\mathcal{PW}_\pi^1} < 3\delta_k$ and because of (5), it follows that the partial sums of the series in (6) form a Cauchy sequence in $\mathcal{PW}_\pi^1$, and thus the series in (6) converges in the $\mathcal{PW}_\pi^1$-norm and consequently uniformly on $\mathbb{R}$. Let $N\in\mathbb{N}$ be arbitrary but fixed. For $t_N = N+1$, it follows that

$$\sum_{l=-N}^{N} f_1(l)\frac{1-\cos(\pi(t_N-l))}{\pi(t_N-l)} = \sum_{l=-N}^{N} f_1(l)\frac{1-(-1)^{N+1-l}}{\pi(N+1-l)}. \tag{7}$$

There exists exactly one $\hat{k}\in\mathbb{N}$ such that $N\in[N_{\hat{k}}, N_{\hat{k}+1})$. We have

$$\sum_{l=-N}^{N} f_1(l)\frac{1-(-1)^{N+1-l}}{\pi(N+1-l)} \geq \sum_{k=\hat{k}}^{\infty}\frac{\delta_k}{\pi}\sum_{l=-N}^{N}\frac{1-(-1)^{N+1-l}}{N+1-l},$$

where we used that $w_{N_{k+1}}(l) = 1$ for all $k\geq\hat{k}$ and all $|l|\leq N$. Further, we have

$$\begin{aligned}\sum_{k=\hat{k}}^{\infty}\frac{\delta_k}{\pi}\sum_{l=-N}^{N}\frac{1-(-1)^{N+1-l}}{N+1-l} &= \frac{1}{\pi}\sum_{k=\hat{k}}^{\infty}\delta_k\sum_{l=0}^{N}\frac{2}{2l+1} \\ &\geq \frac{1}{\pi}\log(2N+3)\sqrt{\bar{\epsilon}_{N_{\hat{k}}}} \geq \frac{1}{\pi}\epsilon_N\log(N)\frac{1}{\sqrt{\epsilon_N}}\end{aligned} \tag{8}$$

because $N\geq N_{\hat{k}}$ and thus $\sqrt{\bar{\epsilon}_{N_{\hat{k}}}} \geq \sqrt{\epsilon_{N_{\hat{k}}}} \geq \sqrt{\epsilon_N}$. From (7)–(8), we see that

$$\sum_{l=-N}^{N} f_1(l)\frac{1-\cos(\pi(t_N-l))}{\pi(t_N-l)} \geq \frac{1}{\pi}\epsilon_N\log(N)\frac{1}{\sqrt{\epsilon_N}}$$

for all $N\in\mathbb{N}$, which in turn implies (4). $\square$

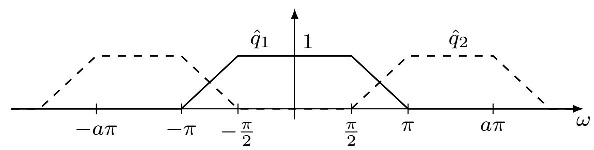

Fig. 1. Definition of $\hat{q}_1$ (solid line) and $\hat{q}_2$ (dashed line).

Next, we analyze the oversampling case for the conjugated Shannon sampling series, i.e., we treat question 3 from Section 3.

For the Shannon sampling series the convergence behavior in the case of oversampling is clear: we have global uniform convergence [17]. However, this is not true for the conjugated Shannon sampling series as the next theorem shows.

Theorem 2. *Let $\{\epsilon_N\}_{N\in\mathbb{N}}$ be an arbitrary sequence of positive numbers converging to zero. For every $\sigma\in(0,\pi]$ there exists a signal $f_\sigma\in\mathcal{PW}_\sigma^1$ such that*

$$\lim_{N\to\infty}\frac{1}{\epsilon_N\log(N)}\left(\max_{t\in\mathbb{R}}\left(\sum_{k=-N}^{N} f_\sigma(k)\frac{1-\cos(\pi(t-k))}{\pi(t-k)}\right)\right) = \infty.$$

Theorem 2 shows that in the case of oversampling, we have the same divergence behavior and speed that was observed in Theorem 1, i.e, the case without oversampling. That is, if we use oversampling as in Theorem 2, we have no improvement. Of course, due to oversampling, we have the freedom to use better, faster decaying kernels than those in Theorem 2. We will analyze this situation in Section 5.

Sketch of Proof. For the proof we use the signal f_1 from Theorem 1, which is defined in (6), and split it in the frequency domain into two parts:

$$\hat{f}_\sigma(\omega) = \begin{cases}\hat{f}_1(\omega), & |\omega|<\sigma, \\ 0, & \sigma\leq|\omega|\leq\pi\end{cases}$$

and

$$\hat{r}_\sigma(\omega) = \begin{cases}0, & |\omega|<\sigma, \\ \hat{f}_1(\omega), & \sigma\leq|\omega|\leq\pi.\end{cases}$$

For the rest of the proof we only present the idea and omit details[1]. Using the properties of the signal f_1, we see that r_σ is in $\mathcal{PW}_\pi^2$. Thus, the divergence is created by the difference signal f_σ. $\square$

Next, we come to the Shannon sampling series for the case of critical sampling at Nyquist rate. In [15] it has been proved that there exists a signal $f\in\mathcal{PW}_\pi^1$ such that $\|S_N f\|_\infty$ diverges strongly, i.e., that $\lim_{N\to\infty}\|S_N f\|_\infty = \infty$, and thus shown that the answer to question Q1 is negative. However, in [15] the authors also raised a question regarding the divergence order. Using the signal f_1 from the proof of Theorem 1, it is possible to answer this question.

Theorem 3. *Let $\{\epsilon_N\}_{N\in\mathbb{N}}$ be an arbitrary sequence of positive numbers converging to zero. There exists a signal $f_2\in\mathcal{PW}_\pi^1$ such that*

$$\lim_{N\to\infty}\frac{1}{\epsilon_N\log(N)}\left(\max_{t\in\mathbb{R}}\left(\sum_{k=-N}^{N} f_2(k)\frac{\sin(\pi(t-k))}{\pi(t-k)}\right)\right) = \infty.$$

Theorem 3 shows that for the Shannon sampling series it is possible to have strong divergence with order $\epsilon_N\log(N)$ for all zero sequences ϵ_N. This answers question 2 from Section 3.

[1] An extended version of this paper is in preparation, where we will include the full proofs of all theorems.

5. OVERSAMPLING AND KERNELS

We now come back to the situation where we know the signal f on an oversampling set. In Theorem 2 we already studied the oversampling case for the conjugated Shannon sampling series and observed that mere oversampling with the standard kernel does not remove the divergence. However, the redundance introduced by oversampling allows us to use other, faster decaying kernels. This introduces a further degree of freedom that can be employed for adaptivity. In addition to the subsequence $\{N_n\}_{n\in\mathbb{N}}$, we now can also choose the reconstruction kernel dependently on the signal f. Thus, question Q1 can be extended in the case of oversampling to also include the adaptive choice of the kernel. We will show in this section that for every arbitrary amount of oversampling the extended question Q1 has to be answered negatively. That is, even the joint optimization of the choice of the subsequence $\{N_n\}_{n\in\mathbb{N}}$ and the reconstruction kernel cannot circumvent the divergence.

We first consider the signal reconstruction problem. In the oversampling case, it is possible to create absolutely convergent sampling series by using other kernels than the sinc-kernel [4, 25, 26]. In particular, all kernels ϕ in the set $\mathcal{M}(a)$, which is defined next, can be used.

Definition 1. $\mathcal{M}(a)$, $a > 1$, is the set of functions $\phi \in \mathcal{B}^1_{a\pi}$ with $\hat{\phi}(\omega) = 1/a$ for $|\omega| \leq \pi$.

The functions in $\mathcal{M}(a)$, $a > 1$, are suitable kernels for the sampling series, because for all $f \in \mathcal{PW}^1_\pi$ and $a > 1$ we have

$$\lim_{N\to\infty} \max_{t\in\mathbb{R}} \left| f(t) - \sum_{k=-N}^{N} f\left(\frac{k}{a}\right) \phi\left(t - \frac{k}{a}\right)\right| = 0$$

if $\phi \in \mathcal{M}(a)$. We introduce the abbreviation

$$(H^a_{N,\phi} f)(t) = \sum_{k=-N}^{N} f\left(\frac{k}{a}\right) (H\phi)\left(t - \frac{k}{a}\right).$$

Theorem 4. *Let $\{\epsilon_N\}_{N\in\mathbb{N}}$ be an arbitrary sequence of positive numbers converging to zero. There exists a universal signal $f_1 \in \mathcal{PW}^1_\pi$ such that for all $a > 1$ and for all $\phi \in \mathcal{M}(a)$ we have*

$$\lim_{N\to\infty} \frac{1}{\epsilon_N \log(N)} \max_{t\in\mathbb{R}} (H^a_{N,\phi} f_1)(t) = \infty.$$

Theorem 4 shows that it is possible to have strong divergence with order $\epsilon_N \log(N)$ for all zero sequences ϵ_N even in the case of oversampling and the full adaptivity regarding the choice of ϕ.

Remark 2. We have the following result. Let $a > 1$ be arbitrary. For every $\phi \in \mathcal{M}(a)$ there exists a constant C_1 such that $\|H^a_{N,\phi} f\|_\infty \leq C_1 \log(N) \|f\|_{\mathcal{PW}^1_\pi}$ for all $N \geq 2$ and all $f \in \mathcal{PW}^1_\pi$. It follows that

$$\lim_{N\to\infty} \frac{\|H^a_{N,\phi} f\|_\infty}{\log(N)} = 0.$$

This shows how sharp the result in Theorem 4 is. Note that the same result is also true for Theorems 1–3.

The proof of Theorem 4 uses the following lemma from [27].

Lemma 1. *For all $a > 1$, $f \in \mathcal{PW}^1_\pi$, $N \in \mathbb{N}$ and $|t| \geq (N+1)/a$ we have*

$$\sum_{k=-N}^{N} \left| f\left(\frac{k}{a}\right) r\left(t - \frac{k}{a}\right)\right| < a^2 \|f\|_\infty,$$

where

$$r(t) = \frac{2}{\pi^2 t^2}\left(\sin(\pi t) - \sin\left(\frac{\pi}{2} t\right)\right). \tag{9}$$

Sketch of Proof of Theorem 4. Let $a > 1$ be arbitrary but fixed. Furthermore, let $\hat{q}_1$ and $\hat{q}_2$ be the functions defined in Figure 1 and $\phi \in \mathcal{M}(a)$ some arbitrary reconstruction kernel. Then we have $\phi = \phi * q_1 + \phi * q_2 = q_1 + \phi * q_2$ and $H\phi = Hq_1 + H(\phi * q_2) = Hq_1 + \phi * (Hq_2)$. Since $Hq_2 \in L^1(\mathbb{R})$, it follows that $s := \phi * (Hq_2) \in L^1(\mathbb{R})$. Moreover, for $N \in \mathbb{N}$ and $f \in \mathcal{PW}^1_\pi$ we have

$$\left|(H^a_{N,\phi} f)(t) - (H^a_{N,q_1} f)(t)\right| \leq C_2 \|f\|_\infty \|s\|_{\mathcal{B}^1_{a\pi}}, \tag{10}$$

where we used Nikol'skiĭ's inequality [8, p. 49]. For $\tau \neq 0$ we can simplify $(Hq_1)(\tau)$, using integration by parts, according to

$$\begin{aligned}(Hq_1)(\tau) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} -i\,\mathrm{sgn}(\omega)\hat{q}_1(\omega)\,\mathrm{e}^{i\omega\tau}\ \mathrm{d}\omega \\ &= \frac{1}{\pi}\int_0^{\pi} \sin(\omega\tau)\hat{q}_1(\omega)\ \mathrm{d}\omega = \frac{1}{\pi\tau} - r(\tau),\end{aligned}$$

where r is defined as in (9). For $|t| \geq (N+1)/a$ we thus obtain

$$(H^a_{N,q_1} f)(t) = \sum_{k=-N}^{N} f\left(\frac{k}{a}\right) \frac{1}{\pi\left(t - \frac{k}{a}\right)} - \sum_{k=-N}^{N} f\left(\frac{k}{a}\right) r\left(t - \frac{k}{a}\right),$$

and since $|\sum_{k=-N}^{N} f\left(\frac{k}{a}\right) r\left(t - \frac{k}{a}\right)| < a^2 \|f\|_\infty$ by Lemma 1, it follows that

$$(H^a_{N,q_1} f)(t) > \sum_{k=-N}^{N} f\left(\frac{k}{a}\right) \frac{1}{\pi\left(t - \frac{k}{a}\right)} - a^2 \|f\|_\infty. \tag{11}$$

Combining (10) and (11) we see that

$$\begin{aligned}(H^a_{N,\phi} f)(t) &\geq (H^a_{N,q_1} f)(t) - C_2 \|f\|_\infty \|s\|_{\mathcal{B}^1_{a\pi}} \\ &> \sum_{k=-N}^{N} f\left(\frac{k}{a}\right) \frac{1}{\pi\left(t - \frac{k}{a}\right)} - (a^2 + C_2 \|s\|_{\mathcal{B}^1_{a\pi}}) \|f\|_\infty\end{aligned}$$

for all $|t| \geq (N+1)/a$ and all $f \in \mathcal{PW}^1_\pi$. Hence, it suffices to concentrate the analysis on $\sum_{k=-N}^{N} f\left(\frac{k}{a}\right) \frac{1}{\pi(t-k/a)}$ in the following. The rest of the proof is similar to the proof of Theorem 1 and omitted due to space constraints. $\square$

6. RELATION TO PRIOR WORK

In the analysis of sampling series for signal reconstruction and system approximation, the focus was on proving weak divergence. However, weak divergence does not exclude the possibility that we have convergence for a certain subsequence and thus that adaptive approaches can be used. In contrast, strong divergence implies that this kind of adaptivity is useless. Although adaptive signal processing has been a very active area of research with numerous interesting and useful results [12–14], the general question of strong divergence has not been addressed. It seems that this is a new direction.

Only recently, the concept of strong divergence gained attention in the analysis of sampling series. In [15] a first result was given, by proving the strong divergence of the peak value of the Shannon sampling series. Here, we consider the adaptive system approximation with oversampling and adaptive choice of the reconstruction kernel, and show strong divergence for the analyzed series. Further, by providing the explicit divergence speed in our theorems, we answer a question about the divergence speed that was raised in [15]. In the case of strong divergence it is interesting to know the size of the set of signals for which it occurs. This question, which was posed in [11], will be answered in [28]. Currently, strong divergence has to be proved from case to case with different tools, because a general theory is still missing.

7. REFERENCES

[1] Claude E. Shannon, "Communication in the presence of noise," in *Proceedings of the IRE*, Jan. 1949, vol. 37, pp. 10–21.

[2] Abdul J. Jerri, "The Shannon sampling theorem–its various extensions and applications: A tutorial review," *Proc. IEEE*, vol. 65, no. 11, pp. 1565–1596, Nov. 1977.

[3] John R. Higgins, "Five short stories about the cardinal series," *Bull. Amer. Math. Soc.*, vol. 12, no. 1, pp. 45–89, 1985.

[4] P. L. Butzer, W. Splettstößer, and R. L. Stens, "The sampling theorem and linear prediction in signal analysis," *Jahresbericht der Deutschen Mathematiker-Vereinigung*, vol. 90, no. 1, pp. 1–70, Jan. 1988.

[5] Farokh Marvasti, Ed., *Nonuniform Sampling: Theory and Practice*, Kluwer Academic / Plenum Publishers, 2001.

[6] J. L. Brown, Jr., "On the error in reconstructing a non-bandlimited function by means of the bandpass sampling theorem," *Journal of Mathematical Analysis and Applications*, vol. 18, pp. 75–84, 1967, Erratum, ibid, vol. 21, 1968, p. 699.

[7] P. L. Butzer and R. L. Stens, "Sampling theory for not necessarily band-limited functions: A historical overview," *SIAM Review*, vol. 34, no. 1, pp. 40–53, Mar. 1992.

[8] John R. Higgins, *Sampling Theory in Fourier and Signal Analysis – Foundations*, Oxford University Press, 1996.

[9] P. L. Butzer, P. J. S. G. Ferreira, J. R. Higgins, S. Saitoh, G. Schmeisser, and R. L. Stens, "Interpolation and sampling: E.T. Whittaker, K. Ogura and their followers," *Journal of Fourier Analysis and Applications*, vol. 17, no. 2, pp. 320–354, 2011.

[10] Holger Boche and Ullrich J. Mönich, "Sampling of deterministic signals and systems," *IEEE Trans. Signal Process.*, vol. 59, no. 5, pp. 2101–2111, May 2011.

[11] Holger Boche and Volker Pohl, *Sampling Theory, a Renaissance*, chapter System Approximations and Generalized Measurements in Modern Sampling Theory, Applied and Numerical Harmonic Analysis. Birkhauser (Springer-Verlag), 2014, to be published, preprint available http://arxiv.org/abs/1410.5872.

[12] Bernard Widrow and Samuel D. Stearns, *Adaptive Signal Processing*, Prentice Hall, 1985.

[13] S. Thomas Alexander, *Adaptive Signal Processing: Theory and Applications*, Springer New York, 1986.

[14] Tülay Adali and Simon Haykin, Eds., *Adaptive Signal Processing: Next Generation Solutions*, John Wiley & Sons, Inc., 2010.

[15] Holger Boche and Brendan Farrell, "Strong divergence of reconstruction procedures for the Paley-Wiener space PW_π^1 and the Hardy space H^1," *Journal of Approximation Theory*, vol. 183, pp. 98–117, July 2014.

[16] Holger Boche and Ullrich J. Mönich, "Sampling-type representations of signals and systems," *Sampling Theory in Signal and Image Processing*, vol. 9, no. 1–3, pp. 119–153, Jan., May, Sept. 2010.

[17] Holger Boche and Ullrich J. Mönich, "Global and local approximation behavior of reconstruction processes for Paley-Wiener functions," *Sampling Theory in Signal and Image Processing*, vol. 8, no. 1, pp. 23–51, Jan. 2009.

[18] L. M. Fink, "Relations between the spectrum and instantaneous frequency of a signal," *Problems of Information Transmission*, vol. 2, no. 4, pp. 11–21, 1966, translation.

[19] D. Ye. Vakman, "On the definition of concepts of amplitude, phase and instantaneous frequency of a signal," *Radio Engineering and Electronic Physics*, vol. 17, no. 5, pp. 754–759, 1972, translation.

[20] Dennis Gabor, "Theory of communication," *Journal of the Institute of Electrical Engineers*, vol. 93, no. 3, pp. 429–457, Nov. 1946.

[21] Norden E. Huang, Zheng Shen, and Steven R. Long, "A new view of nonlinear water waves: The Hilbert spectrum," *Annual Review of Fluid Mechanics*, vol. 31, no. 1, pp. 417–457, Jan. 1999.

[22] Norden E. Huang, Man-Li C. Wu, Steven R. Long, Samuel S. P. Shen, Wendong Qu, Per Gloersen, and Kuang L. Fan, "A confidence limit for the empirical mode decomposition and Hilbert spectral analysis," *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, vol. 459, no. 2037, pp. 2317–2345, Sept. 2003.

[23] Volker Pohl and Holger Boche, *Advanced Topics in System and Signal Theory: A Mathematical Approach*, vol. 4 of *Foundations in Signal Processing, Communications and Networking*, Springer, 2009.

[24] Holger Boche and Ullrich J. Mönich, "Behavior of the quantization operator for bandlimited, nonoversampled signals," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2433–2440, May 2010.

[25] P. L. Butzer, W. Engels, S. Ries, and R. L. Stens, "The Shannon sampling series and the reconstruction of signals in terms of linear, quadratic and cubic splines," *SIAM Journal on Applied Mathematics*, vol. 46, no. 2, pp. 299–323, Apr. 1986.

[26] W. Engels, E. L. Stark, and L. Vogt, "Optimal kernels for a general sampling theorem," *Journal of Approximation Theory*, vol. 50, no. 1, pp. 69–83, May 1987.

[27] Holger Boche and Ullrich J. Mönich, "An impossibility result for linear signal processing under thresholding," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1082–1094, Mar. 2010.

[28] Holger Boche and Ullrich J. Mönich, "Strong divergence of approximation processes in Banach spaces," *in preparation*, 2015.