

Generalized Modulo Hysteresis Encoding

Matthias Beckmann and Jürgen Jeschke

Abstract—Unlimited sensing and its extension via modulo hysteresis provide an efficient encoding scheme for high dynamic range signals by folding the signal’s amplitude into the range of the analog-to-digital converter prior to sampling. Motivated by shortcomings of the modulo hysteresis model, in this work we introduce generalized modulo encoders that flexibly design the folding by incorporating a general transition function and allowing the folding times to depend on the nonideality of the folding transition. We rigorously study mathematical properties of the respective operators and derive conditions under which a bandlimited signal is uniquely determined by its generalized modulo samples. Finally, we propose a recovery algorithm backed by mathematical guarantees and illustrate our theoretical results through numerical experiments.

Index Terms—Unlimited sensing, modulo encoding, generalized modulo operator, Shannon sampling theory, sparse recovery

I. INTRODUCTION

By the classical Shannon sampling theorem [1] any bandlimited signal g with bandwidth $\Omega \frac{\text{rad}}{\text{s}}$ is uniquely determined by equispaced samples $\{g(nT) \mid n \in \mathbb{Z}\}$ with sampling rate $T > 0$ if sampled at or above Nyquist rate, i.e., if $T \leq \frac{\pi}{\Omega}$. Moreover, recovery is achieved by employing the well-known Whittaker-Shannon-Kotelnikov interpolation formula

$$g(t) = \sum_{n \in \mathbb{Z}} g(nT) \operatorname{sinc}\left(\frac{\pi}{T}(t - nT)\right).$$

In hardware, this fundamental principle is implemented by analog-to-digital converters (ADCs) [2]. By design, conventional ADCs operate at a fixed dynamic range and can only capture a quantized signal depending on the available bit-budget resulting in inevitable quantization noise. This imposes a fundamental limit on the digital resolution leading to an intrinsic trade-off between resolution and dynamic range. In particular, if the input signal exceeds the capacity of the ADC, it leads to saturation or clipping and, hence, a permanent loss of information, which compromises the reconstruction quality.

To account for these limitations, the Unlimited Sampling Framework (USF) has been introduced in [3]–[5], a fundamentally different sampling paradigm based on the observation that any function can be decomposed into its *integer part*, corresponding to the quantized signal, and its *fractional part*, representing the quantization noise. The key enabler is then the following mathematical insight: *for smooth functions, their fractional part encodes the integer part* [6]. Hence, USF suggests to capture the fractional part by folding the input

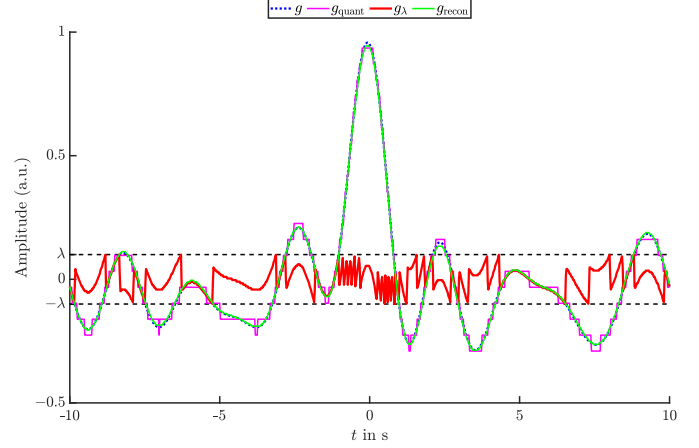


Fig. 1. Illustration of generalized modulo encoding. Capturing the quantized signal g_{quant} with 5 bits yields a mean squared error of $3.6074 \cdot 10^{-4}$. As opposed to this, capturing the modulo signal g_{λ} with the same bit-budget allows for the reconstruction g_{recon} with mean squared error of $3.2712 \cdot 10^{-5}$.

signal prior to sampling into the dynamic range of the ADC by means of modulo arithmetic. More precisely, the *ideal modulo encoder* $\mathcal{M}_{\lambda}$ with threshold $\lambda > 0$ is for $g: \mathbb{R} \rightarrow \mathbb{R}$ defined via

$$\mathcal{M}_{\lambda}g(t) = g(t) - 2\lambda \left\lfloor \frac{g(t) + \lambda}{2\lambda} \right\rfloor,$$

where $\lfloor \cdot \rfloor$ denotes the floor function and $\mathcal{M}_{\lambda}g(t) \in [-\lambda, \lambda]$. The Unlimited Sampling Theorem [3, Theorem 1] then states that $g \in \text{PW}_{\Omega}$ can be recovered (up to an additive constant) from $\{\mathcal{M}_{\lambda}g(nT) \mid n \in \mathbb{Z}\}$ if $T < \frac{1}{\Omega_e}$. The first modulo ADC prototype was introduced in [5] and since then several hardware implementations have been proposed, e.g. [6]–[11], along with various reconstruction approaches, see [12]–[20]. Moreover, USF has been extended to diverse applications like, for instance, imaging [21]–[23], communication [24]–[26], electroencephalography [27], computerized tomography [28]–[30] and radar [31]–[33].

The ideal modulo encoder $\mathcal{M}_{\lambda}$ assumes an instantaneous fold whenever the input signal reaches the dynamic threshold $\pm\lambda$. However, as argued in [34] such exact folding transitions cannot be realized due to hardware imperfections. As opposed to this, in [34]–[36] the modulo hysteresis encoder $\mathcal{M}_H$ with $H = (\lambda, h, \alpha)$ is introduced, which accounts for non-instantaneous folds by incorporating a hysteresis parameter $h \geq 0$ that models an imperfect alignment of the reset threshold and the post-reset value, and a transient parameter $\alpha \geq 0$ accounting for an inexact folding transition. However, $\mathcal{M}_H$ has the following limitations that motivate further refinement:

- I) $\mathcal{M}_H g$ is not always guaranteed to stay in $[-\lambda, \lambda]$;
- II) folding transition is modelled as a straight line;
- III) folding times are independent of the transient nonideality.

This work was supported by Deutsche Forschungsgemeinschaft (DFG) - Project number 530863002.

M. Beckmann is with the Dept. of Mathematics, University of Hamburg, Germany (Email: research@mbeckmann.de).

J. Jeschke is with the Center for Industrial Mathematics, University of Bremen, Germany (Email: jjeschke@uni-bremen.de).

The authors are listed in alphabetical order.

In our preliminary work [37] we addressed I) by proposing the modified modulo hysteresis operator $\mathcal{M}_\lambda^{h,\alpha}$, which makes the folding positions dependent on the folding transition, as also observed in modulo hardware in [9]. In its definition the folding transition is still modelled as a linear function. As this is a rather restrictive assumption that appears unrealistic in practical implementations, in this paper we generalize the operators $\mathcal{M}_H$ and $\mathcal{M}_\lambda^{h,\alpha}$ by incorporating a general folding function j_α . To emphasize their origins, we reuse the notation $\mathcal{M}_H$ with $H = (\lambda, h, j_\alpha)$ for the former and $\mathcal{M}_\lambda^{h,j_\alpha}$ for the latter. As before, we will see that $\mathcal{M}_H g$ may exceed the range $[-\lambda, \lambda]$, while $\mathcal{M}_\lambda^{h,j_\alpha} g$ is guaranteed to stay in $[-\lambda, \lambda]$ under assumptions only on j_α . To allow for a larger class of folding functions, we introduce a reset time $\sigma > 0$ to $\mathcal{M}_\lambda^{h,j_\alpha}$ as an additional model parameter and will see that the resulting operator $\mathcal{M}_{\lambda,\sigma}^{h,j_\alpha}$ exceeds the range $[-\lambda, \lambda]$ in a controlled way, which has also been reported for some modulo ADC prototypes in [9], [11], [34]. We study the mathematical properties of the three new operators we refer to as *generalized modulo hysteresis operators*. To this end, we focus on Lipschitz continuous input functions, which include the classical bandlimited functions, but also other function classes studied in the literature like B-splines, recently discussed in [38]. Our generalized modulo hysteresis operators lead to a *perturbed* decomposition of the signal g into integer and fractional part. Under the additional assumption that g is bandlimited, we formulate conditions under which the signal can be identified by its perturbed fractional part g_λ . For inversion we then describe a generic recovery approach that can be adapted to the specific model to improve reconstruction quality. This yields a flexible and effective encoding scheme for bandlimited signals, which is illustrated in Fig. 1 using only 5 bits for encoding.

Contribution: We advance the existing art and theory of generalized modulo encoders by

- i) developing flexible models for modulo hysteresis with general folding transition functions that can be adapted to pre-described desired properties of the encoding and its actual hardware implementation,
- ii) rigorously analysing the mathematical properties of the introduced generalized modulo hysteresis operators and conditions that guarantee
 - bounded range and folding at $\pm\lambda$,
 - separation of folds,
 - return property for sufficiently large $t \in \mathbb{R}$,
- iii) proving identifiability conditions for bandlimited functions from generalized modulo samples with minimal oversampling,
- iv) describing a generic reconstruction algorithm from discrete data based on orthogonal matching pursuit, which is applicable to all settings.

Our accurate modelling in i) allows us to adapt the reconstruction algorithm in iv) to the respective generalized modulo operator, which leads to improved recovery results especially if the transient parameter is larger than the sampling rate.

Paper overview: This paper is organized as follows. In Section II we introduce and analyze our novel generalized modulo hysteresis encoders. In Section III we compare the

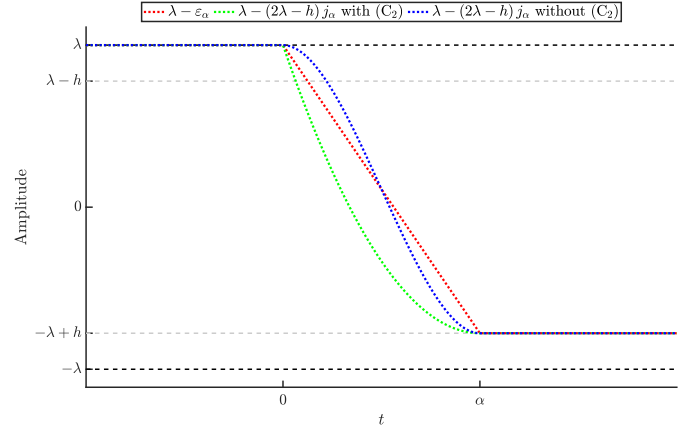


Fig. 2. Illustration of the transition function in [34], [37] (red) and a folding function from Definition 1 satisfying (C₂) (green) and violating (C₂) (blue).

different models and summarize the mathematical properties in Table I. Section IV is devoted to our recovery approach with theoretical guarantees and adaptation to the respective modulo operators. In Section V we provide numerical experiments which showcase the methods developed in this paper. Finally, we summarize our results and point to future research directions in Section VI.

II. GENERALIZED MODULO HYSTERESIS

Operators

We start by recalling the generalized modulo encoder $\mathcal{M}_H$ from [34], where $H = (\lambda, h, \alpha)$ with modulo threshold λ , hysteresis parameter h and transient parameter α . For a real-valued function $g \equiv g(t)$ it is defined via a sequence of folding points $(\tau_p)_{p \in \mathbb{N}}$, where $\tau_1 = \inf\{t > \tau_0 \mid \mathcal{M}_\lambda(g(t) + \lambda) = 0\}$ and $\tau_{p+1} = \inf\{t > \tau_p \mid \mathcal{M}_\lambda(g(t) - g(\tau_p) + h s_p) = 0\}$ with $s_p = \text{sgn}(g(\tau_p) - g(\tau_{p-1}))$. Then, for $t \in \mathbb{R}$ the output is given by $\mathcal{M}_H g(t) = g(t) - (2\lambda - h) \sum_{p \in \mathbb{N}} s_p \varepsilon_\alpha(t - \tau_p)$ with $\varepsilon_0(t) = \mathbb{1}_{[0, \infty)}(t)$ and $\varepsilon_\alpha(t) = \frac{t}{\alpha} \mathbb{1}_{[0, \alpha)}(t) + \mathbb{1}_{[\alpha, \infty)}(t)$.

To account for limitation II), we now generalize this operator by incorporating a function that models the way how a non-instantaneous fold with transient α is applied. In the following, this function is referred to as *folding function*, see Fig. 2 for illustration.

Definition 1. A function $j_\alpha: \mathbb{R} \rightarrow \mathbb{R}$ with transient $\alpha \geq 0$ is called folding function if $j_0 = \mathbb{1}_{[0, \infty)}$ and, for $\alpha > 0$, j_α is continuous with

$$\begin{aligned} \forall t \leq 0: j_\alpha(t) &= 0, \\ \forall t \geq \alpha: j_\alpha(t) &= 1. \end{aligned}$$

When stated explicitly, we additionally assume that

$$\forall t \in [0, \alpha]: \frac{t}{\alpha} \leq j_\alpha(t) \leq 1 \quad (\text{C}_1)$$

and that the right derivative $\partial_+ j_\alpha$ exists on $[0, \alpha]$ with

$$\forall s \leq t \in [0, \alpha]: 0 \leq \partial_+ j_\alpha(t) \leq \partial_+ j_\alpha(s). \quad (\text{C}_2)$$

A. Transient independent folding points

In the following, we consider the triplet $H = (\lambda, h, j_\alpha)$ and modify the definition of the generalized modulo encoder $\mathcal{M}_H$ accordingly. To ensure well-definedness, we restrict ourselves to the function class

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

with arbitrary $\tau_0 \in \mathbb{R}$, 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, whose essential supremum yields a Lipschitz constant of the function.

Definition 2. Let $\lambda > 0$, $h \in (0, 2\lambda)$ and j_α be a folding function with transient $\alpha \geq 0$. For a function $g \in \mathcal{C}_{\lambda, \tau_0}^{0,1}$ with $\tau_0 \in \mathbb{R}$ let $(\tau_p)_{p \in \mathbb{N}}$ be given by

$$\begin{aligned} \tau_1 &= \inf\{t > \tau_0 \mid \mathcal{M}_\lambda(g(t) + \lambda) = 0\}, \\ \tau_{p+1} &= \inf\{t > \tau_p \mid \mathcal{M}_\lambda(g(t) - g(\tau_p) + h s_p) = 0\}, \end{aligned} \quad (1)$$

where $s_p = \text{sgn}(g(\tau_p) - g(\tau_{p-1}))$ if $\tau_p < \infty$ and $\inf \emptyset = \infty$. Then, the output of the generalized modulo encoder $\mathcal{M}_H$ with $H = (\lambda, h, j_\alpha)$ is defined as

$$\mathcal{M}_H g(t) = g(t) - (2\lambda - h) \sum_{\tau_p < \infty} s_p j_\alpha(t - \tau_p). \quad (2)$$

Note that choosing $j_\alpha = \varepsilon_\alpha$ in Definition 2 yields the same generalized modulo encoder as in [34], which justifies the clash of notation. We now study elementary properties of $\mathcal{M}_H$ and begin with its well-definedness by proving that the series in (2) converges pointwise.

Theorem 1. Let $g \in \mathcal{C}_{\lambda, \tau_0}^{0,1}$. Then, the folding points $(\tau_p)_{p \in \mathbb{N}}$ satisfy the separation property

$$\tau_p < \infty \implies \tau_{p+1} - \tau_p \geq \frac{\min\{h, 2\lambda - h\}}{\|g'\|_\infty} > 0$$

and the series $\sum_{p \in \mathbb{N}} s_p j_\alpha(t - \tau_p)$ converges for all $t \in \mathbb{R}$.

Proof. Let $\tau_p < \infty$ and assume that $\tau_{p+1} < \infty$. Otherwise, the separation is trivially true. By definition, τ_{p+1} is the first point larger than τ_p which satisfies $\mathcal{M}_\lambda(g(\tau_{p+1}) - g(\tau_p) + h s_p) = 0$. Because of the continuity of g this is the first point where $g(\tau_{p+1}) - g(\tau_p) + h s_p = 0$ or $g(\tau_{p+1}) - g(\tau_p) + h s_p = 2\lambda s_p$. Therefore, this implies that

$$|g(\tau_{p+1}) - g(\tau_p)| \geq \min\{h, 2\lambda - h\}.$$

As g is Lipschitz continuous, the difference $|g(\tau_{p+1}) - g(\tau_p)|$ can be bounded via $|g(\tau_{p+1}) - g(\tau_p)| \leq \|g'\|_\infty (\tau_{p+1} - \tau_p)$. With this, we obtain the stated separation

$$\tau_{p+1} - \tau_p \geq \frac{\min\{h, 2\lambda - h\}}{\|g'\|_\infty}.$$

Now, for fixed $t \in \mathbb{R}$ there exists $q \in \mathbb{N}$ such that $t \leq \tau_{q+1}$, e.g., $q = \max\{0, \lceil (t - \tau_0) \frac{\|g'\|_\infty}{\min\{h, 2\lambda - h\}} \rceil\}$. Therefore, the series

$$\sum_{p \in \mathbb{N}} s_p j_\alpha(t - \tau_p) = \sum_{p=1}^{q+1} s_p j_\alpha(t - \tau_p).$$

reduces to a finite sum and, hence, converges. $\square$

Theorem 1 shows that $\mathcal{M}_H$ is well-defined on $\mathcal{C}_{\lambda, \tau_0}^{0,1}$. We now study conditions to guarantee that $\mathcal{M}_H g$ returns to g for

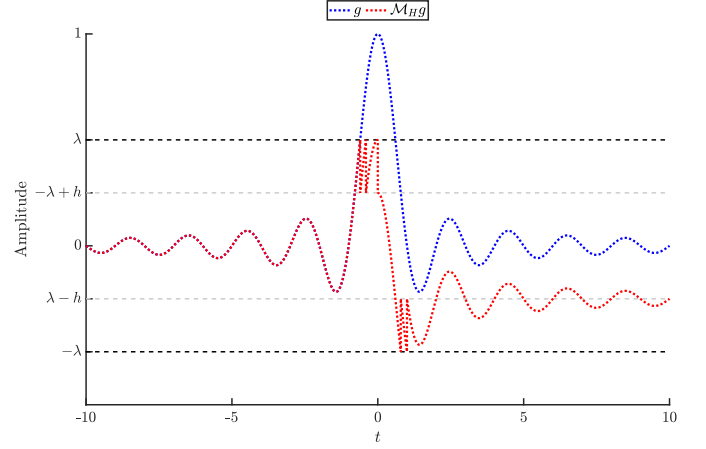


Fig. 3. Illustration of the failure of the return property in the case $h > \lambda$ using $g(t) = \text{sinc}(t)$, where we choose $\lambda = 0.5$, $h = 0.75$ and $\alpha = 0$.

sufficiently large $t \in \mathbb{R}$, see also Fig. 3. To this end, for $\delta > 0$ and $\rho \in \mathbb{R}$ we consider the function class

$$\mathcal{B}_\rho^\delta = \{g: \mathbb{R} \rightarrow \mathbb{R} \mid |g(t)| < \delta \ \forall t \geq \rho\}.$$

Proposition 1. Let $0 < h < \lambda$ and $g \in \mathcal{C}_{\lambda, \tau_0}^{0,1} \cap \mathcal{B}_\rho^{\lambda-h}$ with $\rho > \tau_0$. Then, $\mathcal{M}_H g(t) = g(t)$ for all $t \geq \rho + \alpha$.

Proof. First consider $\alpha = 0$. Then, we can write $\mathcal{M}_H g(t) = g(t) + c(2\lambda - h)$ for some $c \in \mathbb{Z}$ dependent on t . As $g \in \mathcal{B}_\rho^{\lambda-h}$, for $t \geq \rho$ we get $g(t) + c(2\lambda - h) < \lambda - h + c(2\lambda - h)$, which can only be larger than or equal to $-\lambda$ if $c \geq 0$. Analogously, we get $g(t) + c(2\lambda - h) > -\lambda + h + c(2\lambda - h)$, which can only be smaller than or equal to λ if $c \leq 0$. This implies that $c = 0$ so that indeed $\mathcal{M}_H g(t) = g(t)$ for all $t \geq \rho$.

To prove the statement for $\alpha > 0$, consider $H_0 = (\lambda, h, j_0)$. We have already proven that $\mathcal{M}_{H_0} g(t) = g(t)$ for all $t \geq \rho$ and $\tau_p < \rho$ or $\tau_p = \infty$. Because $\mathcal{M}_{H_0} g(t) = \mathcal{M}_H g(t)$ by definition for all $t \notin \bigcup_{\tau_p < \infty} [\tau_p, \tau_p + \alpha) \subset [\tau_0, \rho + \alpha)$ we also have $\mathcal{M}_H g(t) = g(t)$ for all $t \geq \rho + \alpha$. $\square$

If $0 < h < \lambda$ and $g \in \mathcal{B}_\rho^{\lambda-h}$ with $\rho > \tau_0$, Proposition 1 implies that $\tau_p < \rho$ or $\tau_p = \infty$. This in combination with Theorem 1 gives

$$\sum_{p \in \mathbb{N}} s_p j_\alpha(t - \tau_p) = \sum_{p=1}^{q+1} s_p j_\alpha(t - \tau_p)$$

with $q = \max\{0, \lceil (\rho - \tau_0) \frac{\|g'\|_\infty}{h} \rceil\}$ for all $t \in \mathbb{R}$. Therefore, we can guarantee uniform convergence of $\sum_{p \in \mathbb{N}} s_p j_\alpha(\cdot - \tau_p)$ for $0 < h < \lambda$ and $g \in \mathcal{B}_\rho^{\lambda-h}$.

Next, we discuss limitation I) and study conditions under which $\mathcal{M}_H g$ stays in the prescribed range $[-\lambda, \lambda]$. In the case $\alpha = 0$, the condition $g \in \mathcal{C}_{\lambda, \tau_0}^{0,1}$ suffices to guarantee that $|\mathcal{M}_H g(t)| \leq \lambda$ for all $t \in \mathbb{R}$ and $|\mathcal{M}_H g(\tau_p)| = \lambda - h$ if $\tau_p < \infty$. However, by defining the sequence of folding points $(\tau_p)_{p \in \mathbb{N}}$ independently of the transient α it may occur that $\mathcal{M}_H g(t) \notin [-\lambda, \lambda]$ and $|\mathcal{M}_H g(\tau_p)| \neq \lambda$ if $\alpha > 0$, see Fig. 4. To prevent this, we need additional conditions on the first and second derivative of the input function and consider the function class

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

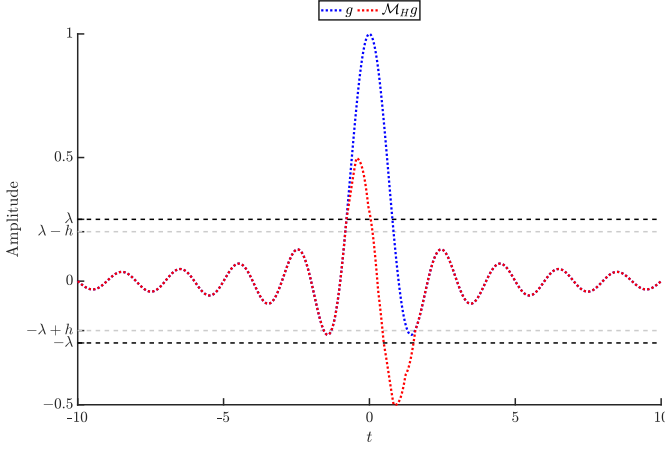


Fig. 4. Example for $\mathcal{M}_H g(t) \notin [-\lambda, \lambda]$ using $g(t) = \text{sinc}(\pi t)$, where we choose $j_\alpha(t) = \frac{t}{\alpha}$ for $t \in [0, \alpha]$, $\lambda = 0.25$, $h = 0.05$ and $\alpha = 0.75$.

Theorem 2. Let $\alpha > 0$ and let j_α satisfy (C₁). Moreover, let $g \in \mathcal{C}_{\lambda, \tau_0}^{1,1}$ with $\|g'\|_\infty \leq \frac{2\lambda-h}{\alpha}$ and $\|g''\|_\infty \leq \frac{2h}{\alpha^2}$. Then, we have $|\mathcal{M}_H g(t)| \leq \lambda$ for all $t \in \mathbb{R}$. If $\tau_0 < \tau_p < \infty$, we further obtain $\tau_{p+1} - \tau_p \geq \alpha$ and $|\mathcal{M}_H g(\tau_p)| = \lambda$.

Proof. We start with showing $\tau_{p+1} - \tau_p \geq \alpha$ for $\tau_0 < \tau_p < \infty$. If $\tau_{p+1} < \infty$, we either have $g(\tau_{p+1}) = g(\tau_p) - h s_p$ and $s_{p+1} = -s_p$ or $g(\tau_{p+1}) = g(\tau_p) + (2\lambda - h)s_p$ and $s_{p+1} = s_p$. In the latter case, the assumption $\|g'\|_\infty \leq \frac{2\lambda-h}{\alpha}$ yields

$$\tau_{p+1} - \tau_p \geq \frac{|g(\tau_{p+1}) - g(\tau_p)|}{\|g'\|_\infty} \geq \alpha.$$

In the former case, we have $s_p \cdot g'(\tau_p) \geq 0$ and $\|g''\|_\infty \leq \frac{2h}{\alpha^2}$ yields

$$\tau_{p+1} - \tau_p \geq \sqrt{\frac{2h}{\|g''\|_\infty}} \geq \alpha.$$

We now prove by induction on $p \in \mathbb{N}$ that $|\mathcal{M}_H g(t)| \leq \lambda$ for $t < \tau_p$ and $\mathcal{M}_H g(\tau_p) = \lambda s_p$ if $\tau_0 < \tau_p < \infty$. Since $\tau_p \rightarrow \infty$ for $p \rightarrow \infty$ due to Theorem 1, this implies that $|\mathcal{M}_H g(t)| \leq \lambda$ for all $t \in \mathbb{R}$. By definition of τ_1 we have $|\mathcal{M}_H g(t)| \leq \lambda$ for all $t < \tau_1$ as well as $\mathcal{M}_H g(\tau_1) = g(\tau_1) = \lambda s_1$ if $\tau_1 < \infty$. Now, let $p \in \mathbb{N}$ such that $|\mathcal{M}_H g(t)| \leq \lambda$ for $t < \tau_p$ and $\mathcal{M}_H g(\tau_p) = \lambda s_p$ if $\tau_p < \infty$. For $t \in (\tau_p, \tau_{p+1})$ we can write $\mathcal{M}_H g(t) = \mathcal{M}_H g(\tau_p) - g(\tau_p) + g(t) - (2\lambda - h)s_p j_\alpha(t - \tau_p)$, where $0 \leq j_\alpha \leq 1$. If $s_p = 1$, we have $g(\tau_p) - h \leq g(t) \leq g(\tau_p) + 2\lambda - h$ as well as $\mathcal{M}_H g(\tau_p) = \lambda$ and can directly conclude that

$$\mathcal{M}_H g(t) \geq \lambda - g(\tau_p) + g(\tau_p) - h - 2\lambda + h = -\lambda.$$

Moreover, for $t \geq \tau_p + \alpha$ we have $j_\alpha(t - \tau_p) = 1$ and, therefore,

$$\begin{aligned} \mathcal{M}_H g(t) &= \mathcal{M}_H g(\tau_p) - g(\tau_p) + g(t) - 2\lambda + h \\ &\leq \lambda - g(\tau_p) + g(\tau_p) + 2\lambda - h - 2\lambda + h = \lambda. \end{aligned}$$

For $t \leq \tau_p + \alpha$, condition (C₁) gives $j_\alpha(t - \tau_p) \geq \frac{t - \tau_p}{\alpha}$ and, hence,

$$\begin{aligned} \mathcal{M}_H g(t) &= \mathcal{M}_H g(\tau_p) - (2\lambda - h)j_\alpha(t - \tau_p) + g(t) - g(\tau_p) \\ &\leq \lambda - (2\lambda - h)\frac{t - \tau_p}{\alpha} + (t - \tau_p)\|g'\|_\infty \leq \lambda. \end{aligned}$$

On the other hand, if $s_p = -1$, we have $g(\tau_p) - (2\lambda - h) \leq g(t) \leq g(\tau_p) + h$ and $\mathcal{M}_H g(\tau_p) = -\lambda$ so that

$$\mathcal{M}_H g(t) \leq -\lambda - g(\tau_p) + g(\tau_p) + h + 2\lambda - h = \lambda.$$

As before, for $t > \tau_p + \alpha$, we have $j_\alpha(t - \tau_p) = 1$ and, hence,

$$\begin{aligned} \mathcal{M}_H g(t) &= \mathcal{M}_H g(\tau_p) - g(\tau_p) + g(t) + 2\lambda - h \\ &\geq -\lambda - g(\tau_p) + g(\tau_p) - (2\lambda - h) + 2\lambda - h = -\lambda. \end{aligned}$$

For $t \leq \tau_p + \alpha$, condition (C₁) gives $j_\alpha(t - \tau_p) \geq \frac{t - \tau_p}{\alpha}$ and

$$\begin{aligned} \mathcal{M}_H g(t) &= \mathcal{M}_H g(\tau_p) + (2\lambda - h)j_\alpha(t - \tau_p) + g(t) - g(\tau_p) \\ &\geq -\lambda + (2\lambda - h)\frac{t - \tau_p}{\alpha} - (t - \tau_p)\|g'\|_\infty \geq \lambda. \end{aligned}$$

Consequently, we have $|\mathcal{M}_H g(t)| \leq \lambda$ for $t < \tau_{p+1}$. Finally, if $\tau_{p+1} < \infty$, we either have $g(\tau_{p+1}) = g(\tau_p) - h s_p$ and $s_{p+1} = -s_p$ implying $\mathcal{M}_H g(\tau_{p+1}) = \lambda s_p + (2\lambda - h)s_p - (2\lambda - h)s_p = \lambda s_{p+1}$ or $g(\tau_{p+1}) = g(\tau_p) + (2\lambda - h)s_p$ and $s_{p+1} = s_p$ so that $\mathcal{M}_H g(\tau_{p+1}) = \lambda s_p - h s_p - (2\lambda - h)s_p = \lambda s_{p+1}$. $\square$

B. Transient dependent folding points

As described in limitation I) and seen in Fig. 4, the above definition of $\mathcal{M}_H$ does not guarantee that $\mathcal{M}_\lambda g(t) \in [-\lambda, \lambda]$ for $t \geq \tau_0$. One reason for this is that the folding points $(\tau_p)_{p \in \mathbb{N}}$ are independent of the folding function j_α if $\alpha > 0$, as already indicated in limitation III). To account for this, Theorem 2 included an additional condition on the growth of the input function. Instead, we now wish to propose a modified definition of the generalized modulo encoder. To this end, we rewrite $\mathcal{M}_H g$ by defining the function sequence $(\eta_n^{(\alpha)})_{n \in \mathbb{N}_0}$ via

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

if $\tau_{n+1} < \infty$ and $\eta_{n+1}^{(\alpha)} = \eta_n^{(\alpha)}$ if $\tau_{n+1} = \infty$, where $\eta_0^{(\alpha)} = g$. With this, we then obtain

$$\tau_{n+1} = \inf\{t > \tau_n \mid |\eta_n^{(0)}(t)| \geq \lambda\}$$

and

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

The idea is to make τ_{n+1} dependent on $\eta_n^{(\alpha)}$ rather than $\eta_n^{(0)}$.

Definition 3. Let $\lambda > 0$, $h < 2\lambda$ and j_α be a folding function with transient $\alpha > 0$ satisfying (C₂). For a function $g \in \mathcal{C}_{\lambda, \tau_0}^{0,1}$ with $\tau_0 \in \mathbb{R}$ we define the two sequences $(\kappa_n)_{n \in \mathbb{N}_0}$ and $(\zeta_n)_{n \in \mathbb{N}_0}$ via

$$\begin{aligned} \kappa_{n+1} &= \inf\{t > \kappa_n \mid |\zeta_n(t)| \geq \lambda\}, \\ \zeta_{n+1} &= \zeta_n - (2\lambda - h) \text{sgn}(\zeta_n(\kappa_{n+1})) j_\alpha(\cdot - \kappa_{n+1}) \end{aligned} \quad (3)$$

and $\zeta_{n+1} = \zeta_n$ if $\kappa_{n+1} = \infty$, where $\kappa_0 = \tau_0$ and $\zeta_0 = g$. Then, the output of the generalized modulo operator $\mathcal{M}_\lambda^{h, j_\alpha}$ is defined as

$$\mathcal{M}_\lambda^{h, j_\alpha} g(t) = \lim_{n \rightarrow \infty} \zeta_n(t). \quad (4)$$

Note that for $\alpha = 0$ and $0 < h < 2\lambda$ Definition 3 would yield the same generalized modulo encoder as Definition 2 so that $\mathcal{M}_\lambda^{h, j_0} = \mathcal{M}_{H_0}$ if $h > 0$. Hence, we require that $\alpha > 0$, in which case also a negative $h < 2\lambda$ is applicable. We now study

basic properties of $\mathcal{M}_\lambda^{h,j_\alpha}$ and begin with its well-definedness by proving that the sequence in (4) converges.

Proposition 2. *For $g \in \mathcal{C}_{\lambda,\tau_0}^{0,1}$ the function sequence $(\zeta_n)_{n \in \mathbb{N}_0}$ converges pointwise and for every $t \in \mathbb{R}$ there exists $p \in \mathbb{N}_0$ with $\mathcal{M}_\lambda^{h,j_\alpha} g(t) = \zeta_p(t)$.*

Proof. We first prove the statement for $g \in \mathcal{C}_{\lambda,\tau_0}^{0,1} \cap L^\infty(\mathbb{R})$ by showing that $(\kappa_n)_{n \in \mathbb{N}_0}$ diverges to infinity. Then, for each $t \in \mathbb{R}$ there exists $p \in \mathbb{N}_0$ such that $t < \kappa_p$ and we obtain $\zeta_n(t) = \zeta_p(t)$ for all $n \geq p$.

We consider the subsequence $(\kappa_{m_n})_{n \in \mathbb{N}_0}$ with $\kappa_{m_1} = \kappa_1$, where $\kappa_{m_{n+1}}$ is the next folding point with different sign, i.e. $\kappa_{m_{n+1}} = \inf_{k > m_n} \{\kappa_k \mid \text{sgn}(\zeta_{k-1}(\kappa_k)) \neq \text{sgn}(\zeta_{m_n}(\kappa_{m_n}))\}$. We now show that $m_{n+1} - m_n \leq n c_g$ by induction on $n \in \mathbb{N}$ with $c_g = \lceil \frac{\|g\|_\infty}{\lambda} \rceil \cdot \lceil \frac{\alpha \|g'\|_\infty}{2\lambda - h} \rceil$. For $n = 1$, we directly have

$$m_2 - m_1 \leq \left\lceil \frac{\|g\|_\infty}{\lambda} \right\rceil \cdot \left\lceil \frac{\alpha \|g'\|_\infty}{2\lambda - h} \right\rceil = c_g,$$

where $\lceil \frac{\|g\|_\infty}{\lambda} \rceil$ bounds the number of folds with distance larger than or equal to α and $\lceil \frac{\alpha \|g'\|_\infty}{2\lambda - h} \rceil$ bounds the number of folds that can happen while resetting in an interval of length α . For $n \geq 2$, we have to incorporate the case that previous non-instantaneous folds may induce additional folds, leading to

$$\begin{aligned} m_{n+1} - m_n &\leq \left\lceil \frac{\|g\|_\infty}{\lambda} \right\rceil \cdot \left\lceil \frac{\alpha \|g'\|_\infty}{2\lambda - h} \right\rceil \\ &\quad + \left\lceil \frac{\partial_+(\zeta_{m_{n-1}}(\kappa_{m_{n-1}}) - g(\kappa_{m_{n-1}}))}{(2\lambda - h) \cdot \partial_+ j_\alpha(0)} \right\rceil \\ &\leq \left\lceil \frac{\|g\|_\infty}{\lambda} \right\rceil \cdot \left\lceil \frac{\alpha \|g'\|_\infty}{2\lambda - h} \right\rceil + \left\lceil \sum_{i=m_{n-1}}^{m_n-1} \frac{\partial_+ j_\alpha(t_i)}{\partial_+ j_\alpha(0)} \right\rceil \end{aligned}$$

with certain points $t_i \in [0, \alpha]$. As Assumption (C₂) gives $\partial_+ j_\alpha(t_i) \leq \partial_+ j_\alpha(0)$, we can bound $m_{n+1} - m_n$ via

$$\begin{aligned} m_{n+1} - m_n &\leq \left\lceil \frac{\|g\|_\infty}{\lambda} \right\rceil \cdot \left\lceil \frac{\alpha \|g'\|_\infty}{2\lambda - h} \right\rceil + m_n - m_{n-1} \\ &\leq n \left\lceil \frac{\|g\|_\infty}{\lambda} \right\rceil \cdot \left\lceil \frac{\alpha \|g'\|_\infty}{2\lambda - h} \right\rceil = n c_g. \end{aligned}$$

By definition of κ_{m_n} , we have

$$\begin{aligned} \kappa_{m_n} &= \inf\{t > \kappa_{m_{n-1}} \mid |\zeta_{m_{n-1}}(t)| \geq \lambda\} \\ &= \inf\{t > \kappa_{m_{n-1}} \mid |\zeta_{m_{n-1}}(t) - \zeta_{m_{n-1}}(\kappa_{m_{n-1}})| \geq 2\lambda\} \end{aligned}$$

and, for $\kappa_{m_{n-1}} < t \leq \kappa_{m_n}$,

$$\begin{aligned} &|\zeta_{m_{n-1}}(t) - \zeta_{m_{n-1}}(\kappa_{m_{n-1}})| \\ &\leq \left((2\lambda - h) \partial_+ j_\alpha(0) (m_n - m_{n-1} + \lceil \frac{\alpha \|g'\|_\infty}{2\lambda - h} \rceil) \right) \\ &\quad \cdot (t - \kappa_{m_{n-1}}) \\ &\leq ((2\lambda - h) \partial_+ j_\alpha(0) (m_n - m_{n-1} + c_g)) (t - \kappa_{m_{n-1}}) \end{aligned}$$

Hence,

$$\kappa_{m_n} \geq \kappa_{m_{n-1}} + \frac{2\lambda}{(2\lambda - h) \partial_+ j_\alpha(0)} (m_n - m_{n-1} + c_g)^{-1}$$

and the above estimate gives

$$\kappa_{m_n} \geq \kappa_{m_{n-1}} + \frac{2\lambda}{(2\lambda - h) \partial_+ j_\alpha(0)} (n c_g)^{-1}.$$

As $\kappa_{m_\ell-1} \geq \kappa_{m_{\ell-1}}$ for all $1 < \ell \leq n$, we can conclude that

$$\kappa_{m_{n+1}} \geq \kappa_1 + \frac{2\lambda}{(2\lambda - h) \partial_+ j_\alpha(0) c_g} \sum_{\ell=1}^n \frac{1}{\ell}.$$

Consequently, the subsequence $(\kappa_{m_n})_{n \in \mathbb{N}}$ and, therefore, also the full sequence $(\kappa_n)_{n \in \mathbb{N}_0}$ diverge to infinity for $n \rightarrow \infty$.

Now, assume that $g \notin L^\infty(\mathbb{R})$. Because g has an essentially bounded derivative and $|g(s)| < \lambda$ for all $s \leq \tau_0$, we can find for any $t \in \mathbb{R}$ a function $\tilde{g} \in \mathcal{C}_{\lambda,\tau_0}^{0,1} \cap L^\infty(\mathbb{R})$ with $g(s) = \tilde{g}(s)$ for all $s \leq t$. This also implies that

$$\mathcal{M}_\lambda^{h,j_\alpha} g(s) = \mathcal{M}_\lambda^{h,j_\alpha} \tilde{g}(s) \quad \forall s \leq t$$

and, hence, the statement is also true for $g \notin L^\infty(\mathbb{R})$. $\square$

As opposed to $\mathcal{M}_H$, we now show that $\mathcal{M}_\lambda^{h,j_\alpha}$ satisfies $\mathcal{M}_\lambda^{h,j_\alpha} g(t) \in [-\lambda, \lambda]$ and $|\mathcal{M}_\lambda^{h,j_\alpha} g(\kappa_n)| = \lambda$ if $\kappa_n < \infty$ for all $g \in \mathcal{C}_{\lambda,\tau_0}^{0,1}$ without additional restrictions on the first and second derivative of g . However, making the folding points depend on j_α with positive transient $\alpha > 0$ has a disadvantageous side effect, namely that additional folds can be caused by the transient of prior folds in the opposite direction. This is illustrated in Fig. 5, where we use $j_\alpha(t) = \frac{t}{\alpha}$ for $t \in [0, \alpha]$ and observe that the folds do not end. To prevent this, we have to make an additional assumption on the second derivative of g , which allows us to show that folds in different directions are separated by α .

Theorem 3. *Let $g \in \mathcal{C}_{\lambda,\tau_0}^{0,1}$. Then, $\mathcal{M}_\lambda^{h,j_\alpha} g(t) \in [-\lambda, \lambda]$ for all $t \in \mathbb{R}$ and $|\mathcal{M}_\lambda^{h,j_\alpha} g(\kappa_n)| = \lambda$ if $\kappa_n < \infty$. If, in addition, $g \in \mathcal{C}_{\lambda,\tau_0}^{1,1}$ with $\|g''\|_\infty \leq \frac{2h}{\alpha^2}$, then, for $n \in \mathbb{N}$ and $\kappa_{n+1} < \infty$ we have*

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

Proof. Since $\mathcal{M}_\lambda^{h,j_\alpha} g(t) = \zeta_{n-1}(t)$ for all $t \leq \kappa_n$, the definition in (3) gives $|\mathcal{M}_\lambda^{h,j_\alpha} g(\kappa_n)| = |\zeta_{n-1}(\kappa_n)| = \lambda$. Now assume that there is $t \in (\kappa_n, \kappa_{n+1})$ so that $|\mathcal{M}_\lambda^{h,j_\alpha} g(t)| > \lambda$. This in turn implies that $|\zeta_n(t)| > \lambda$ and, hence, $t > \kappa_{n+1}$ in contradiction to the assumption. To prove the separation property, as in the proof of Proposition 2 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.,

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

so that $\mathcal{M}_\lambda^{h,j_\alpha} g(\kappa_{m_n}) = \pm \lambda = -\mathcal{M}_\lambda^{h,j_\alpha} g(\kappa_{m_{n+1}})$. Without loss of generality let $\mathcal{M}_\lambda^{h,j_\alpha} g(\kappa_{m_1}) = \lambda$. Then, we obtain $\mathcal{M}_\lambda^{h,j_\alpha} g(\kappa_{m_2-1}) = \lambda$ and $\mathcal{M}_\lambda^{h,j_\alpha} g(\kappa_{m_2}) = -\lambda$. Additionally, the monotonicity of $\partial_+ j_\alpha$ due to Assumption (C₂) gives $g'(\kappa_{m_2-1}) \geq \partial_+(\sum_{\kappa \in K \setminus \{\kappa_{m_2-1}\}} (2\lambda - h) j_\alpha(t - \kappa))$ for $t > \kappa_{m_2-1}$ and with $K = \{\kappa_k \in (\kappa_{m_2-1} - \alpha, \kappa_{m_2-1}]\}$. Therefore, $g'(\kappa_{m_2-1})(t - \kappa_{m_2-1}) \geq \sum_{\kappa \in K \setminus \{\kappa_{m_2-1}\}} (2\lambda -$

$h)(j_\alpha(t - \kappa) - j_\alpha(\kappa_{m_2-1} - \kappa))$ and using (C₂) we can estimate $\zeta_{m_2-1}(t)$ as

$$\begin{aligned} \zeta_{m_2-1}(t) &= \zeta_{m_2-1}(\kappa_{m_2-1}) - \zeta_{m_2-1}(\kappa_{m_2-1}) + \zeta_{m_2-1}(t) \\ &= \lambda - g(\kappa_{m_2-1}) + g(t) - (2\lambda - h) \sum_{n=1}^{m_2-1} \text{sgn}(\zeta_{n-1}(\kappa_n)) \\ &\quad \cdot (j_\alpha(t - \kappa_n) - j_\alpha(\kappa_{m_2-1} - \kappa_n)) \\ &\geq \lambda + g'(\kappa_{m_2-1})(t - \kappa_{m_2-1}) - \frac{\|g''\|_\infty}{2} (t - \kappa_{m_2-1})^2 \\ &\quad - (2\lambda - h) \sum_{\kappa \in K} (j_\alpha(t - \kappa) - j_\alpha(\kappa_{m_2-1} - \kappa)) \\ &\geq \lambda - \frac{h}{\alpha^2} (t - \kappa_{m_2-1})^2 - (2\lambda - h) \cdot j_\alpha(t - \kappa_{m_2-1}) \\ &\geq -\lambda + h - \frac{h}{\alpha^2} (t - \kappa_{m_2-1})^2. \end{aligned}$$

With this we can conclude that

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

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

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

which completes the proof. $\square$

The separation in Theorem 3 can be used to get the return property for sufficiently large t also for the operator $\mathcal{M}_\lambda^{h,j_\alpha}$.

Proposition 3. *Let $0 < h < \lambda$ and $g \in \mathcal{C}_{\lambda,\tau_0}^{1,1} \cap \mathcal{B}_\rho^{\lambda-h}$ with $\rho > \tau_0$ and $\|g''\|_\infty \leq \frac{2h}{\alpha^2}$. Then, $\mathcal{M}_\lambda^{h,j_\alpha} g(t) = g(t)$ for all $t \geq \rho + 2\alpha$.*

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

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

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

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

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

from which follows that $\mathcal{M}_\lambda^{h,j_\alpha} g(t) = g(t)$ for all $t \geq \rho$.

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

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

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

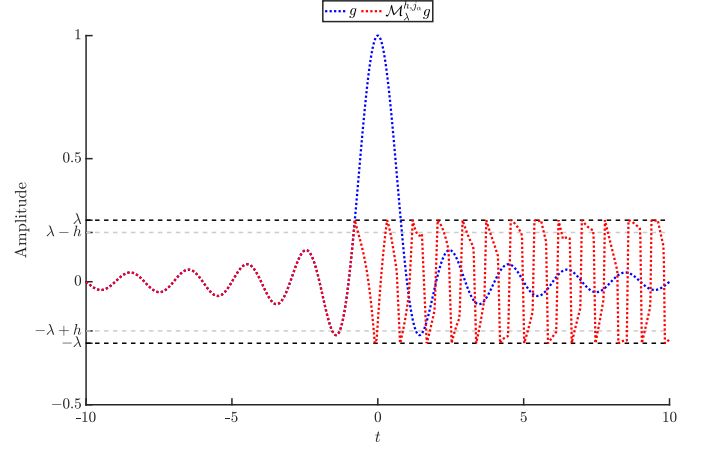


Fig. 5. Example of an endless folding in $\mathcal{M}_\lambda^{h,j_\alpha} g(t)$ using $g(t) = \text{sinc}(\pi t)$, where $j_\alpha(t) = \frac{t}{\alpha}$ for $t \in [0, \alpha]$, $\lambda = 0.25$, $h = 0.05$ and $\alpha = 0.75$.

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

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

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

Case 4: Let $\kappa_n + \alpha > \max\{\rho, \kappa_{n+1}\}$, $\kappa_{n+1} + \alpha > \kappa_{n+2}$. As above, without loss of generality let $\mathcal{M}_\lambda^{h,j_\alpha} g(\kappa_n) = \lambda$. Then, the separation property in Theorem 3 implies that $\lambda = \mathcal{M}_\lambda^{h,j_\alpha} g(\kappa_{n+1}) = \mathcal{M}_\lambda^{h,j_\alpha} g(\kappa_{n+2})$ and $\mathcal{M}_\lambda^{h,j_\alpha} g(\kappa_{n+2} + \alpha) \leq \zeta_{n+2}(\kappa_{n+2} + \alpha)$. This gives us that

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

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

We now show that the behaviour in Fig. 5 is not due to numerical inaccuracies but can be proven to occur. As the construction depends on the specific form of j_α , we restrict the following discussion to the folding function from [37], i.e.,

$$j_\alpha(t) = \begin{cases} \frac{t}{\alpha} & \text{if } t \in [0, \alpha] \\ 0 & \text{otherwise.} \end{cases}$$

Endless folding may happen if folds in one direction are triggered while folds in the opposite direction are still incomplete. More precisely, we make the following assumption.

Assumption 1. *There are $c_0 \in \mathbb{N}$, $c_0 < c_1 \in \mathbb{N}$, $c_2 \in \{0, 1\}$, $c_3 = c_1 + c_2$, $c_4 \in \{0, 1\}$ as well as $t_1 \in \mathbb{R}$, $t_2 \in [t_1, t_1 + \alpha)$, $t_3 \in (\max\{t_2, t_1 + \frac{\alpha}{2}\}, t_1 + \alpha)$ and $t_4 \in [t_3, t_3 + \alpha)$ so that*

$$\begin{aligned} \zeta_n(t) &= g(t) + \sigma r(t) - (2\lambda - h) \sigma (c_0 - c_1 j_\alpha(t - t_1) \\ &\quad - c_2 j_\alpha(t - t_2) + c_3 j_\alpha(t - t_3) + c_4 j_\alpha(t - t_4)) \end{aligned}$$

for all $t \geq t_3$ and

$$\sigma \cdot \zeta_n(t) < \lambda \quad \forall t > t_4,$$

where $\sigma \in \{\pm 1\}$ and $r: \mathbb{R} \rightarrow \mathbb{R}$ is a function with $r(t) \geq 0$ for all $t \in \mathbb{R}$, $r(t) = 0$ for all $t \geq t_1 + \alpha$ and $\|\partial_+ r\|_\infty \leq \frac{2\lambda - h}{\alpha}$.

Using Assumption 1 we can guarantee the continuation of foldings as follows.

Proposition 4. Let $h < \frac{\lambda}{2}$ and $j_\alpha(t) = \frac{t}{\alpha}$ for $t \in [0, \alpha]$. Moreover, let $g \in \mathcal{C}_{\lambda, \tau_0}^{1,1}$ and $n \geq 5$ such that

- 1) Assumption 1 is satisfied with $\sigma \in \{\pm 1\}$ and $t_i = \kappa_{m_i}$ for $m_1 \leq m_2 \leq m_3 \leq m_4 = n$,
- 2) for all $t > \kappa_n - 2\alpha$ there exists at most one $s \in [t, t + 2\alpha]$ with $g'(s) = 0$,
- 3) $|g(t)| < h$ for all $t \geq t_3$,
- 4) $|g(t) - g(s)| < h$ for all $t, s \geq t_3$,
- 5) $|g'(t)| < \frac{2\lambda - h}{\alpha}$ for all $t \geq t_1$.

Then, for all $t \in \mathbb{R}$ there exists $q \in \mathbb{N}$ so that $t < \kappa_q < \infty$.

Proof. Without loss of generality, assume that $\sigma = -1$. Then, for $t \in (t_3, t_1 + \alpha]$ we get

$$\begin{aligned} \zeta_n(t) &= \zeta_n(t) - \zeta_n(t_3) - \lambda \\ &= -\lambda + g(t) - g(t_3) + r(t_3) - r(t) \\ &\quad - (2\lambda - h)(c_1 j_\alpha(t - t_1) + c_2 j_\alpha(t - t_2)) \\ &\quad - c_3 j_\alpha(t - t_3) - c_4 j_\alpha(t - t_4) - c_1 j_\alpha(t_3 - t_1) \\ &\quad - c_2 j_\alpha(t_3 - t_2)) \\ &< -\lambda + h - r(t) + r(t_3) + (2\lambda - h)c_4 j_\alpha(t - t_4) \\ &\leq -\lambda + h + 2 \frac{2\lambda - h}{\alpha} \left(t_1 + \alpha - t_1 - \frac{\alpha}{2} \right) = \lambda. \end{aligned}$$

Therefore, we have $\kappa_{n+1} > t_1 + \alpha$. In the same way we obtain $\zeta_n(t) < \lambda$ for $t \in (t_3, t_3 + \frac{\alpha}{2}]$ and, hence, $\kappa_{n+1} > t_3 + \frac{\alpha}{2}$. On the other hand, we have

$$\begin{aligned} \zeta_n(t_3 + \alpha) &= g(t_3 + \alpha) + (2\lambda - h)(c_0 + c_4 j_\alpha(t_3 + \alpha - t_4)) \\ &\geq -h + c_0(2\lambda - h) > \lambda. \end{aligned}$$

Thus, we conclude that $\kappa_{n+1} \in (\max\{t_1 + \alpha, t_3 + \frac{\alpha}{2}\}, t_3 + \alpha)$ and $\text{sgn}(\zeta_n(\kappa_{n+1})) = 1 = -\sigma = -\text{sgn}(\zeta_n(\kappa_n))$.

Let us first assume that $\kappa_{n+1} \in [t_2 + \alpha, t_3 + \alpha)$. Then, we have $\kappa_{n+1} = \dots = \kappa_{n+m}$ with

$$\begin{aligned} m &= \left\lfloor \frac{\partial_+ \zeta_n(\kappa_{n+1})}{\frac{2\lambda - h}{\alpha}} + 1 \right\rfloor = \left\lfloor \frac{g'(\kappa_{n+1})}{\frac{2\lambda - h}{\alpha}} + c_3 + c_4 + 1 \right\rfloor \\ &= \begin{cases} c_3 + c_4 + 1 & \text{if } g'(\kappa_{n+1}) \geq 0 \\ c_3 + c_4 & \text{otherwise.} \end{cases} \end{aligned}$$

For $m = c_3 + c_4 + 1$ setting $\tilde{t}_3 = \tilde{t}_4 = \kappa_{n+m}$, $\tilde{t}_1 = t_3$, $\tilde{t}_2 = t_4$, $\tilde{c}_0 = c_1 + c_2 - c_0$, $\tilde{c}_1 = c_3$, $\tilde{c}_2 = c_4$, $\tilde{c}_3 = c_3 + c_4$, $\tilde{c}_4 = 1$ as well as $\tilde{r}(t) = 0$ gives

$$\begin{aligned} \zeta_{n+m}(t) &= g(t) + \tilde{r}(t) - (2\lambda - h)(\tilde{c}_0 - \tilde{c}_1 j_\alpha(t - \tilde{t}_1) \\ &\quad - \tilde{c}_2 j_\alpha(t - \tilde{t}_2) + \tilde{c}_3 j_\alpha(t - \tilde{t}_3) \\ &\quad + \tilde{c}_4 j_\alpha(t - \tilde{t}_4)) \end{aligned} \quad (5)$$

for all $t \geq \tilde{t}_3$ and $\zeta_{n+m}(t) < \lambda$ for all $t > \tilde{t}_4$ as $|g'(t)| < \frac{2\lambda - h}{\alpha}$ for all $t \geq t_1$. Since $\text{sgn}(\zeta_n(\kappa_n)) = -\text{sgn}(\zeta_n(\kappa_{n+1}))$, we have $\tilde{t}_3 = \kappa_{n+m} = \kappa_{n+1} > \kappa_n = t_4 = t_2$. Moreover, $\tilde{t}_3 = \kappa_{n+m} = \kappa_{n+1} \geq t_3 + \frac{\alpha}{2} = \tilde{t}_1 + \frac{\alpha}{2}$. For $m = c_3 + c_4$ we distinguish two cases. If there exists no $t \in (\kappa_{n+m}, t_3 + \alpha)$ with $\zeta_{n+m}(t) = \lambda$, we get the same equation for ζ_{n+m} as in

(5) with $\tilde{c}_4 = 0$. On the other hand, if $t \in (\kappa_{n+m}, t_3 + \alpha)$ exists with $\zeta_{n+m}(t) = \lambda$, then κ_{n+m+1} must be the first such point and setting $\tilde{t}_3 = \kappa_{n+m}$, $\tilde{t}_4 = \kappa_{n+m+1}$, $\tilde{t}_1 = t_3$, $\tilde{t}_2 = t_4$, $\tilde{c}_0 = c_1 + c_2 - c_0$, $\tilde{c}_1 = c_3$, $\tilde{c}_2 = c_4$, $\tilde{c}_3 = c_3 + c_4$, $\tilde{c}_4 = 1$ as well as $\tilde{r}(t) = 0$ gives

$$\begin{aligned} \zeta_{n+m+1}(t) &= g(t) + \tilde{r}(t) - (2\lambda - h)(\tilde{c}_0 - \tilde{c}_1 j_\alpha(t - \tilde{t}_1) \\ &\quad - \tilde{c}_2 j_\alpha(t - \tilde{t}_2) + \tilde{c}_3 j_\alpha(t - \tilde{t}_3) \\ &\quad + \tilde{c}_4 j_\alpha(t - \tilde{t}_4)) \end{aligned} \quad (6)$$

for all $t \geq \tilde{t}_3$ and $\zeta_{n+m+1}(t) < \lambda$ for all $t > \tilde{t}_4$. As before, we obtain $\tilde{t}_3 > \tilde{t}_2$ and $\tilde{t}_3 > \tilde{t}_1 + \frac{\alpha}{2}$.

Let us now assume that $\kappa_{n+1} \in (t_3 + \frac{\alpha}{2}, t_2 + \alpha)$, which can only happen if $c_2 = 1$ and $t_1 < t_2$. As $|g'(t)| < \frac{2\lambda - h}{\alpha}$ for all $t \geq t_1$ and $j_\alpha(s) = \frac{s}{\alpha}$ for $s \in [0, \alpha]$, we get $g'(t_1) < 0$ and $g'(t_2) \geq 0$ so that there exists $\tilde{t} \in (t_1, t_2]$ with $g'(\tilde{t}) = 0$. Using 2) yields $g'(t) > 0$ for $t \in (\tilde{t}, \tilde{t} + 2\alpha)$ and, hence, also for $t \in (t_2, t_2 + \alpha)$ so that, in particular, $g'(\kappa_{n+1}) \geq 0$ and ζ_n is monotonously increasing on $(t_1 + \alpha, t_2 + \alpha)$. Therefore, we have $\kappa_{n+1} = \dots = \kappa_{n+m}$ with

$$\begin{aligned} m &= \left\lfloor \frac{\partial_+ \zeta_n(\kappa_{n+1})}{\frac{2\lambda - h}{\alpha}} + 1 \right\rfloor = \left\lfloor \frac{g'(\kappa_{n+1})}{\frac{2\lambda - h}{\alpha}} - c_2 + c_3 + c_4 + 1 \right\rfloor \\ &= c_3 + c_4. \end{aligned}$$

If there exists no $t \in (\kappa_{n+m}, t_3 + \alpha)$ with $\zeta_{n+m}(t) = \lambda$, we get the same equation for ζ_{n+m} as in (5) with $\tilde{c}_4 = 0$ and $\tilde{r}(t) = c_2(2\lambda - h)(1 - j_\alpha(t - t_2))$. If there does exist $t \in (\kappa_{n+m}, t_3 + \alpha)$ with $\zeta_{n+m}(t) = \lambda$, then κ_{n+m+1} must be the first such point and we get the same equation for ζ_{n+m+1} as in (6) with $\tilde{r}(t) = c_2(2\lambda - h)(1 - j_\alpha(t - t_2))$.

All in all, Assumption 1 is satisfied for $\zeta_{n+m+\ell}$ with opposite sign $-\sigma$ and folding points $\kappa_{m_3}, \kappa_n, \kappa_{n+m}, \kappa_{n+m+\ell}$, where either $\ell = 0$ or $\ell = 1$. In particular, we have $\kappa_{n+m} \in [\kappa_n + \frac{\alpha}{2}, \kappa_n + \alpha)$. Repeating this argument shows that for all $t \in \mathbb{R}$ there exists some $q \in \mathbb{N}$ with $t < \kappa_q < \infty$. $\square$

We wish to stress that endless folding does not occur only in a niche folding regime. In fact, for any choice of parameters $\lambda > 0$, $h < \frac{\lambda}{2}$ and $\alpha > 0$ one can find functions satisfying the assumptions of Proposition 4. For example, for $\lambda = 1$, $h = 0.35$ and $\alpha = \frac{\pi}{2}$ the assumptions are satisfied for the benign function $g(t) = \text{sinc}(t)$.

C. Transient dependent folding points and reset time

The definition of the generalized modulo operator requires the rather restrictive condition (C₂) on the folding function j_α with $\alpha > 0$, which, in particular, implies that j_α cannot be differentiable at 0, whereas it may be arbitrarily smooth at α . We now modify the modulo operator $\mathcal{M}_\lambda^{h, j_\alpha}$ to allow for more general folding functions not necessarily satisfying (C₂). To this end, we introduce a minimal reset time $\sigma > 0$ between two consecutive folding points.

Definition 4. Let $\lambda > 0$, $h < 2\lambda$, $\sigma > 0$ and let j_α be a folding function with transient $\alpha \geq 0$. For a function $g \in \mathcal{C}_{\lambda, \tau_0}^{0,1}$ with $\tau_0 \in \mathbb{R}$ we define the two sequences $(\nu_n)_{n \in \mathbb{N}_0}$ and $(\xi_n)_{n \in \mathbb{N}_0}$ via

$$\begin{aligned} \nu_{n+1} &= \inf\{t > \nu_n + \sigma \mid |j_\alpha(t)| \geq \lambda\}, \\ \xi_{n+1} &= \xi_n - (2\lambda - h) \text{sgn}(\xi_n(\nu_{n+1})) j_\alpha(\cdot - \nu_{n+1}) \end{aligned} \quad (7)$$

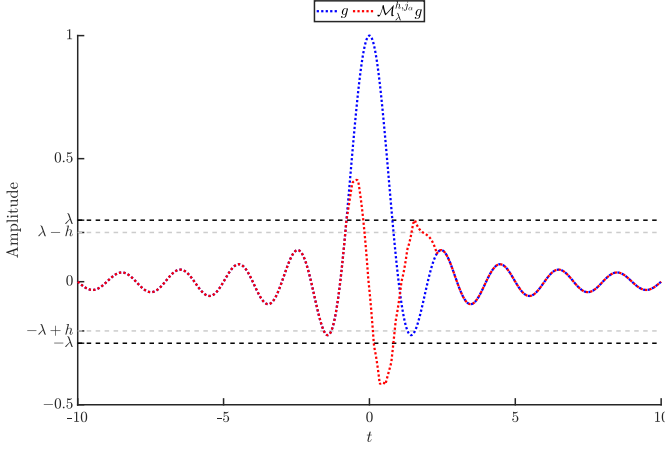


Fig. 6. Illustration of $\mathcal{M}_{\lambda, \sigma}^{h, j_\alpha} g(t)$ using $g(t) = \text{sinc}(\pi t)$, where we choose $j_\alpha(t) = \frac{t}{\alpha}$ for $t \in [0, \alpha]$, $\lambda = 0.25$, $h = 0.05$, $\alpha = 0.75$ and $\sigma = 0.2$.

and $\xi_{n+1} = \xi_n$ if $\nu_{n+1} = \infty$, where $\nu_0 = \tau_0 - \sigma$ and $\xi_0 = g$. Then, the output of the delayed generalized modulo operator $\mathcal{M}_{\lambda, \sigma}^{h, j_\alpha}$ is defined as

$$\mathcal{M}_{\lambda, \sigma}^{h, j_\alpha} g(t) = \lim_{n \rightarrow \infty} \xi_n(t). \quad (8)$$

As we now enforce a minimum spacing of $\sigma > 0$ between folding points, for each $t \in \mathbb{R}$ we can find $n \in \mathbb{N}$ so that $t \leq \tau_0 + n \cdot \sigma$ and $t \leq \nu_{n+1}$. This implies $\mathcal{M}_{\lambda, \sigma}^{h, j_\alpha} g(t) = \xi_{n+1}(t)$ and, hence, the well-definedness of $\mathcal{M}_{\lambda, \sigma}^{h, j_\alpha}$ without additional assumptions. We now collect further properties of $\mathcal{M}_{\lambda, \sigma}^{h, j_\alpha}$.

Proposition 5. Let $g \in \mathcal{C}_{\lambda, \tau_0}^{0,1}$. Then, $|\mathcal{M}_{\lambda, \sigma}^{h, j_\alpha} g(\nu_n)| \geq \lambda$ for all $\nu_n < \infty$ and $|\mathcal{M}_{\lambda, \sigma}^{h, j_\alpha} g(t)| < \lambda$ for all $t \in [\nu_n + \sigma, \nu_{n+1}]$.

Proof. Assume that $\nu_n < \infty$ and $|\mathcal{M}_{\lambda, \sigma}^{h, j_\alpha} g(\nu_n)| < \lambda$. By definition this also implies $|\xi_{n-1}(\nu_n)| < \lambda$ and, therefore, $\nu_n \notin \{t > \nu_{n-1} + \sigma \mid |\xi_{n-1}(t)| \geq \lambda\}$ in contradiction to the definition of ν_n . Now, assume that there exists $s \in [\nu_n + \sigma, \nu_{n+1}]$ with $|\mathcal{M}_{\lambda, \sigma}^{h, j_\alpha} g(s)| \geq \lambda$. This also implies $|\xi_n(s)| \geq \lambda$ and, hence, $s \in \{t > \nu_n + \sigma \mid |\xi_n(t)| \geq \lambda\}$. But this gives $s \geq \nu_{n+1}$ in contradiction to the assumption. $\square$

While well-definedness of $\mathcal{M}_{\lambda, \sigma}^{h, j_\alpha}$ does not require the conditions (C₁) and (C₂), to guarantee a bounded range and the return property, we need additional assumptions that are weaker than (C₁) and (C₂), but result in a tension between modelling flexibility and theoretical guarantees.

Theorem 4. Let $0 < \sigma \leq \alpha$ and j_α satisfy

$$\forall t \in [0, \alpha]: j_\alpha(t) \leq 1. \quad (\text{C}_1^*)$$

and

$$\forall s \leq t \in [\sigma, \alpha]: 0 \leq \partial_+ j_\alpha(t) \leq \partial_+ j_\alpha(s) \quad (\text{C}_2^*)$$

Moreover, let $g \in \mathcal{C}_{\lambda, \tau_0}^{1,1}$ satisfy $\|g'\|_\infty \leq (2\lambda - h) \cdot \frac{j_\alpha(\sigma)}{\sigma}$ and $\|g''\|_\infty \leq \frac{2h}{\alpha^2}$. Then, we have $\mathcal{M}_{\lambda, \sigma}^{h, j_\alpha} g(t) \in [-\lambda - r, \lambda + r]$ with

$$r = \sup_{t \in (0, \sigma)} (t \cdot \|g'\|_\infty - (2\lambda - h) \cdot j_\alpha(t)) \geq 0,$$

and

$$\mathcal{M}_{\lambda, \sigma}^{h, j_\alpha} g(\nu_n) \cdot \mathcal{M}_{\lambda, \sigma}^{h, j_\alpha} g(\nu_{n+1}) < 0 \implies \nu_{n+1} - \nu_n \geq \alpha$$

as well as $\mathcal{M}_{\lambda, \sigma}^{h, j_\alpha} g(\nu_n) = \pm\lambda$ for all $n \in \mathbb{N}$ with $\nu_n < \infty$. If, in addition, $g \in \mathcal{B}_\rho^{\lambda-h}$ with $\rho > \tau_0$, we further obtain $\mathcal{M}_{\lambda, \sigma}^{h, j_\alpha} g(t) = g(t)$ for all $t \geq \rho + 2\alpha$.

Proof. By definition, we have $\mathcal{M}_{\lambda, \sigma}^{h, j_\alpha} g(t) = g(t) \in (-\lambda, \lambda)$ for all $t < \nu_1$ and $g(\nu_1) = \pm\lambda$ if $\nu_1 < \infty$ as $\nu_1 = \kappa_1$. Without loss of generality let $g(\nu_1) = \lambda$ so that $g'(\nu_1) \geq 0$. For $t \in (\nu_1, \nu_1 + \sigma)$ we obtain

$$\begin{aligned} \mathcal{M}_{\lambda, \sigma}^{h, j_\alpha} g(t) &= g(t) - (2\lambda - h) \cdot j_\alpha(t - \nu_1) \\ &= \lambda - g(\nu_1) + g(t) - (2\lambda - h) \cdot j_\alpha(t - \nu_1) \\ &\leq \lambda + (t - \nu_1) \cdot \|g'\|_\infty - (2\lambda - h) \cdot j_\alpha(t - \nu_1) \\ &\leq \lambda + r \end{aligned}$$

and $g'(\nu_1) \geq 0$ in combination with (C₁^{*}) gives

$$\begin{aligned} \mathcal{M}_{\lambda, \sigma}^{h, j_\alpha} g(t) &= \lambda - g(\nu_1) + g(t) - (2\lambda - h) \cdot j_\alpha(t - \nu_1) \\ &\geq \lambda - \frac{\|g''\|_\infty}{2} (t - \nu_1)^2 - (2\lambda - h) \cdot j_\alpha(t - \nu_1) \\ &\geq -\lambda + h - \frac{h}{\alpha^2} (t - \nu_1)^2 > -\lambda. \end{aligned}$$

Moreover,

$$\mathcal{M}_{\lambda, \sigma}^{h, j_\alpha} g(\nu_1 + \sigma) \leq \lambda + \sigma \cdot \|g'\|_\infty - (2\lambda - h) \cdot j_\alpha(\sigma) \leq \lambda$$

and

$$\mathcal{M}_{\lambda, \sigma}^{h, j_\alpha} g(\nu_1 + \sigma) \geq \lambda - \frac{\sigma^2}{2} \|g''\|_\infty - (2\lambda - h) \cdot j_\alpha(\sigma) \geq -\lambda.$$

We already know that $|\mathcal{M}_{\lambda, \sigma}^{h, j_\alpha} g(t)| < \lambda$ for all $t \in [\nu_1 + \sigma, \nu_2]$ and, hence, $\mathcal{M}_{\lambda, \sigma}^{h, j_\alpha} g(t) \in [-\lambda, \lambda + r]$ for all $t < \nu_2$. Moreover, $\mathcal{M}_{\lambda, \sigma}^{h, j_\alpha} g(\nu_2) = \pm\lambda$ if $\nu_2 < \infty$ and $\mathcal{M}_{\lambda, \sigma}^{h, j_\alpha} g(\nu_2) = -\lambda$ can only happen if $\nu_2 \geq \nu_1 + \alpha$.

Now let $2 \leq \ell \leq n \in \mathbb{N}$ such that $\nu_{\ell-1} + \alpha \leq \nu_\ell \leq \nu_n < \infty$ and $\mathcal{M}_{\lambda, \sigma}^{h, j_\alpha} g(\nu_\ell) = \dots = \mathcal{M}_{\lambda, \sigma}^{h, j_\alpha} g(\nu_n) = \pm\lambda$. Without loss of generality assume that $\mathcal{M}_{\lambda, \sigma}^{h, j_\alpha} g(\nu_n) = \xi_{n-1}(\nu_n) = \lambda$. Then, for $t \in (\nu_n, \nu_n + \sigma)$ we get, as before and using the monotonicity of j_α on $[\sigma, \alpha]$ due to Assumption (C₂^{*}),

$$\begin{aligned} \mathcal{M}_{\lambda, \sigma}^{h, j_\alpha} g(t) &= \xi_{n-1}(t) - (2\lambda - h) \cdot j_\alpha(t - \nu_n) \\ &= \lambda - \xi_{n-1}(\nu_n) + \xi_{n-1}(t) \\ &\quad - (2\lambda - h) \cdot j_\alpha(t - \nu_n) \\ &\leq \lambda - g(\nu_n) + g(t) - (2\lambda - h) \cdot j_\alpha(t - \nu_n) \\ &\leq \lambda + r \end{aligned}$$

and $\mathcal{M}_{\lambda, \sigma}^{h, j_\alpha} g(\nu_n + \sigma) \leq \lambda$. Additionally, the monotonicity of $\partial_+ j_\alpha$ on $[\sigma, \alpha]$ due to Assumption (C₂^{*}) gives $g'(\nu_n) \geq$

TABLE I

CONDITIONS FOR THE MODULO OPERATORS TO GUARANTEE CERTAIN PROPERTIES FOR $g \in \mathcal{C}_{\lambda, \tau_0}^{0,1}$ ("–", IF NO FURTHER CONDITIONS ARE NEEDED).

		$\mathcal{M}_H g$	$\mathcal{M}_\lambda^{h, j_\alpha} g$	$\mathcal{M}_{\lambda, \sigma}^{h, j_\alpha} g$
well-definedness	condition	–	(C ₂)	–
		(C ₁)	(C ₂)	(C ₁ [*]); (C ₂ [*])
folding at $\pm \lambda$ & bounded range	condition	$g \in \mathcal{C}_{\lambda, \tau_0}^{1,1}; \ g''\ _\infty \leq \frac{2h}{\alpha^2};$ $\ g'\ _\infty \leq \frac{2\lambda-h}{\alpha}$		$g \in \mathcal{C}_{\lambda, \tau_0}^{1,1}; \ g''\ _\infty \leq \frac{2h}{\alpha^2};$ $\ g'\ _\infty \leq \frac{(2\lambda-h) \cdot j_\alpha(\sigma)}{\sigma}$
	range	$[-\lambda, \lambda]$	$[-\lambda, \lambda]$	$[-\lambda - r, \lambda + r]$
separation of folds	condition	–	(C ₂) $g \in \mathcal{C}_{\lambda, \tau_0}^{1,1}; \ g''\ _\infty \leq \frac{2h}{\alpha^2}$	–
	amount	$\frac{\min\{h, 2\lambda-h\}}{\ g'\ _\infty}$	α for folds in different directions	σ
return property	condition	$g \in \mathcal{B}_\rho^{\lambda-h}$	(C ₂) $g \in \mathcal{C}_{\lambda, \tau_0}^{1,1} \cap \mathcal{B}_\rho^{\lambda-h}; \ g''\ _\infty \leq \frac{2h}{\alpha^2}$	(C ₁ [*]); (C ₂ [*]) $g \in \mathcal{C}_{\lambda, \tau_0}^{1,1} \cap \mathcal{B}_\rho^{\lambda-h}; \ g''\ _\infty \leq \frac{2h}{\alpha^2};$ $\ g'\ _\infty \leq \frac{(2\lambda-h) \cdot j_\alpha(\sigma)}{\sigma}$
	interval	$[\rho + \alpha, \infty)$	$[\rho + 2\alpha, \infty)$	$[\rho + 2\alpha, \infty)$

$\partial_+ \left(\sum_{\nu \in K \setminus \{\nu_n\}} (2\lambda - h) \cdot j_\alpha(t - \nu) \right)$ for $t > \nu_n$ and with $K = \{\nu_m \in (\nu_n - \alpha, \nu_n]\}$. With this we get for $t > \nu_n$

$$\begin{aligned} \xi_n(t) &\geq \xi_n(\nu_n) + g'(\nu_n)(t - \nu_n) - \frac{\|g''\|_\infty}{2}(t - \nu_n)^2 \\ &\quad - (2\lambda - h) \sum_{\nu \in K} (j_\alpha(t - \nu) - j_\alpha(\nu_n - \nu)) \\ &\geq \lambda - \frac{h}{\alpha^2}(t - \nu_n)^2 - (2\lambda - h) \cdot j_\alpha(t - \nu_n) \\ &\geq -\lambda + h - \frac{h}{\alpha^2}(t - \nu_n)^2. \end{aligned}$$

Therefore, for $t \in (\nu_n, \nu_n + \sigma]$ we have $\xi_n(t) \geq -\lambda$ and $\xi_n(t) = -\lambda$ can only occur for $t \geq \nu_n + \alpha$. We already know that $|\mathcal{M}_{\lambda, \sigma}^{h, j_\alpha} g(t)| < \lambda$ for all $t \in [\nu_n + \sigma, \nu_{n+1})$ and, hence, $\mathcal{M}_{\lambda, \sigma}^{h, j_\alpha} g(t) \in [-\lambda, \lambda + r]$ for all $t < \nu_{n+1}$. Moreover, $\mathcal{M}_{\lambda, \sigma}^{h, j_\alpha} g(\nu_{n+1}) = \pm \lambda$ if $\nu_{n+1} < \infty$ and $\mathcal{M}_{\lambda, \sigma}^{h, j_\alpha} g(\nu_{n+1}) = -\lambda$ can only happen if $\nu_{n+1} \geq \nu_n + \alpha$.

In conclusion, we have $\mathcal{M}_{\lambda, \sigma}^{h, j_\alpha} g(t) \in [-\lambda - r, \lambda + r]$ for all $t \in \mathbb{R}$ and

$$\mathcal{M}_{\lambda, \sigma}^{h, j_\alpha} g(\nu_n) \cdot \mathcal{M}_{\lambda, \sigma}^{h, j_\alpha} g(\nu_{n+1}) < 0 \implies \nu_{n+1} - \nu_n \geq \alpha$$

as well as $\mathcal{M}_{\lambda, \sigma}^{h, j_\alpha} g(\nu_n) = \pm \lambda$ if $\nu_n < \infty$. With this, for $g \in \mathcal{B}_\rho^{\lambda-h}$ with $\rho > \tau_0$ finally follows that $\mathcal{M}_{\lambda, \sigma}^{h, j_\alpha} g(t) = g(t)$ for all $t \geq \rho + 2\alpha$ as in the proof of Proposition 3. $\square$

III. COMPARISON OF THE MODULO OPERATORS

As mentioned above, the generalized modulo encoder $\mathcal{M}_H$ and the generalized modulo operator $\mathcal{M}_\lambda^{h, j_\alpha}$ agree for $\alpha = 0$. We now state a condition under which this also holds for $\alpha > 0$ and for the delayed generalized modulo operator $\mathcal{M}_{\lambda, \sigma}^{h, j_\alpha}$.

Theorem 5. *Let $\alpha \geq \sigma > 0$, $0 < h < 2\lambda$ and j_α satisfy (C₁) and (C₂). Moreover, let $g \in \mathcal{C}_{\lambda, \tau_0}^{1,1}$ with $\|g'\|_\infty < \frac{2\lambda-h}{\alpha}$ and $\|g''\|_\infty < \frac{2h}{\alpha^2}$. Then, we have $\mathcal{M}_H g = \mathcal{M}_\lambda^{h, j_\alpha} g = \mathcal{M}_{\lambda, \sigma}^{h, j_\alpha} g$.*

Proof. We first show that $\mathcal{M}_H g = \mathcal{M}_\lambda^{h, j_\alpha} g$. By definition we have $\tau_1 = \kappa_1$ and, therefore, $\mathcal{M}_H g(t) = \mathcal{M}_\lambda^{h, j_\alpha} g(t)$ for

all $t \leq \tau_1 = \kappa_1$. Now assume that for some $p \in \mathbb{N}$ we have $\tau_n = \kappa_n < \infty$ for all $n \leq p$ and $\mathcal{M}_H g(t) = \mathcal{M}_\lambda^{h, j_\alpha} g(t)$ for all $t \leq \tau_p = \kappa_p$. From Theorem 2 follows that $\tau_p \geq \tau_{p-1} + \alpha$ and furthermore $\tau_p + \alpha \leq \tau_{p+1}$. The separation in Theorem 3 gives $\kappa_p + \alpha \leq \kappa_{p+1}$ if $\mathcal{M}_\lambda^{h, j_\alpha} g(\kappa_p) \cdot \mathcal{M}_\lambda^{h, j_\alpha} g(\kappa_{p+1}) < 0$. Therefore, we now assume that $\mathcal{M}_\lambda^{h, j_\alpha} g(\kappa_p) \cdot \mathcal{M}_\lambda^{h, j_\alpha} g(\kappa_{p+1}) > 0$, in which case the range property in Theorem 3 gives $\mathcal{M}_\lambda^{h, j_\alpha} g(\kappa_p) = \mathcal{M}_\lambda^{h, j_\alpha} g(\kappa_{p+1})$ and, thus, $\zeta_p(\kappa_p) = \zeta_p(\kappa_{p+1})$. Consequently, we obtain

$$\begin{aligned} \kappa_{p+1} &= \inf\{t > \kappa_p \mid |\zeta_p(t)| \geq \lambda\} \\ &= \inf\{t > \kappa_p \mid |\zeta_p(t) - \zeta_p(\kappa_p)| \geq 0\}, \end{aligned}$$

where

$$\begin{aligned} |\zeta_p(t) - \zeta_p(\kappa_p)| &\leq |g(t) - g(\kappa_p)| - (2\lambda - h) \cdot j_\alpha(t - \kappa_p) \\ &\leq \|g'\|_\infty(t - \kappa_p) - (2\lambda - h) \\ &\leq \frac{2\lambda-h}{\alpha} \cdot (t - \kappa_p) - (2\lambda - h) \end{aligned}$$

and, thus,

$$\begin{aligned} \kappa_{p+1} &\geq \inf\{t > \kappa_p \mid \frac{2\lambda-h}{\alpha} \cdot (t - \kappa_p) - (2\lambda - h) \geq 0\} \\ &= \kappa_p + \alpha \end{aligned}$$

Consequently, $\kappa_p + \alpha \leq \kappa_{p+1}$ holds in all possible cases and $\zeta_p(\kappa_p) = \mathcal{M}_\lambda^{h, j_\alpha} g(\kappa_p) = \mathcal{M}_H g(\kappa_p) = s_p \lambda$ gives

$$\begin{aligned} \tau_{p+1} &= \inf\{t > \tau_p \mid \mathcal{M}_\lambda(g(t) - g(\tau_p) + h s_p) = 0\} \\ &= \inf\{t > \tau_p \mid \mathcal{M}_\lambda(g(t) - g(\kappa_p) - (2\lambda - h) s_p) = 0\} \\ &= \inf\{t > \tau_p \mid \mathcal{M}_\lambda(g(t) - g(\kappa_p) \\ &\quad - (2\lambda - h) \operatorname{sgn}(\zeta_p(\kappa_p))) = 0\} \\ &= \inf\{t > \tau_p \mid \mathcal{M}_\lambda(\zeta_p(t) - \zeta_p(\kappa_p)) = 0\} \\ &= \inf\{t > \tau_p \mid \mathcal{M}_\lambda(\zeta_p(t) - \lambda \cdot \operatorname{sgn}(\zeta_p(\kappa_p))) = 0\} \\ &= \inf\{t > \tau_p \mid |\zeta_p(t)| = \lambda\} = \kappa_{p+1} \end{aligned}$$

so that $\mathcal{M}_H g(t) = \mathcal{M}_\lambda^{h, j_\alpha} g(t)$ for all $t \leq \tau_{p+1} = \kappa_{p+1}$. By induction we can conclude that $\mathcal{M}_H g = \mathcal{M}_\lambda^{h, j_\alpha} g$.

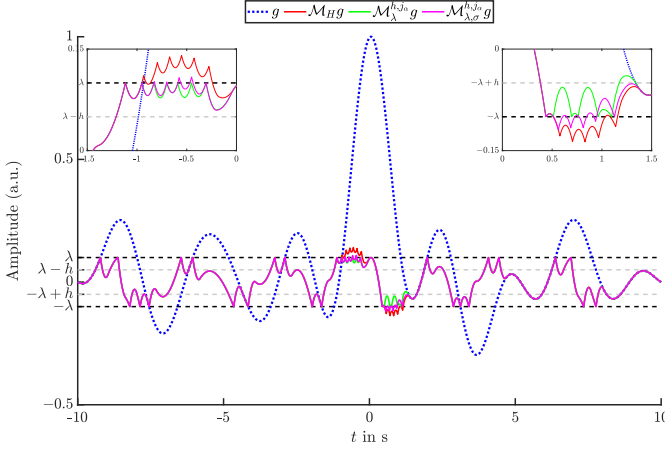


Fig. 7. Illustration of $\mathcal{M}_H g$, $\mathcal{M}_\lambda^{h,j_\alpha} g$ and $\mathcal{M}_{\lambda,\sigma}^{h,j_\alpha} g$ for $j_\alpha = j_\alpha^{(2)}$ and $g \in \text{PW}_\Omega$ with $\Omega = \pi \frac{\text{rad}}{\text{s}}$, where $\lambda = 0.1$, $h = 0.05$, $\alpha = 0.25$, $\sigma = 0.125$.

We now prove that $\mathcal{M}_\lambda^{h,j_\alpha} g = \mathcal{M}_{\lambda,\sigma}^{h,j_\alpha} g$. By definition we have $\nu_1 = \kappa_1$ and $\mathcal{M}_\lambda^{h,j_\alpha} g = \mathcal{M}_{\lambda,\sigma}^{h,j_\alpha} g$ for all $t \leq \nu_1 = \kappa_1$. Assume that for some $p \in \mathbb{N}$ we have $\nu_n = \kappa_n < \infty$ for all $n \leq p$. Hence, $\mathcal{M}_\lambda^{h,j_\alpha} g(t) = \mathcal{M}_{\lambda,\sigma}^{h,j_\alpha} g(t)$ for all $t \leq \nu_p = \kappa_p$ and $\zeta_p = \xi_p$. On the one hand,

$$\begin{aligned} \nu_{p+1} &= \inf\{t > \nu_p + \sigma \mid |\xi_p(t)| \geq \lambda\} \\ &= \inf\{t > \kappa_p + \sigma \mid |\zeta_p(t)| \geq \lambda\} \\ &\geq \inf\{t > \kappa_p \mid |\zeta_p(t)| \geq \lambda\} = \kappa_{p+1}. \end{aligned}$$

On the other hand, since $\tau_p = \kappa_p$ and $\tau_{p+1} = \kappa_{p+1}$, Theorem 2 gives $\kappa_{p+1} = \tau_{p+1} \geq \tau_p + \alpha = \kappa_p + \alpha$ and we obtain

$$\begin{aligned} \nu_{p+1} &= \inf\{t > \kappa_p + \sigma \mid |\zeta_p(t)| \geq \lambda\} \\ &\leq \inf\{t > \kappa_p + \alpha \mid |\zeta_p(t)| \geq \lambda\} \\ &= \inf\{t > \kappa_p \mid |\zeta_p(t)| \geq \lambda\} = \kappa_{p+1}. \end{aligned}$$

This gives $\mathcal{M}_\lambda^{h,j_\alpha} g(t) = \mathcal{M}_{\lambda,\sigma}^{h,j_\alpha} g(t)$ for all $t \leq \nu_{p+1} = \kappa_{p+1}$ and, by induction, we conclude that $\mathcal{M}_\lambda^{h,j_\alpha} g = \mathcal{M}_{\lambda,\sigma}^{h,j_\alpha} g$. $\square$

If $\alpha > 0$ and the assumptions of Theorem 5 are not satisfied, the generalized modulo operators $\mathcal{M}_H$, $\mathcal{M}_\lambda^{h,j_\alpha}$ and $\mathcal{M}_{\lambda,\sigma}^{h,j_\alpha}$ in general result in different output functions. For comparison, the conditions required to guarantee the properties studied in this paper for the three operators are summarized in Table I.

For a function g not meeting the assumptions of Theorem 5 we now study the behaviour of the proposed modulo operators numerically. This requires an approximation of the folding points $(\tau_n)_{n \in \mathbb{N}}$, $(\kappa_n)_{n \in \mathbb{N}}$ and $(\nu_n)_{n \in \mathbb{N}}$. To this end, we choose a sufficiently small sampling rate $d > 0$ and set

$$\begin{aligned} \tau_{n+1} &= \inf\{kd > \tau_n \mid |\eta_n^{(0)}(kd)| \geq \lambda, k \in \mathbb{Z}\}, \\ \kappa_{n+1} &= \inf\{kd > \kappa_n \mid |\zeta_n(kd)| \geq \lambda, k \in \mathbb{Z}\}, \\ \nu_{n+1} &= \inf\{kd > \nu_n + \sigma \mid |\xi_n(kd)| \geq \lambda, k \in \mathbb{Z}\}. \end{aligned} \quad (9)$$

Moreover, we consider one of the three folding functions $j_\alpha^{(1)}, j_\alpha^{(2)}, j_\alpha^{(3)}: \mathbb{R} \rightarrow \mathbb{R}$, for $t \in [0, \alpha]$ given by

$$j_\alpha^{(1)}(t) = \frac{t}{\alpha}, j_\alpha^{(2)}(t) = 1 - \left(\frac{\alpha-t}{\alpha}\right)^2, j_\alpha^{(3)}(t) = 3\left(\frac{t}{\alpha}\right)^2 - 2\left(\frac{t}{\alpha}\right)^3.$$

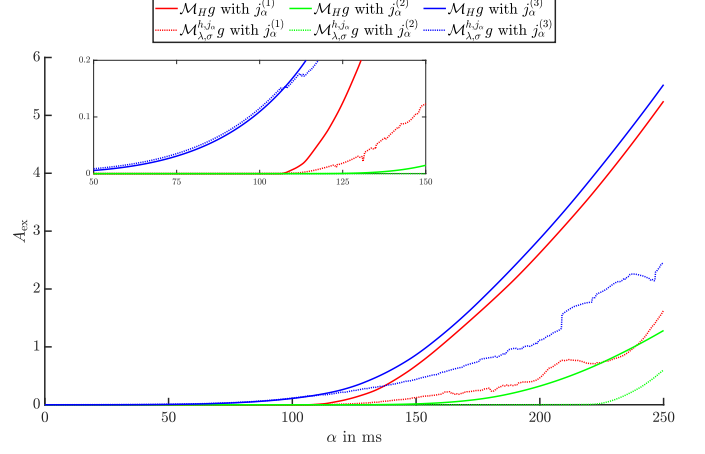


Fig. 8. Exceedance A_{ex} for $\mathcal{M}_H g$ (solid) and $\mathcal{M}_{\lambda,\sigma}^{h,j_\alpha} g$ (dotted) with $\sigma = \frac{\alpha}{2}$ for random function $g \in \text{PW}_\Omega$ with $\Omega = \pi \frac{\text{rad}}{\text{s}}$, where $d = 0.0463\text{ms}$ and $j_\alpha = j_\alpha^{(1)}$ (red), $j_\alpha = j_\alpha^{(2)}$ (green) or $j_\alpha = j_\alpha^{(3)}$ (blue).

Because condition (C₂) is not satisfied for $j_\alpha^{(3)}$, we cannot combine it with the generalized modulo operator $\mathcal{M}_\lambda^{h,j_\alpha}$.

We are mainly interested in studying the exceedance of the range $[-\lambda, \lambda]$. In Theorem 3 we have shown that $\mathcal{M}_\lambda^{h,j_\alpha} g$ always stays within $[-\lambda, \lambda]$. This can indeed be observed numerically in Figure 7. In contrast to this, $\mathcal{M}_H g$ and $\mathcal{M}_{\lambda,\sigma}^{h,j_\alpha} g$ leave the range $[-\lambda, \lambda]$. To quantify the severeness, we calculate the exceedance area of $g_\lambda \in \{\mathcal{M}_H g, \mathcal{M}_{\lambda,\sigma}^{h,j_\alpha} g\}$ via

$$A_{\text{ex}} = \int_{|g_\lambda(t)| > \lambda} (|g_\lambda(t)| - \lambda) dt$$

for all three folding functions. The results are shown in Figure 8 for varying values of α . We observe that for all folding functions the exceedance of $\mathcal{M}_H g$ is more severe than for $\mathcal{M}_{\lambda,\sigma}^{h,j_\alpha} g$ and the difference is more pronounced for larger values of α . Moreover, the exceedance grows with smaller initial slopes of the folding function j_α . Indeed, $j_\alpha^{(3)}$ yields the largest exceedance area followed by $j_\alpha^{(1)}$ and, lastly, $j_\alpha^{(2)}$.

IV. RECONSTRUCTION USING OMP

We first state a condition for the sampling rate $T > 0$ to uniquely determine $g \in \text{PW}_\Omega \cap \mathcal{C}_{\lambda,\tau_0}^{0,1}$ with bandwidth $\Omega > 0$ from discrete samples of $g_\lambda \in \{\mathcal{M}_H g, \mathcal{M}_\lambda^{h,j_\alpha} g, \mathcal{M}_{\lambda,\sigma}^{h,j_\alpha} g\}$,

$$\{g_\lambda(kT) \mid k \in \mathbb{Z}\},$$

where we assume that g_λ is well-defined and that there exists $\rho > 0$ such that $g_\lambda(t) = g(t)$ for all $|t| \geq \rho$, c.f. Table I for sufficient conditions in the different cases. In [37, Lemma 3] we have shown that any $g \in \text{PW}_\Omega$ is determined by $\{Ag(kT) \mid k \in \mathbb{Z}\}$ if the oversampling condition $T < \frac{\pi}{\Omega}$ is met, where $\mathcal{A}: \text{PW}_\Omega \rightarrow L^2(\mathbb{R})$ is a general operator with

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

for some bounded set $M \subset \mathbb{R}$. Hence, applying [37, Lemma 3] to our setting directly gives the following identifiability result.

Proposition 6. For $g \in \text{PW}_\Omega \cap \mathcal{C}_{\lambda, \tau_0}^{0,1}$ let g_λ be well-defined and $\rho > 0$ with $g_\lambda(t) = g(t)$ for all $|t| \geq \rho$. Then, g is uniquely determined by the samples $\{g_\lambda(kT) \mid k \in \mathbb{Z}\}$ if $T < \frac{\pi}{\Omega}$. $\square$

We now adapt the recovery approach from [37] to our setting. To this end, we assume that we are given discrete samples

$$\{g_\lambda(kT) \mid k \in \mathbb{Z}\}$$

with $T < \frac{\pi}{\Omega}$ and $g_\lambda(kT) = g(kT)$ for all $|k| \geq K \in \mathbb{N}$. Then, there exist $L_\lambda \in \mathbb{N}$ folding points $\kappa_1, \dots, \kappa_{L_\lambda} \in [-K, K]T$ and signs $\sigma_1, \dots, \sigma_{L_\lambda} \in \{\pm 1\}$ such that $g_\lambda = g - s_\lambda$ with $s_\lambda = (2\lambda - h) \sum_{l=1}^{L_\lambda} \sigma_l j_\alpha(\cdot - \kappa_l)$. As in [37], we can represent the samples of the folding function j_α by the sum of multiple instantaneous folds on $T\mathbb{Z}$. Hence, defining $\underline{g} = g(\cdot + T) - g$, we obtain $\underline{g}_\lambda = \underline{g} - s_\lambda$ with

$$s_\lambda(nT) = \sum_{\ell \in \mathbb{L}_\lambda} c_\ell \delta(nT - t_\ell)$$

for suitable $\mathbb{L}_\lambda \subseteq \{-K, \dots, K\}$, $c_\ell \in \mathbb{R}$ and $t_\ell \in (T\mathbb{Z}) \cap [-KT, KT]$, where $|\mathbb{L}_\lambda| \leq L_\lambda (\lceil \frac{\alpha}{T} \rceil + 1)$. We now consider the discrete time Fourier transform (DTFT)

$$\widehat{\underline{g}}(\omega) = \sum_{n \in \mathbb{Z}} \underline{g}(nT) e^{-inT\omega} \quad \text{for } \omega \in \mathbb{R}.$$

As $g \in \text{PW}_\Omega$ implies $\underline{g} \in \text{PW}_\Omega$, Poisson's summation formula

$$\sum_{n \in \mathbb{Z}} \mathcal{F} \underline{g}(\omega + \frac{2\pi}{T}n) = T \sum_{n \in \mathbb{Z}} \underline{g}(nT) e^{-inT\omega}$$

implies that $\widehat{\underline{g}}$ is $\frac{2\pi}{T}$ -periodic and, for $\omega \in [0, \frac{2\pi}{T})$,

$$\widehat{\underline{g}}(\omega) = \begin{cases} \frac{1}{T} \mathcal{F} \underline{g}(\omega) & \text{for } 0 \leq \omega \leq \Omega, \\ 0 & \text{for } \Omega < \omega < \frac{2\pi}{T} - \Omega, \\ \frac{1}{T} \mathcal{F} \underline{g}(\omega - \frac{2\pi}{T}) & \text{for } \frac{2\pi}{T} - \Omega \leq \omega < \frac{2\pi}{T}. \end{cases}$$

In particular, for all $\Omega < \omega < \frac{2\pi}{T} - \Omega$ we obtain

$$\widehat{\underline{g}}_\lambda(\omega) = -\widehat{s}_\lambda(\omega) = -\sum_{\ell \in \mathbb{L}_\lambda} c_\ell e^{-i\omega t_\ell}.$$

and, with $\omega_m = \frac{2\pi m}{NT} \in [0, \frac{2\pi}{T})$ for $m = 0, \dots, N-1$ and $N \geq K$,

$$\widehat{\underline{g}}_\lambda(\omega_m) = -\sum_{\ell \in \mathbb{L}_\lambda} c_\ell e^{-i\omega_m t_\ell} \quad \forall m \in \mathbb{I}_{N_\Omega, N},$$

where $\omega_0 = \frac{2\pi}{NT}$ and $\mathbb{I}_{N_\Omega, N} = \{N_\Omega + 1, \dots, N - N_\Omega - 1\}$ with $N_\Omega = \lfloor \frac{\Omega NT}{2\pi} \rfloor$. For sake of brevity, we set $g[n] = g((n-K)T)$ and

$$\widehat{\underline{g}}_\lambda[m] = e^{-i\omega_0 m K} \widehat{\underline{g}}_\lambda(\omega_m) = \sum_{n \in \mathbb{Z}} g_\lambda[n] e^{-i\omega_0 m n}.$$

Hence, with $\mathbf{s} = (\widehat{\underline{g}}_\lambda[m])_{m \in \mathbb{I}_{N_\Omega, N}} \in \mathbb{C}^M$, $M = N - 2N_\Omega - 1$, we can find $\mathbf{c} = (c_k)_{k=-K}^K \in \mathbb{C}^{2K+1}$ with $c_k = 0$ for $k \notin \mathbb{L}_\lambda$ by solving the minimization problem

$$\text{minimize } \|\mathbf{c}\|_0 \quad \text{such that } \mathbf{V}\mathbf{c} = \mathbf{s}, \quad (10)$$

where $\mathbf{V} \in \mathbb{C}^{M \times (2K+1)}$ is a Vandermonde matrix with entries $V_{m-N_\Omega, n+1} = e^{-i\omega_0 m n}$ for $m \in \mathbb{I}_{N_\Omega, N}$, $n = 0, \dots, 2K$. The solution vector $\mathbf{c} \in \mathbb{C}^{2K+1}$ of (10) encodes the sought parameters $\{c_\ell\}_{\ell \in \mathbb{L}_\lambda}$ as its non-zero vector components and the set $\{t_\ell\}_{\ell \in \mathbb{L}_\lambda}$ is determined by the corresponding indices.

As explained in [29], [39] in the context of the modulo Radon transform [28], the optimization problem (10) can be

Algorithm 1 (OMP)

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

- 1: $\mathbf{c}^{(0)} = \mathbf{0} \in \mathbb{C}^J$, $\mathcal{S}^{(0)} = \emptyset$, $i = 1$
- 2: **while** $\|\mathbf{V}^*(\mathbf{s} - \mathbf{V}\mathbf{c}^{(i-1)})\|_\infty > \varepsilon$ **do**
- 3: $j^{(i)} = \arg \max_{1 \leq j \leq J} |\mathbf{V}^*(\mathbf{s} - \mathbf{V}\mathbf{c}^{(i-1)})|_j|$
- 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: $\mathbf{c}^{(i_{\text{end}})} \in \mathbb{C}^J$

solved by applying the orthogonal matching pursuit (OMP) algorithm, which was first proposed in [40]. Here we use a variant as described in Algorithm 1. Given the estimates $\{c_\ell, t_\ell\}_{\ell \in \mathbb{L}_\lambda}$, we can form $s_\lambda(kT) = \sum_{\ell \in \mathbb{L}_\lambda} c_\ell \delta(kT - t_\ell)$ for $k \in \mathbb{Z}$ and compute the samples $s_\lambda(kT)$ by applying anti-differences, i.e., $s_\lambda(kT) = \sum_{j < k} s_\lambda(jT)$. Adding $s_\lambda(kT)$ to $g_\lambda(kT)$ finally gives $g(kT)$ and we can recover g via Shannon interpolation, i.e.,

$$g = \sum_{k \in \mathbb{Z}} g(kT) \text{sinc}\left(\frac{\pi}{T}(\cdot - kT)\right).$$

In [29, Theorem 1] it is shown that Algorithm 1 with exact arithmetic and $\varepsilon = 0$ recovers the parameters $\{c_\ell, t_\ell\}_{\ell \in \mathbb{L}_\lambda}$ if $\max_{\ell \in \mathbb{L}_\lambda + K} |\ell| < N - 2N_\Omega - 1$. As $\max_{\ell \in \mathbb{L}_\lambda + K} |\ell| \leq 2K - 1$ because of $g_\lambda(kT) = g(kT)$ for all $|k| \geq K$ by choice of $K \in \mathbb{N}$, we set $N = K + K'$ and require that $2K - 1 < N - 2N_\Omega - 1$. Defining the oversampling factor $\text{OF} = \frac{\pi}{\Omega T} > 1$, this can be guaranteed if

$$K' > \frac{1 + \text{OF}^{-1}}{1 - \text{OF}^{-1}} K.$$

In total, we obtain the following reconstruction result for OMP.

Theorem 6. For $g \in \text{PW}_\Omega \cap \mathcal{C}_{\lambda, \tau_0}^{0,1}$ let g_λ be well-defined and $\rho > 0$ with $g_\lambda(t) = g(t)$ for all $|t| \geq \rho$. Assume we are given discrete samples $\{g_\lambda(kT) \mid k \in \mathbb{Z}\}$ with $T < \frac{\pi}{\Omega}$. Then, g is exactly recovered by OMP with $\varepsilon = 0$ and $N = K + K'$, where

$$K \geq \rho T^{-1}, \quad K' > \frac{1 + \text{OF}^{-1}}{1 - \text{OF}^{-1}} K. \quad \square$$

Note that the above statement does not guarantee that OMP only selects correct folding times. However, for falsely selected positions the corresponding coefficients will be zero. To guarantee stable recovery of the folding times, additional conditions are needed as the restricted isometry property or mutual coherence conditions. This will depend on the separation of folds and is beyond the scopes of this paper.

A. Iterated forward differences

For $\alpha > 0$, the above reconstruction approach splits one non-instantaneous fold into up to $\lceil \frac{\alpha}{T} \rceil + 1$ instantaneous folds. This makes the reconstruction via OMP more challenging, especially for large values of α compared to T . To deal with this case, we now modify our reconstruction approach for

Algorithm 2 (SAOMP)

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

```

1:  $\mathbf{c}^{(0)} = 0 \in \mathbb{C}^J$ ,  $\mathcal{S}^{(0)} = \emptyset$ ,  $\delta = \nu$ ,  $i = 1$ 
2: while  $i \leq i_{\max}$  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:    $\mathcal{S}^{(i)} = \mathcal{S}^{(i-1)} \cup \{j \in \{1, \dots, J\} \mid |\mathbf{r}_j^{(i)}| \geq \delta \|\mathbf{r}^{(i)}\|_{\infty}\}$ 
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:    $\mathcal{S}^{(i)} = \{j \in \{1, \dots, J\} \mid |\mathbf{c}_j^{(i)}| \geq \mu \|\mathbf{c}^{(i)}\|_{\infty}\}$ 
7:   for  $j \notin \mathcal{S}^{(i)}$ :  $\mathbf{c}_j^{(i)} = 0$ 
8:    $\delta = \delta + \frac{1-\nu}{i_{\max}}$ ,  $i = i + 1$ 
9: end while

```

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

certain smooth folding functions. More precisely, we assume that $j_{\alpha} \in \mathcal{C}^{p-2}(\mathbb{R})$ for some $p \geq 2$ and $j_{\alpha}|_{[0, \alpha]} \in \mathcal{P}_{p-1}$, i.e., j_{α} is a polynomial of degree $p-1$ on $[0, \alpha]$. Note that such a folding function cannot satisfy assumption (C₂) so that in the following we have to deal with $\mathcal{M}_H$ or $\mathcal{M}_{\lambda, \sigma}^{h, j_{\alpha}}$. Moreover, we consider the forward difference operator $\Delta_T: \mathcal{C}(\mathbb{R}) \rightarrow \mathcal{C}(\mathbb{R})$ given by $\Delta_T f = f(\cdot + T) - f$ and the divided difference operator $[x_0, \dots, x_p]: \mathcal{C}(\mathbb{R}) \rightarrow \mathbb{R}$ with $x_j = (j+m)T$ for $j = 0, \dots, p$ and some $m \in \mathbb{Z}$, defined via the recursion

$$[x_j, \dots, x_k](f) = \frac{[x_{j+1}, \dots, x_k](f) - [x_j, \dots, x_{k-1}](f)}{x_k - x_j}$$

for $0 \leq j < k \leq p$, where $[x_j](f) = f(x_j)$ for $j = 0, \dots, p$. By induction one can show that

$$\forall 0 \leq k \leq p \forall 0 \leq j \leq p-k: [x_j, \dots, x_{j+k}](f) = \frac{\Delta_T^k f(x_j)}{k! T^k}$$

and, if $f \in \mathcal{C}^p([x_0, x_L])$, the mean-value theorem gives

$$[x_0, \dots, x_p](f) = \frac{f^{(p)}(\tau)}{p!} \quad \text{for some } \tau \in [x_0, x_p].$$

Hence, if $[x_0, x_p] \subseteq [0, \alpha]$ we observe that $\Delta_T^p j_{\alpha}(mT) = 0$. The same holds if $[x_0, x_p] \subset (-\infty, 0]$ or $[x_0, x_p] \subset [\alpha, \infty)$ so that $\Delta_T^p j_{\alpha}(mT) \neq 0$ only if $0 \in (x_0, x_p)$ or $\alpha \in (x_0, x_p)$. Analogously, for $j_{\alpha}^{(\kappa)} = j_{\alpha}(\cdot - \kappa)$ located at $\kappa \in \mathbb{R}$ we have $\Delta_T^p j_{\alpha}^{(\kappa)}(mT) \neq 0$ only if $\kappa \in (x_0, x_p)$ or $\kappa + \alpha \in (x_0, x_p)$, which is equivalent to $m \in (\frac{\kappa}{T} - p, \frac{\kappa}{T}) \cup (\frac{\kappa + \alpha}{T} - p, \frac{\kappa + \alpha}{T})$. Consequently, if $j_{\alpha} \in \mathcal{C}^{p-2}(\mathbb{R})$ is a polynomial of degree $p-1$ on $[0, \alpha]$, applying the iterated forward difference operator Δ_T^p splits one non-instantaneous fold into at most $2p$ instantaneous folds, which is advantageous over the previous approach if

$$\left\lceil \frac{\alpha}{T} \right\rceil > 2p - 1.$$

To explain the full adapted recovery approach, we again assume that we are given discrete samples

$$\{g_{\lambda}(kT) \mid k \in \mathbb{Z}\}$$

with $T < \frac{\pi}{\Omega}$ and $g_{\lambda}(kT) = g(kT)$ for all $|k| \geq K \in \mathbb{N}$. Then, there exist $L_{\lambda} \in \mathbb{N}$ folding points $\kappa_1, \dots, \kappa_{L_{\lambda}} \in [-K, K]T$

and signs $\sigma_1, \dots, \sigma_{L_{\lambda}} \in \{\pm 1\}$ such that $g_{\lambda} = g - s_{\lambda}$ with $s_{\lambda} = (2\lambda - h) \sum_{l=1}^{L_{\lambda}} \sigma_l j_{\alpha}(\cdot - \kappa_l)$. Now, defining $\underline{g} = \Delta_T^p g$, we obtain $\underline{g}_{\lambda} = \underline{g} - \underline{s}_{\lambda}$ with

$$\underline{s}_{\lambda}(nT) = \sum_{\ell \in \mathbb{L}_{\lambda}} c_{\ell} \delta(nT - t_{\ell})$$

for suitable $\mathbb{L}_{\lambda} \subseteq \{-K, \dots, K\}$ and $t_{\ell} \in (T\mathbb{Z}) \cap [-KT, KT]$, where $|\mathbb{L}_{\lambda}| \leq 2pL_{\lambda}$. As before, DTFT in combination with the bandlimitedness of $g \in \text{PW}_{\Omega}$ gives

$$\widehat{\underline{g}}_{\lambda}[m] = - \sum_{\ell \in \mathbb{L}_{\lambda}} c_{\ell} e^{-i\omega_0 m(\ell + K)} \quad \forall m \in \mathbb{I}_{N_{\Omega}, N},$$

where the left-hand side is computable from given modulo data $g[n] = g((n-K)T)$ via

$$\widehat{\underline{g}}_{\lambda}[m] = \sum_{n \in \mathbb{Z}} \underline{g}_{\lambda}[n] e^{-i\omega_0 mn}.$$

Consequently, we end up with the same minimization problem

$$\text{minimize } \|\mathbf{c}\|_0 \quad \text{such that } \mathbf{V}\mathbf{c} = \mathbf{s},$$

which can be solved by applying OMP. Given the estimates $\{c_{\ell}, t_{\ell}\}_{\ell \in \mathbb{L}_{\lambda}}$, we can form $\underline{s}_{\lambda}(kT) = \sum_{\ell \in \mathbb{L}_{\lambda}} c_{\ell} \delta(kT - t_{\ell})$ for $k \in \mathbb{Z}$ and compute the samples $s_{\lambda}(kT)$ by iteratively applying anti-differences p -times. Adding $s_{\lambda}(kT)$ to $\underline{g}_{\lambda}(kT)$ finally gives $g(kT)$ and we can recover g via Shannon interpolation, i.e.,

$$g = \sum_{k \in \mathbb{Z}} g(kT) \text{sinc}\left(\frac{\pi}{T}(\cdot - kT)\right).$$

B. Reconstruction adapted to folding function

The ability to model the folding function allows us to adapt the dictionary for the Vandermonde matrix $\mathbf{V}$ so that a fold can be represented by *one* column of the matrix. This is especially useful if the transient α is much larger than the sampling rate T . To this end, we assume that the folding points satisfy $\kappa_1, \dots, \kappa_{L_{\lambda}} \in (T\mathbb{Z}) \cap [-KT, KT]$. As this is usually not satisfied, we encounter a small error by approximating folds not exactly located at sampling points with ones located at the nearest sampling points. Now, for $n \in \mathbb{Z}$ we have

$$s_{\lambda}(nT) = \sum_{\ell \in \mathbb{L}_{\lambda}} c_{\ell} j_{\alpha}(nT - t_{\ell}), \quad (11)$$

with $c_{\ell} \in (2\lambda - h)\mathbb{Z}$ and applying forward differences gives

$$\underline{s}_{\lambda}(nT) = \sum_{\ell \in \mathbb{L}_{\lambda}} c_{\ell} (j_{\alpha}((n+1)T - t_{\ell}) - j_{\alpha}(nT - t_{\ell})),$$

where $t_{\ell} \in (T\mathbb{Z}) \cap [-KT, KT]$ and $\mathbb{L}_{\lambda} \subseteq \{-K, \dots, K\}$ with $|\mathbb{L}_{\lambda}| = L_{\lambda}$. Thereon, DTFT in combination with the bandlimitedness of $g \in \text{PW}_{\Omega}$ yields

$$\widehat{\underline{g}}_{\lambda}[m] = - \sum_{\ell \in \mathbb{L}_{\lambda}} c_{\ell} \sum_{k=0}^{\lceil \frac{\alpha}{T} \rceil - 1} j_{\alpha}(kT) e^{-i\omega_0 m(k + \ell + K)}$$

for all $m \in \mathbb{I}_{N_{\Omega}, N}$. Hence, with $\mathbf{s} = (\widehat{\underline{g}}_{\lambda}[m])_{m \in \mathbb{I}_{N_{\Omega}, N}} \in \mathbb{C}^M$ we obtain the minimization problem

$$\text{minimize } \|\mathbf{c}\|_0 \quad \text{such that } \mathbf{V}\mathbf{c} = \mathbf{s},$$

however with $\mathbf{V}_{m-N_{\Omega}, n+1} = \sum_{k=0}^{\lceil \frac{\alpha}{T} \rceil - 1} j_{\alpha}(kT) e^{-i\omega_0 m(k+n)}$. Having obtained the solution vector $\mathbf{c} \in \mathbb{C}^{2K+1}$, we can reconstruct s_{λ} using (11) and, afterwards, g via $g = g_{\lambda} + s_{\lambda}$.

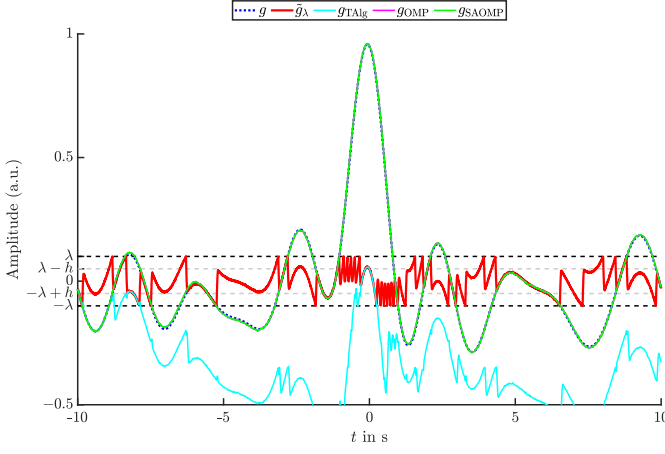


Fig. 9. Illustration of TALg, OMP and SAOMP reconstruction for random $g \in \text{PW}_\Omega$ with $\Omega = \pi \frac{\text{rad}}{\text{s}}$ based on noisy modulo samples $\tilde{g}_\lambda(nT)$ with an SNR of 30dB and $\lambda = 0.1$, $h = 0.05$, $\alpha = 50\text{ms}$, $T = 20.8\text{ms}$, $j_\alpha = j_\alpha^{(2)}$.

V. NUMERICAL EXPERIMENTS

In practice, only finitely many samples are available and we assume that we are given $\{g_\lambda(kT) \mid k = -K, \dots, K'\}$. Then, we set $N = K + K' + 1 - p$ and approximate the left-hand side $\hat{g}_\lambda[m]$ for $m = 0, \dots, N - 1$ via

$$\hat{g}_\lambda[m] \approx \sum_{n=0}^{N-1} g_\lambda[n] e^{-i\omega_0 mn},$$

which can be efficiently computed with the fast Fourier transform (FFT). Moreover, as in [37], we speed up the recovery by applying the stagewise arithmetic orthogonal matching pursuit (SAOMP) algorithm proposed in [41], see Algorithm 2, which has additional parameters ν to find multiple folding points in one iteration and μ to remove incorrectly selected ones. Note that $\nu = 1$ and $\mu = 0$ yields the original OMP algorithm. As a general rule, choosing a smaller value for ν allows for more folding points to be picked in one iteration and thereby can yield a faster recovery. However, choosing ν too small might result in selecting false entries of the dictionary and can hinder accurate recovery. To account for this, the parameter μ is incorporated for pruning the dictionary entry selection, whose value should be smaller than the ration of the largest and smallest correct non-zero entry in $\mathbf{c}$, which depends on the number of folds occurring at approximately the same time and on the flatness of the folding function. The parameter ε is independent of ν and μ and accounts for noise. Its value should be smaller than the smallest non-zero entry in $\mathbf{c}$.

We now present some numerical simulations for a proof of concept and to illustrate the versatility of our reconstruction approach. To this end, we randomly select a function $g \in \text{PW}_\Omega$ with $\Omega = \pi$ and approximate g_λ numerically. To guarantee a good discretization of the generalized modulo operators, we choose $d > 0$ sufficiently small to compute the folding points as in (9). For reconstruction, however, we only use the modulo samples $\{g_\lambda(kT) \mid k = -K, \dots, K'\}$ with $T \gg d$. Moreover, for simplicity, we mainly focus on the delayed generalized modulo operator $\mathcal{M}_{\lambda, \sigma}^{h, j_\alpha}$ with $\sigma = \frac{\alpha}{2}$, which allows us to compare all three folding functions $j_\alpha^{(1)}$, $j_\alpha^{(2)}$, $j_\alpha^{(3)}$ from above.

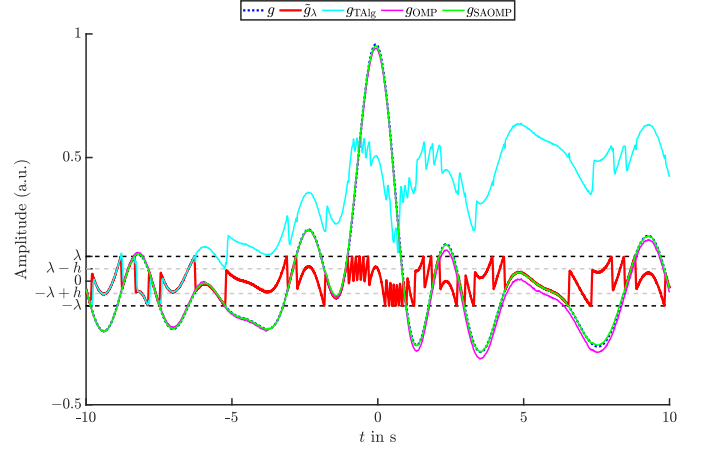


Fig. 10. Illustration of TALg, OMP and SAOMP reconstruction for random $g \in \text{PW}_\Omega$ with $\Omega = \pi \frac{\text{rad}}{\text{s}}$ based on noisy modulo samples $\tilde{g}_\lambda(nT)$ with an SNR of 30dB and $\lambda = 0.1$, $h = 0.05$, $\alpha = 50\text{ms}$, $T = 20.8\text{ms}$, $j_\alpha = j_\alpha^{(3)}$.

In our first experiments we compare our OMP and SAOMP approach with the thresholding algorithm (TAlg) from [34], which was designed for the generalized modulo encoder $\mathcal{M}_H$ with $j_\alpha^{(1)}$. In [37] we have already seen that TALg is in general not able to recover g from $g_\lambda = \mathcal{M}_{\lambda, j_\alpha}^{h, j_\alpha} g$ with $j_\alpha = j_\alpha^{(1)}$, while SAOMP gives satisfactory results. The case $g_\lambda = \mathcal{M}_{\lambda, \sigma}^{h, j_\alpha} g$ with $j_\alpha = j_\alpha^{(2)}$ is shown in Fig. 9, where we chose $\lambda = 0.1$, $h = 0.05$, $\alpha = 0.05$ and $T = 0.0208$. Moreover, we added Gaussian noise yielding noisy samples $\tilde{g}_\lambda(nT)$ with a signal-to-noise ratio (SNR) of 30dB. We observe that OMP and SAOMP yield the same near perfect reconstruction with a mean squared error (MSE) of $2.1 \cdot 10^{-5}$. However, OMP needs 122 iterations and 2.6 seconds, whereas SAOMP only requires 86 iterations and 1.3 seconds. In contrast to this, TALg is not able to reconstruct g in this scenario. The results for $j_\alpha = j_\alpha^{(3)}$ are shown in Fig. 10, where we use the same parameters and again add Gaussian noise yielding an SNR of 30dB. As before, TALg fails, while OMP and SAOMP give satisfactory results. Now, however, SAOMP performs better than OMP with an MSE of $1.6 \cdot 10^{-5}$, while OMP yields an MSE of $2.8 \cdot 10^{-4}$.

In Fig. 11 we repeat the experiment from Fig. 9 with additional shot noise [42] at up to 100 positions and with a maximal magnitude of 0.5 yielding an SNR of 0.8dB. Despite this very low SNR, both OMP and SAOMP are able to reconstruct the signal with a MSE of $4.5 \cdot 10^{-5}$. This is because each position of shot noise is interpreted by the algorithms as two subsequent folds and is therefore automatically removed.

When increasing the quotient $\frac{\alpha}{T}$, each non-instantaneous fold is split into a growing number of instantaneous ones and the reconstruction via standard SAOMP becomes more challenging. In this case, our modified approaches with iterated forward differences and adapted dictionaries yield better results. To illustrate this, in Fig. 12 we recover a random $g \in \text{PW}_\pi$ from $g_\lambda(nT) = \mathcal{M}_{\lambda, \sigma}^{h, j_\alpha} g(nT)$ with $j_\alpha = j_\alpha^{(1)}$, $\lambda = 0.1$, $h = 0.05$, $\alpha = 0.1$ and $T = 0.0052$. While standard SAOMP fails, SAOMP with an adapted dictionary yields an MSE of $2.5 \cdot 10^{-5}$. The reconstruction quality for SAOMP with iterated forward differences, where $p = 2$, is even better

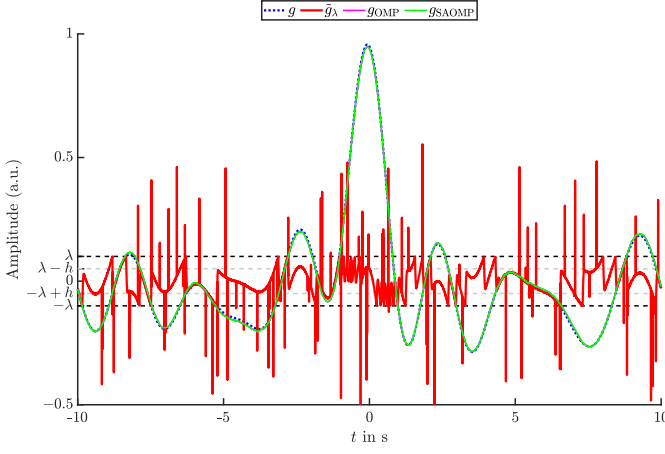


Fig. 11. Illustration of (SA)OMP reconstruction for random $g \in \text{PW}_\Omega$ with $\Omega = \pi \frac{\text{rad}}{\text{s}}$ based on noisy modulo samples $\tilde{g}_\lambda(nT)$ with shot noise at up to 100 positions and $\lambda = 0.1$, $h = 0.05$, $\alpha = 50\text{ms}$, $T = 20.8\text{ms}$, $j_\alpha = j_\alpha^{(2)}$.

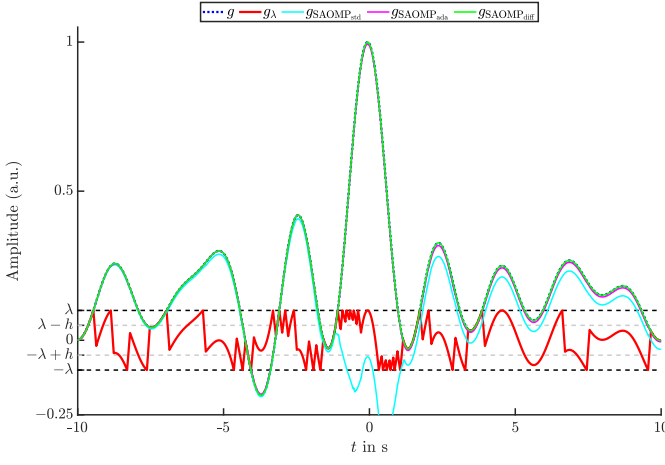


Fig. 12. Illustration of SAOMP reconstruction (standard with and without iterated forward differences vs. adapted dictionary) for random $g \in \text{PW}_\Omega$ with $\Omega = \pi \frac{\text{rad}}{\text{s}}$ based on samples $g_\lambda(nT)$ and $\lambda = 0.1$, $h = 0.05$, $\alpha = 100\text{ms}$, $T = 5.2\text{ms}$ and $j_\alpha = j_\alpha^{(1)}$, where the dictionary is adapted to $j_\alpha^{(1)}$.

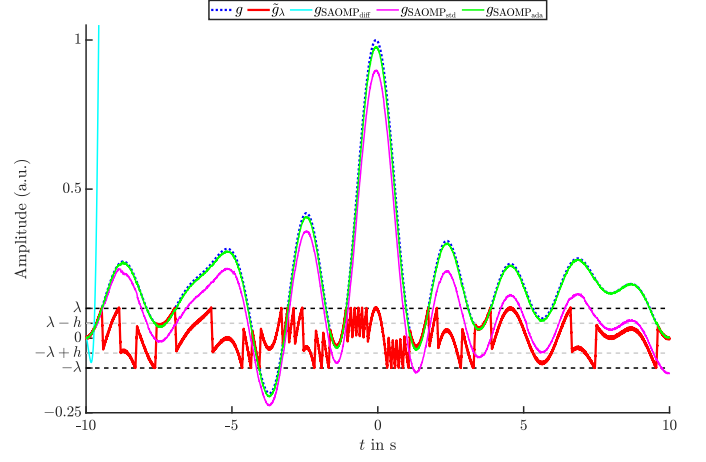


Fig. 13. Illustration of SAOMP reconstruction (standard with and without iterated forward differences vs. adapted dictionary) for random $g \in \text{PW}_\Omega$, $\Omega = \pi \frac{\text{rad}}{\text{s}}$, based on noisy samples $\tilde{g}_\lambda(nT)$ with an SNR of 30dB and $\lambda = 0.1$, $h = 0.05$, $\alpha = 50\text{ms}$, $T = 5.2\text{ms}$, $j_\alpha = j_\alpha^{(1)}$.

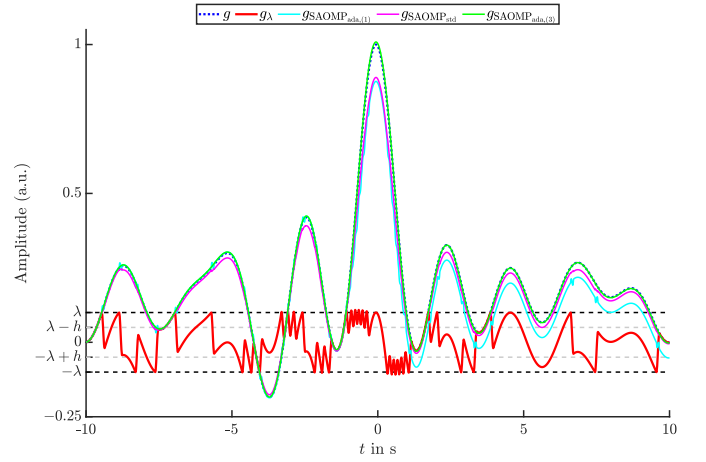


Fig. 14. Illustration of SAOMP reconstruction (standard with and without iterated forward differences vs. adapted dictionary) for random $g \in \text{PW}_\Omega$ with $\Omega = \pi \frac{\text{rad}}{\text{s}}$ based on samples $g_\lambda(nT)$ and $\lambda = 0.1$, $h = 0.05$, $\alpha = 100\text{ms}$, $T = 5.2\text{ms}$ and $j_\alpha = j_\alpha^{(3)}$, where the dictionary is falsely adapted to $j_\alpha^{(1)}$ (magenta) and correctly to $j_\alpha^{(3)}$ (cyan).

yielding an MSE of $6.1 \cdot 10^{-10}$. This, however, is only true in the noise-free setting, as shown in Fig. 13, where we chose $\alpha = 0.05$ and added Gaussian noise yielding noisy samples $\tilde{g}_\lambda(nT)$ with an SNR of 30dB. In this case, SAOMP with iterated forward differences fails, while SAOMP with adapted dictionary yields the best result with an MSE of $9.0 \cdot 10^{-5}$.

When considering $g_\lambda = \mathcal{M}_{\lambda, \sigma}^{h, j_\alpha} g$ with $j_\alpha = j_\alpha^{(3)}$, the correct number of forward differences is $p = 4$. While the iterated forward differences do not create additional errors in the noise-free setting, the necessary iterated anti-differences amplify small inaccuracies in the identification of folding points and, consequently, render this approach inappropriate. Hence, in Fig. 14 we compare standard SAOMP with adapted dictionaries, where in addition to $j_\alpha^{(3)}$ we falsely adapt to $j_\alpha^{(1)}$. We observe that SAOMP with correctly adapted dictionary yields a satisfying reconstruction with an MSE of $7.3 \cdot 10^{-6}$, while the wrong dictionary leads to an MSE of $2.5 \cdot 10^{-3}$, worse than standard SAOMP with an MSE of $8.9 \cdot 10^{-4}$. This

illustrates the necessity of correctly adapting the dictionary and, thus, the need for an accurate model. The same remains true when adding noise, as illustrated in Fig. 15.

In our last numerical experiment, we study the influence of the reset time σ on the reconstruction quality of SAOMP with adapted dictionary. As a larger value for σ yields a larger separation of folds, we expect that reconstruction becomes easier with increasing σ . This can indeed be observed numerically, as illustrated in Fig. 16, where we plot the median relative MSE as a function of σ . More precisely, we computed the relative MSE for the SAOMP reconstruction with adapted dictionary from $g_\lambda(nT) = \mathcal{M}_{\lambda, \sigma}^{h, j_\alpha} g(nT)$ with $j_\alpha = j_\alpha^{(1)}$, $\lambda = 0.1$, $h = 0.05$, $\alpha = 0.1$, $T = 0.0052$ and varying $\sigma \in [\frac{\alpha}{2}, 2\alpha]$ for 25 random functions $g \in \text{PW}_\pi$. We observe that for $\frac{\alpha}{2} \leq \sigma \leq \alpha$ the median relative MSE is constant and of the same level as for $g_\lambda = \mathcal{M}_h g$. Increasing σ , however, leads to a significant drop yielding better results than for $g_\lambda = \mathcal{M}_\lambda^{h, j_\alpha} g$.

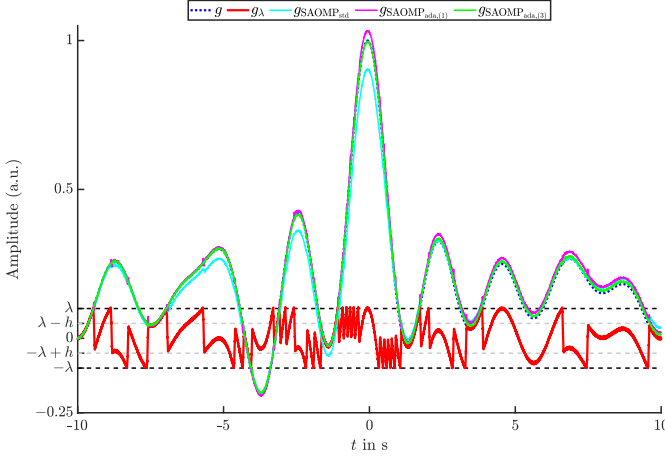


Fig. 15. Illustration of SAOMP reconstruction (standard with and without iterated forward differences vs. adapted dictionary) for random $g \in \text{PW}_\Omega$, $\Omega = \pi \frac{\text{rad}}{\text{s}}$, based on noisy samples $\hat{g}_\lambda(nT)$ with an SNR of 30dB and $\lambda = 0.1$, $h = 0.05$, $\alpha = 50\text{ms}$, $T = 5.2\text{ms}$, $j_\alpha = j_\alpha^{(3)}$.

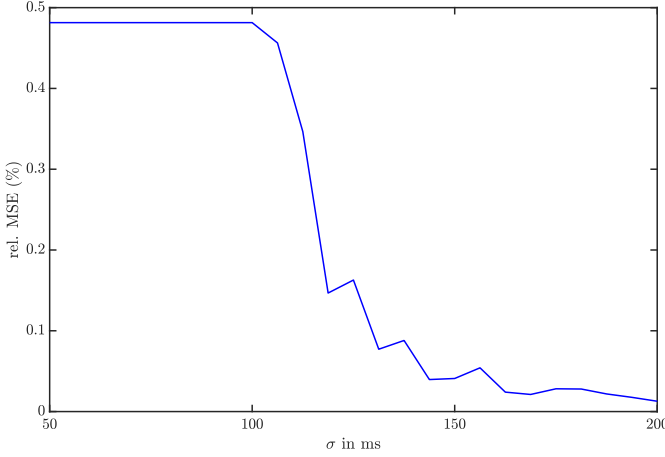


Fig. 16. Median relative MSE of SAOMP reconstructions with adapted dictionary for 25 random $g \in \text{PW}_\Omega$ with $\Omega = \pi \frac{\text{rad}}{\text{s}}$ based on samples $\hat{g}_\lambda(nT)$ with $g_\lambda = \mathcal{M}_{\lambda, \sigma}^{h, j_\alpha} g$, $\lambda = 0.1$, $h = 0.05$, $\alpha = 100\text{ms}$, $T = 5.2\text{ms}$, $j_\alpha = j_\alpha^{(1)}$ and varying $\sigma \in [\frac{\alpha}{2}, 2\alpha]$. For $g_\lambda = \mathcal{M}_H g$ the median relative MSE is 0.48% and for $g_\lambda = \mathcal{M}_{\lambda, \sigma}^{h, j_\alpha} g$ the median relative MSE is 0.37%

VI. CONCLUSION

The unlimited sensing framework overcomes the intrinsic trade-off between dynamic range and digital resolution in conventional Shannon-Nyquist sampling by exploiting that for any bandlimited signal its fractional part encodes the integer part. To account for hardware imperfections, modulo hysteresis operators were proposed leading to a perturbed decomposition of the signal. In this paper we introduced three generalized modulo hysteresis operators, which differ from the standard linear folding transition by incorporating a general transition function, which for two of the operators can take any form between zero and α . The operators differ in the way they determine the folding times. While the first one determines the folding points in relation to an ideal folding transition, the other two make the position of the folds depend on the transient nonideality either with or without an additional reset

time. We rigorously analysed the mathematical properties of the three operators and studied the question under which conditions the perturbed fractional part still encodes the whole signal. In particular, we proved conditions for identifiability of bandlimited signals from discrete generalized modulo samples. Moreover, we described a mathematically backed reconstruction algorithm based on orthogonal matching pursuit, which can be adapted to the specific modulo operator and folding function. This allowed us to also reconstruct signals in the challenging regime where the transition parameter is large compared to the sampling rate. Our numerical results indicate that the reconstruction quality depends on the utilized model and recovery works best for transient depending folding points with reset time larger than the transient parameter.

Our work raises interesting questions and possible extensions for future research. Key directions include

- validating our theoretical results through hardware experiments to study the applicability of our models in practice,
- theoretically and numerically studying the stability and noise performance of our reconstruction approach,
- designing recovery algorithms tailored to the specific model that allow for non-bandlimited input functions.

REFERENCES

- [1] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, 1948.
- [2] R. H. Walden, "Analog-to-digital converter survey and analysis," *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 4, pp. 539–550, 1999.
- [3] A. Bhandari, F. Krahmer, and R. Raskar, "On unlimited sampling," in *IEEE International Conference on Sampling Theory and Applications (SampTA)*, 2017, pp. 31–35.
- [4] —, "On unlimited sampling and reconstruction," *IEEE Transactions on Signal Processing*, vol. 69, pp. 3827–3839, 2020.
- [5] 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.
- [6] Y. Zhu, R. Guo, and A. Bhandari, "Ironing the modulo folds: Unlimited sensing with 100x bandwidth expansion," in *IEEE International Conference on Sampling Theory and Applications (SampTA)*, 2025, pp. 1–5.
- [7] S. Mulletti, E. Reznitskiy, S. Savariego, M. Namer, N. Glazer, and Y. C. Eldar, "A hardware prototype of wideband high-dynamic range analog-to-digital converter," *IET Circuits, Devices & Systems*, vol. 17, no. 4, pp. 181–192, 2023.
- [8] Y. Zhu and A. Bhandari, "Unleashing dynamic range and resolution in unlimited sensing framework via novel hardware," in *IEEE SENSORS*, 2024, pp. 1–4.
- [9] J. Zhu, J. Ma, Z. Liu, F. Qu, Z. Zhu, and Q. Zhang, "A modulo sampling hardware prototype and reconstruction algorithms evaluation," *IEEE Transactions on Instrumentation and Measurement*, vol. 74, p. 2005511, 2025.
- [10] R. Guo, Y. Zhu, and A. Bhandari, "Sub-nyquist USF spectral estimation: k frequencies with $6k + 4$ modulo samples," *IEEE Transactions on Signal Processing*, vol. 72, pp. 5065–5076, 2025.
- [11] Z. Li, W. Yan, L. Gan, G. Li, and H. Liu, "An FPGA-calibrated modulo ADC for high-dynamic-range signal acquisition," *IEEE Open Journal of Circuits and Systems*, vol. 7, pp. 296–307, 2026.
- [12] S. Rudresh, A. Adiga, B. A. Shenoy, and C. S. Seelamantula, "Wavelet-based reconstruction for unlimited sampling," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2018, pp. 4584–4588.
- [13] 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.
- [14] A. Weiss, E. Huang, O. Ordentlich, and G. W. Wornell, "Blind modulo analog-to-digital conversion," *IEEE Transactions on Signal Processing*, vol. 70, pp. 4586–4601, 2022.

- [15] 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.
- [16] S. B. Shah, S. Mulleti, and Y. C. Eldar, "Lasso-based fast residual recovery for modulo sampling," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2023, pp. 1–5.
- [17] 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.
- [18] E. Azar, S. Mulleti, and Y. C. Eldar, "Unlimited sampling beyond modulo," *Applied and Computational Harmonic Analysis*, vol. 74, p. 101715, 2025.
- [19] F. P. Patricio, F. Krahmer, and P. Catala, "Stable retrieval for unlimited sampling via adaptive local representations," in *IEEE Statistical Signal Processing Workshop (SSP)*, 2025, pp. 181–185.
- [20] H. Kobayashi and R. Hayakawa, "Reconstruction of piecewise-constant sparse signals for modulo sampling," 2026, arXiv:2602.16418.
- [21] A. Bhandari and F. Krahmer, "HDR imaging from quantization noise," in *IEEE International Conference on Image Processing (ICIP)*, 2020, pp. 101–105.
- [22] C. Zhou, H. Zhao, J. Han, C. Xu, C. Xu, T. Huang, and B. Shi, "UnModNet: Learning to unwrap a modulo image for high dynamic range imaging," in *Advances in Neural Information Processing Systems (NeurIPS)*, 2020, pp. 1–12.
- [23] B. Monroy and J. Bacca, "Deep lightweight unrolled network for high dynamic range modulo imaging," *IEEE Transactions on Computational Imaging*, vol. 12, pp. 406–415, 2026.
- [24] L. G. Ordoñez, P. Ferrand, M. Duarte, M. Guillaud, and G. Yang, "On full-duplex radios with modulo-ADCs," *IEEE Open Journal of the Communications Society*, vol. 2, pp. 1279–1297, 2021.
- [25] Z. Liu, A. Bhandari, and B. Clerckx, " λ -MIMO: Massive MIMO via modulo sampling," *IEEE Transactions on Communications*, vol. 71, no. 11, pp. 6301–6315, 2023.
- [26] —, "Full-duplex beyond self-interference: The unlimited sensing way," *IEEE Communications Letters*, vol. 29, no. 1, pp. 165–169, 2025.
- [27] T. Geng, F. Ji, Pratibha, and W. P. Tay, "Modulo EEG signal recovery using transformer," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2023, pp. 1–5.
- [28] 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.
- [29] 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.
- [30] M. Beckmann and C. Dittert, "On an analytical inversion formula for the modulo Radon transform," in *Scale Space and Variational Methods in Computer Vision (SSVM)*, 2025, pp. 56–69.
- [31] 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.
- [32] T. Feuillen and A. Bhandari, "Quadrature radars naturally benefit from unlimited sensing," in *IEEE Statistical Signal Processing Workshop (SSP)*, 2025, pp. 251–255.
- [33] Q. Zhang, J. Zhu, F. Qu, and D. W. Soh, "One-bit-aided modulo sampling for DOA estimation," in *IEEE International Radar Conference (RADAR)*, 2025, pp. 1–6.
- [34] 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.
- [35] D. Florescu and A. Bhandari, "Unlimited sampling via generalized thresholding," in *IEEE International Symposium on Information Theory (ISIT)*, 2022, pp. 1606–1611.
- [36] —, "Multi-dimensional unlimited sampling and robust reconstruction," *Applied and Computational Harmonic Analysis*, vol. 79, p. 101796, 2025.
- [37] M. Beckmann and J. Jeschke, "Orthogonal matching pursuit based reconstruction for modulo hysteresis operators," in *IEEE International Conference on Sampling Theory and Applications (SampTA)*, 2025, pp. 1–5.
- [38] R. Guo and A. Bhandari, "Unlocking off-the-grid sparse recovery with unlimited sensing: Simultaneous super-resolution in time and amplitude," *IEEE Journal of Selected Topics in Signal Processing*, vol. 19, no. 8, pp. 1808–1821, 2025.
- [39] 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.
- [40] S. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Transactions on Signal Processing*, vol. 41, no. 12, pp. 3397–3415, 1993.
- [41] Y. Zhang and G. Sun, "Stagewise arithmetic orthogonal matching pursuit," *International Journal of Wireless Information Networks*, vol. 25, pp. 221–228, 2018.
- [42] A. Bhandari, "Unlimited sampling with sparse outliers: Experiments with impulsive and jump or reset noise," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2022, pp. 5403–5407.