

Orthogonal Matching Pursuit based Reconstruction for Modulo Hysteresis Operators

Matthias Beckmann and Jürgen Jeschke

Center for Industrial Mathematics, University of Bremen, Germany

research@mbeckmann.de • jjeschke@uni-bremen.de

Abstract—Unlimited sampling provides an acquisition scheme for high dynamic range signals by folding the signal into the dynamic range of the analog-to-digital converter (ADC) using modulo non-linearity prior to sampling to prevent saturation. Recently, a generalized scheme called modulo hysteresis was introduced to account for hardware non-idealities. The encoding operator, however, does not guarantee that the output signal is within the dynamic range of the ADC. To resolve this, we propose a modified modulo hysteresis operator and show identifiability of bandlimited signals from modulo hysteresis samples. We propose a recovery algorithm based on orthogonal matching pursuit and validate our theoretical results through numerical experiments.

Index Terms—Unlimited sampling, generalized modulo operator, Shannon sampling theory, orthogonal matching pursuit.

I. INTRODUCTION

The well-known Shannon sampling theorem [1] allows for the recovery of a bandlimited function g from discrete samples $\{g(nT) \mid n \in \mathbb{Z}\}$ with sufficiently small sampling rate $T > 0$, namely if sampled at or above the Nyquist rate, which is implemented in hardware by analog-to-digital converters (ADCs). The reconstruction formula, however, assumes access to perfect samples and any distortion in the data leads to artifacts in the reconstruction. One particular bottleneck in practice is that ADCs operate at a fixed dynamic range and an input signal exceeding the threshold leads to saturation or clipping and, hence, we suffer a permanent information loss.

To overcome these limitations, the Unlimited Sampling Framework (USF) was introduced in [2], [3], [4] along with mathematically backed reconstruction algorithms, where, prior to sampling, the input function is folded into the dynamic range using modulo arithmetic. Since then several reconstruction approaches have been proposed, e.g. [5], [6], [7], [8], and USF has been extended to various applications like imaging [9], computerized tomography [10], [11] and radar [12]. More general encoding operators were considered in [13].

The above approaches assume an instantaneous fold whenever the input signal g reaches the dynamic threshold, yielding an encoding that operates pointwise, i.e., the output at time t is determined by $g(t)$. As opposed to this, in [14], [15] a generalized modulo encoder is proposed that accounts for non-instantaneous folds due to hardware imperfections. This yields an encoding operator \mathcal{M}_H with memory in the sense that $\mathcal{M}_H g(t)$ not only depends on $g(t)$, but on $g(\tau)$ for $\tau \leq t$. Because of this, \mathcal{M}_H is also called *modulo hysteresis operator*.

This work was supported by the Deutsche Forschungsgemeinschaft (DFG), project number 530863002.

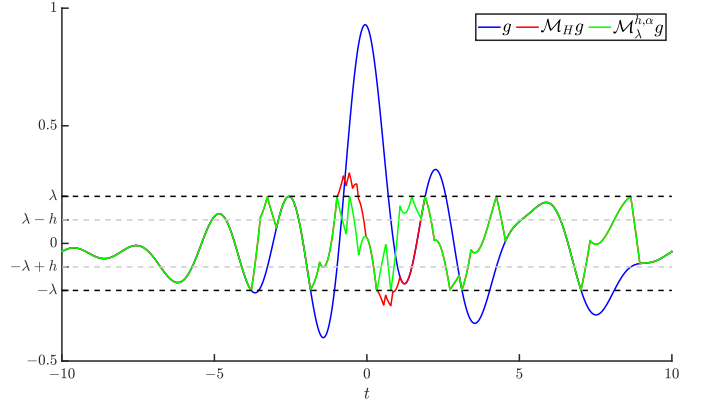


Fig. 1. Illustration of generalized modulo encoder \mathcal{M}_H and modified modulo hysteresis operator $\mathcal{M}_\lambda^{h,\alpha}$ for random input $g \in \text{PW}_\Omega$ with $\lambda = 0.2$, $h = 0.1$ and $\alpha = 0.3$, where $\mathcal{M}_\lambda^{h,\alpha} g \in [-\lambda, \lambda]$, but $|\mathcal{M}_H g(t)| > \lambda$ for some t .

The model proposed in [14], [15], however, does not guarantee that the output is within the dynamic range of the ADC, see Fig. 1 for illustration. Hence, in this work, we propose a modified modulo hysteresis operator $\mathcal{M}_\lambda^{h,\alpha}$ that stays within the given dynamic range. In Section II, we show that the input signal g is uniquely determined by the discrete samples $\{\mathcal{M}_\lambda^{h,\alpha} g(nT) \mid n \in \mathbb{Z}\}$ with sufficiently small $T > 0$ and, in Section III, explain how g can be recovered using orthogonal matching pursuit (OMP). Finally, our theoretical results are supported by numerical experiments in Section IV.

II. MODULO HYSTERESIS OPERATORS

For a function $g : \mathbb{R} \rightarrow \mathbb{R}$, the ideal modulo encoder \mathcal{M}_λ with modulo threshold $\lambda > 0$ is pointwisely defined as in [2] via

$$\mathcal{M}_\lambda g(t) = g(t) - 2\lambda \left\lfloor \frac{g(t) + \lambda}{2\lambda} \right\rfloor \quad \text{for } t \in \mathbb{R},$$

where $\lfloor \cdot \rfloor$ denotes the floor function for real numbers, and satisfies $\mathcal{M}_\lambda g(t) \in [-\lambda, \lambda]$ for all $t \in \mathbb{R}$. The typical input space is the Paley-Wiener space PW_Ω of square-integrable bandlimited functions with bandwidth $\Omega > 0$. In this case, [3] shows that the input function can be recovered from output samples if the sampling rate satisfies the oversampling condition $T < \frac{1}{2\Omega e}$, where the reconstruction is based on iterated forward differences in time domain. A first hardware validation of a modulo encoder is reported in [4] together with a Fourier domain recovery approach.

To account for non-ideal hardware implementations, in [14] a generalized modulo encoder \mathcal{M}_H with parameter triplet $H = (\lambda, h, \alpha)$ is introduced, where the *hysteresis* parameter $h \geq 0$ models an imperfect alignment of the reset threshold and the post-reset value and the *transient* parameter $\alpha \geq 0$ models an inexact folding transition. For a function g , the output $\mathcal{M}_H g$ is defined via a sequence of folding points $(\tau_i)_{i \in \mathbb{N}}$, given by

$$\begin{aligned}\tau_1 &= \min\{t > \tau_0 \mid \mathcal{M}_\lambda(g(t) + \lambda) = 0\}, \\ \tau_{i+1} &= \min\{t > \tau_i \mid \mathcal{M}_\lambda(g(t) - g(\tau_i) + h s_i) = 0\}\end{aligned}$$

with $s_i = \text{sgn}(g(\tau_i) - g(\tau_{i-1}))$ and $\tau_0 \in \mathbb{R}$. Then, for $t \in \mathbb{R}$,

$$\mathcal{M}_H g(t) = g(t) - (2\lambda - h) \sum_{i \in \mathbb{N}} s_i \varepsilon_\alpha(t - \tau_i),$$

where $\varepsilon_0(t) = \mathbb{1}_{[0, \infty)}(t)$ models an instantaneous folding transition and, for $\alpha > 0$, $\varepsilon_\alpha(t) = \frac{t}{\alpha} \mathbb{1}_{[0, \alpha)}(t) + \mathbb{1}_{[\alpha, \infty)}(t)$ models the folding transition as a straight line with slope $1/\alpha$.

Note that the output $\mathcal{M}_H g(t)$ depends on all folding points $\tau_j < t$ and, hence, \mathcal{M}_H is not acting pointwisely. Moreover, the above definition does not guarantee that $\mathcal{M}_\lambda g(t) \in [-\lambda, \lambda]$ for $t \geq \tau_0$, as illustrated in Fig. 1. One reason for this is that the folding points $(\tau_i)_{i \in \mathbb{N}}$ are independent of the transient α . To account for this, we now wish to propose a modified definition of the modulo hysteresis encoder. To this end, we rewrite $\mathcal{M}_H g$ by defining the function sequence $(\eta_n^{(\alpha)})_{n \in \mathbb{N}_0}$ via

$$\eta_{n+1}^{(\alpha)} = \eta_n^{(\alpha)} - (2\lambda - h) \text{sgn}(\eta_n^{(\alpha)}(\tau_{n+1})) \varepsilon_\alpha(\cdot - \tau_{n+1}),$$

where $\eta_0^{(\alpha)} = g$. With this, we then obtain

$$\tau_i = \inf\{t > \tau_{i-1} \mid |\eta_{i-1}^{(0)}(t)| \geq \lambda\}$$

and

$$\mathcal{M}_H g(t) = \lim_{n \rightarrow \infty} \eta_n^{(\alpha)}(t).$$

The idea is now to make τ_i dependent on $\eta_{i-1}^{(\alpha)}$ instead of $\eta_{i-1}^{(0)}$.

Definition 1 (Modified Modulo Hysteresis). *Let $\lambda > 0$, $h \geq 0$ and $\alpha \geq 0$. For a function $g : \mathbb{R} \rightarrow \mathbb{R}$ and starting point $\tau_0 \in \mathbb{R}$, we define the two sequences $(\kappa_n)_{n \in \mathbb{N}_0}$ and $(\zeta_n)_{n \in \mathbb{N}_0}$ via*

$$\kappa_{n+1} = \inf\{t > \kappa_n \mid |\zeta_n(t)| \geq \lambda\}$$

and

$$\zeta_{n+1} = \zeta_n - (2\lambda - h) \text{sgn}(\zeta_n(\kappa_{n+1})) \varepsilon_\alpha(\cdot - \kappa_{n+1}),$$

where we set $\kappa_0 = \tau_0$ and $\zeta_0 = g$. With this, we define the modified modulo hysteresis operator $\mathcal{M}_\lambda^{h, \alpha}$ as

$$\mathcal{M}_\lambda^{h, \alpha} g(t) = \lim_{n \rightarrow \infty} \zeta_n(t) \quad \text{for } t \in \mathbb{R}.$$

When considering the input function space

$$\mathcal{C}_{\lambda, \tau_0}^{0,1} = \{g \in \mathcal{C}^{0,1}(\mathbb{R}) \mid |g(t)| < \lambda \ \forall |t| \geq |\tau_0|\},$$

where $\mathcal{C}^{0,1}(\mathbb{R})$ denotes the space of Lipschitz continuous functions on \mathbb{R} , which are differentiable almost everywhere with essentially bounded derivative, one can show that $\mathcal{M}_\lambda^{h, \alpha} g$ is well-defined for all $g \in \mathcal{C}_{\lambda, \tau_0}^{0,1}$ and, as opposed to $\mathcal{M}_H g$, it

satisfies $\mathcal{M}_\lambda^{h, \alpha} g(t) \in [-\lambda, \lambda]$ for all $t \in \mathbb{R}$ and

$$\kappa_0 < \kappa_n < \infty \implies |\mathcal{M}_\lambda^{h, \alpha} g(\kappa_n)| = \begin{cases} \lambda & \text{if } \alpha > 0, \\ \lambda - h & \text{if } \alpha = 0, \end{cases}$$

see also Fig. 1 for illustration. Due to space limitations, we omit the proof in this paper and instead now focus on the identifiability from discrete modulo hysteresis samples $\{\mathcal{M}_\lambda^{h, \alpha} g(kT) \mid k \in \mathbb{Z}\}$ with sampling rate $T > 0$. To this end, we consider the space

$$\mathcal{C}_{\lambda, \tau_0}^{1,1} = \{g \in \mathcal{C}_{\lambda, \tau_0}^{0,1} \mid g' \in \mathcal{C}^{0,1}(\mathbb{R})\}$$

and state a separation property of the folding points $(\kappa_n)_{n \in \mathbb{N}_0}$.

Lemma 1. *Let $\alpha > 0$ and $g \in \mathcal{C}_{\lambda, \tau_0}^{1,1}$ with $\|g''\|_\infty \leq \frac{2h}{\alpha^2}$. Then, for $n \in \mathbb{N}_0$ we have*

$$\mathcal{M}_\lambda^{h, \alpha} g(\kappa_n) \cdot \mathcal{M}_\lambda^{h, \alpha} g(\kappa_{n+1}) < 0 \implies \kappa_{n+1} - \kappa_n \geq \alpha.$$

Proof. Consider the subsequence $(\kappa_{m_n})_{n \in \mathbb{N}}$, where $\kappa_{m_1} = \kappa_1$ and $\kappa_{m_{n+1}}$ is the next folding point with different sign, i.e.,

$$\kappa_{m_{n+1}} = \min_{k > m_n} \{\kappa_k \mid \text{sgn}(\zeta_k(\kappa_k)) \neq \text{sgn}(\zeta_{m_n}(\kappa_{m_n}))\}$$

so that $\mathcal{M}_\lambda^{h, \alpha} g(\kappa_{m_n}) = \pm \lambda = -\mathcal{M}_\lambda^{h, \alpha} g(\kappa_{m_{n+1}})$. Now, without loss of generality let $\mathcal{M}_\lambda^{h, \alpha} g(\kappa_{m_1}) = \lambda$. Then, we obtain $\mathcal{M}_\lambda^{h, \alpha} g(\kappa_{m_2-1}) = \lambda$ and $\mathcal{M}_\lambda^{h, \alpha} g(\kappa_{m_2}) = -\lambda$. Additionally, $g'(\kappa_{m_2-1}) \geq p \frac{2\lambda - h}{\alpha}$ with $p = |\{\kappa_k \in (\kappa_{m_2-1} - \alpha, \kappa_{m_2-1})\}|$. Therefore, we can estimate $\zeta_{m_2-1}(t)$ for $t > \kappa_{m_2-1}$ as

$$\begin{aligned}\zeta_{m_2-1}(t) &\geq \zeta_{m_2-1}(\kappa_{m_2-1}) + g'(\kappa_{m_2-1})(t - \kappa_{m_2-1}) \\ &\quad - \frac{\|g''\|_\infty}{2} (t - \kappa_{m_2-1})^2 - (p+1) \frac{2\lambda - h}{\alpha} (t - \kappa_{m_2-1}) \\ &\geq \lambda - \frac{h}{\alpha^2} (t - \kappa_{m_2-1})^2 - \frac{2\lambda - h}{\alpha} (t - \kappa_{m_2-1}).\end{aligned}$$

With this we can conclude that

$$\begin{aligned}\kappa_{m_2} &\geq \inf\left\{t > \kappa_{m_2-1} \mid \lambda - \frac{h}{\alpha^2} (t - \kappa_{m_2-1})^2 \right. \\ &\quad \left. - \frac{2\lambda - h}{\alpha} (t - \kappa_{m_2-1}) = -\lambda\right\} \\ &= \kappa_{m_2-1} + \alpha.\end{aligned}$$

We now assume that we have $\kappa_{m_n} \geq \kappa_{m_{n-1}} + \alpha$ for some $n \geq 2$. Again without loss of generality let $\mathcal{M}_\lambda^{h, \alpha} g(\kappa_{m_n}) = \lambda$. Since $\kappa_{m_{n+1}-1} \geq \kappa_{m_n} > \kappa_{m_n-1} + \alpha$ we can proceed with the same arguments as before to obtain

$$\begin{aligned}\kappa_{m_{n+1}} &\geq \inf\left\{t > \kappa_{m_{n+1}-1} \mid \lambda - \frac{h}{\alpha^2} (t - \kappa_{m_{n+1}-1})^2 \right. \\ &\quad \left. - \frac{2\lambda - h}{\alpha} (t - \kappa_{m_{n+1}-1}) = -\lambda\right\} \\ &= \kappa_{m_{n+1}-1} + \alpha,\end{aligned}$$

which completes the proof. \square

With the above condition on the second derivative on g we can characterize the behaviour of $\mathcal{M}_\lambda^{h, \alpha} g(t)$ for large $t > |\tau_0|$.

Lemma 2. *If $g \in \mathcal{C}_{\lambda-h, \tau_0}^{1,1}$ with $\|g''\|_\infty \leq \frac{2h}{\alpha^2}$ and $0 \leq h < \lambda$, we have $\mathcal{M}_\lambda^{h, \alpha} g(t) = g(t)$ for all $t \geq |\tau_0| + 2\alpha$.*

Algorithm 1 (OMP)

Input: Signal $\mathbf{s} \in \mathbb{C}^M$ and dictionary matrix $\mathbf{V} \in \mathbb{C}^{M \times N}$, error tolerance $\varepsilon > 0$

- 1: $\mathbf{c}^{(0)} = 0 \in \mathbb{C}^N$, $\mathcal{S}^{(0)} = \emptyset$, $i = 1$
- 2: **while** $\|\mathbf{V}^*(\mathbf{s} - \mathbf{V}\mathbf{c}^{(i-1)})\|_\infty > \varepsilon$ **do**
- 3: $j^{(i)} = \arg \max_{1 \leq j \leq N} |[\mathbf{V}^*(\mathbf{s} - \mathbf{V}\mathbf{c}^{(i-1)})]_j|$
- 4: $\mathcal{S}^{(i)} = \mathcal{S}^{(i-1)} \cup \{j^{(i)}\}$
- 5: $\mathbf{c}^{(i)} = \arg \min_{\mathbf{c}} \{\|\mathbf{s} - \mathbf{V}\mathbf{c}\|_2 \mid \text{supp}(\mathbf{c}) \subseteq \mathcal{S}^{(i)}\}$
- 6: $i = i + 1$
- 7: **end while**

Output: $\mathbf{c}^{(i_{\text{end}})} \in \mathbb{C}^N$

Proof. Let κ_n be the largest folding point smaller than or equal to $|\tau_0|$. We now go through the different possible cases.

Case 1: If $\kappa_n + \alpha \leq |\tau_0|$, there exists $k \in \mathbb{Z}$ such that

$$\begin{aligned} |\mathcal{M}_\lambda^{h,\alpha} g(|\tau_0|)| &= |k(2\lambda - h) + g(|\tau_0|)| \\ &\geq |k|(2\lambda - h) - |g(|\tau_0|)| \\ &> |k|(2\lambda - h) - \lambda + h. \end{aligned}$$

As $|\mathcal{M}_\lambda^{h,\alpha} g(|\tau_0|)| \leq \lambda$, we get $k = 0$, which implies that $\mathcal{M}_\lambda^{h,\alpha} g(|\tau_0|) = g(|\tau_0|)$ and, thus, $\zeta_n(t) = g(t)$ for all $t \geq |\tau_0|$. Inserting this into the definition of κ_{n+1} gives

$$\kappa_{n+1} = \inf\{t > |\tau_0| \mid |g(t)| \geq \lambda\} = \infty,$$

from which follows that $\mathcal{M}_\lambda^{h,\alpha} g(t) = g(t)$ for all $t \geq |\tau_0|$. As $\kappa_n \leq |\tau_0|$, this is the only case for $\alpha = 0$. So now let $\alpha > 0$.

Case 2: If $\kappa_n + \alpha > |\tau_0|$, but $\kappa_n + \alpha \leq \kappa_{n+1}$, we get, for some $k \in \mathbb{Z}$,

$$\begin{aligned} |\mathcal{M}_\lambda^{h,\alpha} g(\kappa_n + \alpha)| &= |k(2\lambda - h) + g(\kappa_n + \alpha)| \\ &> |k|(2\lambda - h) - \lambda + h \end{aligned}$$

so that $\kappa_{n+1} = \infty$ and $\mathcal{M}_\lambda^{h,\alpha} g(t) = g(t)$ for all $t \geq \kappa_n + \alpha$. Since $\kappa_n \leq |\tau_0|$, this also holds for all $t \geq |\tau_0| + \alpha$.

Case 3: If $\kappa_n + \alpha > \max\{|\tau_0|, \kappa_{n+1}\}$, but $\kappa_{n+1} + \alpha \leq \kappa_{n+2}$, we get, for some $k \in \mathbb{Z}$,

$$\begin{aligned} |\mathcal{M}_\lambda^{h,\alpha} g(\kappa_{n+1} + \alpha)| &= |k(2\lambda - h) + g(\kappa_{n+1} + \alpha)| \\ &> |k|(2\lambda - h) - \lambda + h \end{aligned}$$

so that $\kappa_{n+2} = \infty$ and $\mathcal{M}_\lambda^{h,\alpha} g(t) = g(t)$ for all $t \geq \kappa_{n+1} + \alpha$. Since $\kappa_{n+1} \leq |\tau_0| + \alpha$, this also holds for all $t \geq |\tau_0| + 2\alpha$.

Case 4: Let $\kappa_n + \alpha > \max\{|\tau_0|, \kappa_{n+1}\}$, $\kappa_{n+1} + \alpha > \kappa_{n+2}$. As above, without loss of generality let $\mathcal{M}_\lambda^{h,\alpha} g(\kappa_n) = \lambda$. Lemma 1 implies that $\lambda = \mathcal{M}_\lambda^{h,\alpha} g(\kappa_{n+1}) = \mathcal{M}_\lambda^{h,\alpha} g(\kappa_{n+2})$ and $\mathcal{M}_\lambda^{h,\alpha} g(\kappa_{n+2} + \alpha) \leq \zeta_{n+2}(\kappa_{n+2} + \alpha)$. This gives us that

$$\begin{aligned} \mathcal{M}_\lambda^{h,\alpha} g(\kappa_{n+2} + \alpha) &\leq \lambda + \zeta_{n+2}(\kappa_{n+2} + \alpha) - \zeta_{n+2}(\kappa_n) \\ &\leq \lambda + g(\kappa_{n+2} + \alpha) - g(\kappa_n) - 3(2\lambda - h) \\ &< \lambda + 2(\lambda - h) - 3(2\lambda - h) = -3\lambda + h < -2\lambda \end{aligned}$$

in contradiction to $\mathcal{M}_\lambda^{h,\alpha} g(t) \in [-\lambda, \lambda]$. \square

Algorithm 2 (SAOMP)

Input: Signal $\mathbf{s} \in \mathbb{C}^M$ and dictionary matrix $\mathbf{V} \in \mathbb{C}^{M \times N}$, error tolerance $\varepsilon > 0$, initial threshold $\nu \in [0, 1]$, pruning threshold $\mu \in [0, 1]$, maximal iteration number $i_{\text{max}} \in \mathbb{N}$

- 1: $\mathbf{c}^{(0)} = 0 \in \mathbb{C}^N$, $\mathcal{S}^{(0)} = \emptyset$, $\delta = \nu$, $i = 1$
- 2: **while** $i \leq i_{\text{max}}$ **and** $\|\mathbf{V}^*(\mathbf{s} - \mathbf{V}\mathbf{c}^{(i-1)})\|_\infty > \varepsilon$ **do**
- 3: $\mathbf{r}^{(i)} = \mathbf{V}^*(\mathbf{s} - \mathbf{V}\mathbf{c}^{(i-1)})$
- 4: $\mathcal{S}^{(i)} = \mathcal{S}^{(i-1)} \cup \{j \in \{1, \dots, N\} \mid |\mathbf{r}_j^{(i)}| \geq \delta \|\mathbf{r}^{(i)}\|_\infty\}$
- 5: $\mathbf{c}^{(i)} = \arg \min_{\mathbf{c}} \{\|\mathbf{s} - \mathbf{V}\mathbf{c}\|_2 \mid \text{supp}(\mathbf{c}) \subseteq \mathcal{S}^{(i)}\}$
- 6: $\mathcal{S}^{(i)} = \{j \in \{1, \dots, N\} \mid |\mathbf{c}_j^{(i)}| \geq \mu \|\mathbf{c}^{(i)}\|_\infty\}$
- 7: **for** $j \notin \mathcal{S}^{(i)}$: $\mathbf{c}_j^{(i)} = 0$
- 8: $\delta = \delta + \frac{1-\nu}{i_{\text{max}}}$, $i = i + 1$
- 9: **end while**

Output: $\mathbf{c}^{(i_{\text{end}})} \in \mathbb{C}^N$

We can now derive a condition for the sampling rate T to uniquely determine $g \in \text{PW}_\Omega$ with bandwidth $\Omega > 0$ from modulo hysteresis samples. In [3, Lemma 1], it is shown that any function $g \in \text{PW}_\Omega$ is uniquely determined by the samples $\{g(kT_\varepsilon) \mid k \in \mathbb{Z} \setminus J\}$ if $0 < T_\varepsilon \leq \frac{\pi}{\Omega + \varepsilon}$ for arbitrarily small $\varepsilon > 0$ and any finite set $J \subset \mathbb{Z}$. We can apply this to show that if $T < \frac{\pi}{\Omega}$, g is also uniquely determined by the samples $\{\mathcal{A}g(kT) \mid k \in \mathbb{Z}\}$ for a general operator $\mathcal{A} : \text{PW}_\Omega \rightarrow L^2(\mathbb{R})$ with

$$\mathcal{A}g(t) = g(t) \quad \forall t \in \mathbb{R} \setminus M$$

for some bounded set $M \subset \mathbb{R}$.

Lemma 3. Any $g \in \text{PW}_\Omega$ is determined by $\{\mathcal{A}g(kT) \mid k \in \mathbb{Z}\}$ if the oversampling condition $T < \frac{\pi}{\Omega}$ is met.

Proof. Since $T < \frac{\pi}{\Omega}$, there exists $\varepsilon > 0$ so that $T \leq \frac{\pi}{\Omega + \varepsilon}$. Now define $J = M \cap \mathbb{Z}$, which is finite as $M \subset \mathbb{R}$ is bounded. Since $\{\mathcal{A}g(kT) \mid k \in \mathbb{Z} \setminus J\} = \{g(kT) \mid k \in \mathbb{Z} \setminus J\}$ by assumption on \mathcal{A} , the statement follows from [3, Lemma 1]. \square

Since the modulo hysteresis operator $\mathcal{M}_\lambda^{h,\alpha}$ satisfies the above conditions on \mathcal{A} if we restrict ourselves to functions $g \in \text{PW}_\Omega \cap \mathcal{C}_{\lambda-h, \tau_0}^{1,1}$ with $\|g''\|_\infty \leq \frac{2h}{\alpha^2}$ and $0 \leq h < \lambda$ due to Lemma 2, we get the following identifiability result for $\mathcal{M}_\lambda^{h,\alpha}$.

Corollary 1. Let $0 \leq h < \lambda$. Then, any $g \in \text{PW}_\Omega \cap \mathcal{C}_{\lambda-h, \tau_0}^{1,1}$ with $\|g''\|_\infty \leq \frac{2h}{\alpha^2}$ is uniquely determined by the modulo hysteresis samples $\{\mathcal{M}_\lambda^{h,\alpha} g(kT) \mid k \in \mathbb{Z}\}$ if $T < \frac{\pi}{\Omega}$. \square

III. RECONSTRUCTION VIA OMP

For our reconstruction approach we assume that we are given $2K + 1$ modulo hysteresis samples of $g \in \text{PW}_\Omega \cap \mathcal{C}_{\lambda, \tau_0}^{0,1}$,

$$\{\mathcal{M}_\lambda^{h,\alpha} g(kT) \mid k = -K, \dots, K\}$$

with sampling rate $0 < T < \frac{\pi}{\Omega}$, where $K \in \mathbb{N}$ is large enough such that $|g(t)| < \lambda$ for all $|t| \geq KT$. Let us first consider the folding function $\varepsilon_\alpha(\cdot - \kappa)$ located at fixed $-KT < \kappa < KT$

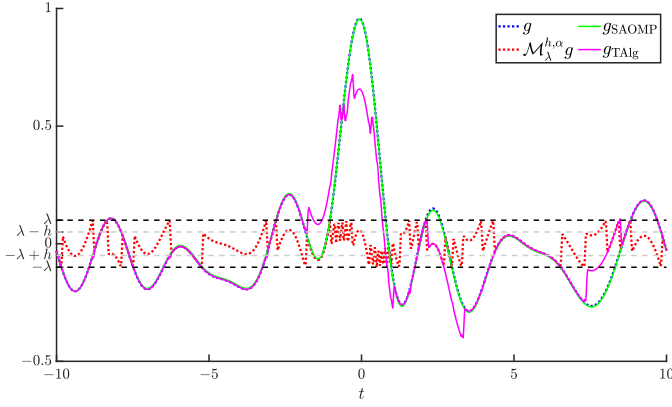


Fig. 2. Illustration of SAOMP and TAlg reconstruction for random $g \in \text{PW}_\Omega$ with $\Omega = 6.3 \frac{\text{rad}}{\text{s}}$ based on modulo hysteresis samples $\mathcal{M}_\lambda^{h,\alpha} g(nT)$ with $\lambda = 0.1$, $h = 0.05$, $\alpha = 50\text{ms}$ and $T = 20.8\text{ms}$.

and set $\varepsilon_\alpha^{(\kappa)}[n] = \varepsilon_\alpha((n-K)T - \kappa)$ for $n = 0, \dots, N$ with $N = 2K$. Let $\Delta : \mathbb{R}^{N+1} \rightarrow \mathbb{R}^N$ denote the forward difference operator defined by $\Delta z[k] = z[k+1] - z[k]$. Then, there are at most $L = \lceil \frac{\alpha}{T} \rceil + 1$ entries, where $\Delta \varepsilon_\alpha^{(\kappa)}$ is nonzero, so that $\Delta \varepsilon_\alpha^{(\kappa)}[n] = \Delta \varepsilon_0^{(\tilde{\kappa}_1)}[n] + \dots + \Delta \varepsilon_0^{(\tilde{\kappa}_L)}[n]$ for suitable $\tilde{\kappa}_1, \dots, \tilde{\kappa}_L \in \mathbb{T}\mathbb{Z}$. Intuitively, one linear fold is split into at most L instantaneous folds. Based on this observation we can adapt the recovery approach proposed in [16].

To this end, set $g[n] = g((n-K)T)$, for $n = 0, \dots, N$, and $g_\lambda[n] = \mathcal{M}_\lambda^{h,\alpha} g((n-K)T)$. Then, there exist $L_\lambda \in \mathbb{N}$, $\kappa_1, \dots, \kappa_{L_\lambda} \in [-K, K]T$ and $\sigma_1, \dots, \sigma_{L_\lambda} \in \{\pm 1\}$ such that $g_\lambda[n] = g[n] - s_\lambda[n]$ with $s_\lambda[n] = (2\lambda - h) \sum_{l=1}^{L_\lambda} \sigma_l \varepsilon_\alpha^{(\kappa_l)}[n]$. We apply Δ to get $\underline{g}_\lambda[n] = \Delta g_\lambda[n]$ for $n = 0, \dots, N-1$ and, thereon, calculate its discrete Fourier transform (DFT) $\hat{\underline{g}}_\lambda[m] = \sum_{n=0}^{N-1} \underline{g}_\lambda[n] \exp(-\frac{2\pi i m n}{N})$ for $m = 0, \dots, N-1$. Based on the above observation, the signal $\hat{\underline{g}}_\lambda$ can be written as

$$\hat{\underline{g}}_\lambda[m] = \sum_{\ell \in \mathbb{L}_\lambda} c_\ell \exp\left(-i \frac{\omega_0 m}{T} t_\ell\right)$$

with $\omega_0 = \frac{2\pi}{T}$, $\mathbb{L}_\lambda \subseteq \{0, \dots, N\}$ and $t_\ell \in (\mathbb{T}\mathbb{Z}) \cap [0, NT]$, where $|\mathbb{L}_\lambda| \leq L_\lambda (\lceil \frac{\alpha}{T} \rceil + 1)$. Now, the bandlimitedness of g gives

$$\hat{\underline{g}}_\lambda[m] = -\hat{\underline{s}}_\lambda[m] \quad \text{for } m \in \mathbb{E}_{N_\Omega, N}^c$$

with indices $\mathbb{E}_{N_\Omega, N}^c = \{N_\Omega + 1, \dots, N - N_\Omega - 1\}$ and effective bandwidth $N_\Omega = \lceil \frac{\Omega(N+1)T}{2\pi} \rceil$. Defining the vector $\mathbf{s} \in \mathbb{C}^M$ for $M = N - 2N_\Omega - 1$ via $\mathbf{s}_{m-N_\Omega} = \hat{\underline{g}}_\lambda[m]$ for $m \in \mathbb{E}_{N_\Omega, N}^c$, we can find $\mathbf{c} = (c_n)_{n=0}^{N-1} \in \mathbb{C}^N$ with $c_n = 0$ for $n \notin \mathbb{L}_\lambda$ by solving the minimization problem

$$\text{minimize } \|\mathbf{c}\|_0 \quad \text{subject to} \quad \mathbf{V}\mathbf{c} = \mathbf{s}, \quad (1)$$

where $\mathbf{V} \in \mathbb{C}^{M \times N}$ is a Vandermonde matrix with entries $\mathbf{V}_{m-N_\Omega, n+1} = e^{-i\omega_0 m n}$ for $m \in \mathbb{E}_{N_\Omega, N}^c$, $n = 0, \dots, N-1$. As explained in [16], (1) can be efficiently solved via the orthogonal matching pursuit (OMP) algorithm [17], see Algorithm 1. Applying anti-difference operator $\mathbf{S} : \mathbb{R}^N \rightarrow \mathbb{R}^{N+1}$, defined by $\mathbf{S}z[k] = \sum_{j < k} z[j]$, and subtracting $\mathbf{S}\mathbf{c}$ from $\mathbf{g}_\lambda = (g_\lambda[n])_{n=0}^N$ finally recovers the samples $\mathbf{g} = (g[n])_{n=0}^N$.

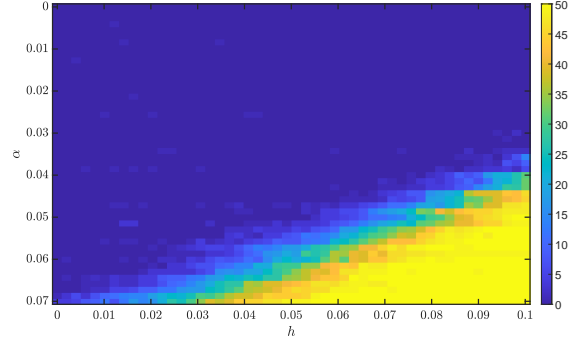


Fig. 3. Success of SAOMP reconstruction for $\mathcal{M}_\lambda^{h,\alpha} g(nT)$ with $\lambda = 0.1$, $T = 20.8\text{ms}$ and different values for α and h showing how often the MSE is larger than 0.001 for 50 random functions $g \in \text{PW}_\Omega$ with $\Omega = 6.3 \frac{\text{rad}}{\text{s}}$.

In [11], a recovery guarantee for solving (1) using OMP is shown if the index of the maximal nonzero entry of $\mathbf{c} \in \mathbb{C}^N$ satisfies $\max\{n \mid \mathbf{c}_n \neq 0\} \leq M - 1$. By means of Lemma 2 we know that this is satisfied if $g \in \mathcal{C}_{\lambda-h, \tau_0}^{1,1}$ with $\|g''\|_\infty \leq \frac{2h}{\alpha^2}$ and

$$\left\lceil \frac{|\tau_0| + 2\alpha}{T} \right\rceil + K \leq N - 2N_\Omega - 2.$$

To accelerate the recovery, we replace OMP by the so-called stagewise arithmetic orthogonal matching pursuit (SAOMP) algorithm proposed in [18]. It has additional parameters ν to find multiple folding points in one iteration and μ to remove incorrectly selected ones, see Algorithm 2. Note that SAOMP agrees with OMP when choosing $\nu = 1$, $\mu = 0$ and $i_{\max} = N$.

IV. NUMERICAL EXPERIMENTS

In our numerical experiments we randomly select a function $g \in \text{PW}_\Omega$ with $\Omega = 6.3$ and approximate its modulo hysteresis output $\mathcal{M}_\lambda^{h,\alpha} g$ numerically. Thereon, we reconstruct g from its samples $\mathcal{M}_\lambda^{h,\alpha} g(nT)$ using our SAOMP approach and compare our results with the thresholding algorithm (TAlg) proposed in [14]. One example is shown in Fig. 2, where we choose $\lambda = 0.1$, $h = 0.05$, $\alpha = 0.05$ and $T = 0.0208$.

While recovery with TAlg performs better for $\alpha \ll 1$ very close to zero, TAlg fails for larger α , whereas SAOMP still recovers, see Fig. 2. For a more detailed analysis, we investigate the influence of α and h on the reconstruction success of our SAOMP approach. To this end, we set $\lambda = 0.1$ and vary α between 0 and 0.07 and h between 0 and 0.1. We then calculate the mean squared error (MSE) for the SAOMP reconstructions of 50 randomly chosen functions. The results are depicted in Fig. 3, where the colors indicate for how many functions the MSE of SAOMP is larger than 0.001.

We observe that SAOMP gives satisfactory reconstructions for a large range of parameters. However, our approach shows instabilities for a large number of folds, in particular if α is much larger than T . In this case, a modified dictionary matrix \mathbf{V} can reduce the number of non-zero elements in the coefficient vector \mathbf{c} and allows for better reconstruction results. This approach, however, is beyond the scope of this paper and calls for more in-depth research in the future.

REFERENCES

- [1] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, 1948.
- [2] A. Bhandari, F. Krahmer, and R. Raskar, "On unlimited sampling," in *IEEE International Conference on Sampling Theory and Applications (SampTA)*, 2017, pp. 31–35.
- [3] —, "On unlimited sampling and reconstruction," *IEEE Transactions on Signal Processing*, vol. 69, pp. 3827–3839, 2020.
- [4] A. Bhandari, F. Krahmer, and T. Poskitt, "Unlimited sampling from theory to practice: Fourier-Prony recovery and prototype ADC," *IEEE Transactions on Signal Processing*, vol. 70, pp. 1131–1141, 2021.
- [5] S. Rudresh, A. Adiga, B. A. Shenoy, and C. S. Seelamantula, "Wavelet-based reconstruction for unlimited sampling," in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2018, pp. 4584–4588.
- [6] E. Romanov and O. Ordentlich, "Above the nyquist rate, modulo folding does not hurt," *IEEE Signal Processing Letters*, vol. 26, no. 8, pp. 1167–1171, 2019.
- [7] E. Azar, S. Mulleti, and Y. C. Eldar, "Residual recovery algorithm for modulo sampling," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2022, pp. 5722–5726.
- [8] R. Guo and A. Bhandari, "ITER-SIS: Robust unlimited sampling via iterative signal sieving," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2023, pp. 1–5.
- [9] A. Bhandari and F. Krahmer, "HDR imaging from quantization noise," in *IEEE International Conference on Image Processing (ICIP)*, 2020, pp. 101–105.
- [10] M. Beckmann, A. Bhandari, and F. Krahmer, "The modulo Radon transform: Theory, algorithms and applications," *SIAM Journal on Imaging Sciences*, vol. 15, no. 2, pp. 455–490, 2022.
- [11] M. Beckmann, A. Bhandari, and M. Iske, "Fourier-domain inversion for the modulo Radon transform," *IEEE Transactions on Computational Imaging*, vol. 10, pp. 653–665, 2024.
- [12] T. Feuillen, B. Shankar, and A. Bhandari, "Unlimited sampling radar: Life below the quantization noise," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2023, pp. 1–5.
- [13] E. Azar, S. Mulleti, and Y. C. Eldar, "Unlimited sampling beyond modulo," *Applied and Computational Harmonic Analysis*, vol. 74, p. 101715, 2025.
- [14] D. Florescu, F. Krahmer, and A. Bhandari, "The surprising benefits of hysteresis in unlimited sampling: Theory, algorithms and experiments," *IEEE Transactions on Signal Processing*, vol. 70, pp. 616–630, 2022.
- [15] D. Florescu and A. Bhandari, "Unlimited sampling via generalized thresholding," in *IEEE International Symposium on Information Theory (ISIT)*, 2022, pp. 1606–1611.
- [16] M. Beckmann and A. Bhandari, "MR. TOMP: Inversion of the modulo Radon transform (MRT) via orthogonal matching pursuit (OMP)," in *IEEE International Conference on Image Processing (ICIP)*, 2022, pp. 3748–3752.
- [17] S. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Transactions on Signal Processing*, vol. 41, no. 12, pp. 3397–3415, 1993.
- [18] Y. Zhang and G. Sun, "Stagewise arithmetic orthogonal matching pursuit," *International Journal of Wireless Information Networks*, vol. 25, pp. 221–228, 2018.