Line spectral estimation with unlimited sensing

Wang, Hongwei; Fang, Jun; Li, Hongbin; Leus, Geert; Zhu, Ruixiang; Gan, Lu

DOI
10.1016/j.sigpro.2025.110205

Publication date
2026

Document Version
Final published version

Published in
Signal Processing

Signal Processing 238 (2026) 110205

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Line spectral estimation with unlimited sensing ☆

Hongwei Wang [a,*], Jun Fang [a], Hongbin Li [b], Geert Leus [c], Ruixiang Zhu [d], Lu Gan [e]

[a] National Key Laboratory of Wireless Communications, University of Electronic Science and Technology of China, Chengdu 611731, China
[b] Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, NJ 07030, USA
[c] Faculty of Electrical Engineering, Mathematics and Computer Science, Delft University of Technology, 2826 CD Delft, The Netherlands
[d] Department of Electronic and Electrical Engineering, University College London, London, WC1E 6BT, UK
[e] Department of Electronic and computer Engineering, Brunel University, Uxbridge UB8 3PH, UK

ARTICLE INFO

Keywords:
Unlimited sensing
Line spectral estimation
Modulo samples

ABSTRACT

In the paper, we consider the line spectral estimation problem in an unlimited sensing framework (USF), where a modulo analog-to-digital converter (ADC) is employed to fold the input signal back into a bounded interval before quantization. Such an operation is mathematically equivalent to taking the modulo of the input signal with respect to the interval. To overcome the noise sensitivity of higher-order difference-based methods, we explore the properties of the first-order difference of modulo samples, and develop two line spectral estimation algorithms based on the first-order difference, which are robust against noise. Specifically, we show that, with a high probability, the first-order difference of the original samples is equivalent to that of the modulo samples. By utilizing this property, line spectral estimation is solved via a robust sparse signal recovery approach. The second algorithms is built on our finding that, with a sufficiently high sampling rate, the first-order difference of the original samples can be decomposed as a sum of the first-order difference of the modulo samples and a sequence whose elements are confined to three possible values. This decomposition enables us to formulate the line spectral estimation problem as a mixed integer linear program that can be efficiently solved. Simulation results show that both proposed methods are robust against noise and achieve a significant performance improvement over the higher-order difference-based method. methods.

1. Introduction

Line spectral estimation (LSE) is a fundamental problem in statistical signal processing, which is aimed to detect and identify the components of a mixture of complex sinusoids from a finite number of samples. This problem can be found in numerous applications, such as direction of arrival estimation in array signal processing [1–3], bearing and range estimation in radar signal processing [4,5], speech analysis in acoustic speech signal processing [6], and many others [7]. Traditional LSE solutions include Prony's method [8], MUltiple SIgnal Classification (MUSIC) [9,10], and Estimation of Signal Parameters via Rotational Invariance Techniques (ESPRIT) [11]. In addition, the LSE problem can be framed as a sparse recovery problem and solved using the compressed sensing approach [12].

In general, samples of continuous signals are obtained via an analog-to-digital converter (ADC). Conventional ADCs, however, suffer from a fundamental bottleneck due to their limited dynamic range (DR). When the amplitude of the input signal exceeds the DR of a conventional ADC, clipping/saturation occurs, resulting in distortion. In such cases, traditional LSE methods experience considerable performance degradation or may even fail [13]. To avoid clipping and saturation, automatic gain control is usually employed at the receiver to maintain a constant output signal level. Nevertheless, due to the limited DR, a weak signal could be buried beneath the quantization noise in the presence of a strong signal.

Recently, unlimited sensing was proposed to address the limited DR issue of conventional sampling systems [14–17]. The potential applications of the unlimited sensing include the high dynamic range imaging [18], wireless communications [19], radar systems [20], and many others. In an unlimited sensing framework, a modulo ADC is employed to map the input signal to a bounded interval $[-\lambda, \lambda]$ before quantization. Such an operation is mathematically equivalent to taking the modulo of the input signal with respect to λ, which eliminates the saturation problem of conventional sampling systems. However, the modulo operation, which is a nonlinear mapping from the input to

☆ This research was supported by the National Natural Science Foundation of China under Grants No. 62103083. The work of H. Li was supported in part by the National Science Foundation under Grants No. CCF-2316865, ECCS-2212940, and ECCS-2332534.

* Corresponding author.
E-mail addresses: tianhangxinxiang@163.com, hongwei_wang@uestc.edu.cn (H. Wang).

https://doi.org/10.1016/j.sigpro.2025.110205
Received 1 December 2024; Received in revised form 9 July 2025; Accepted 11 July 2025
Available online 23 July 2025

the output, introduces a different type of information loss. Therefore, how to recover the original input signal or infer the parameters of interest from modulo samples is crucial to realizing the potential of the unlimited sensing approach.

The above problem has been extensively investigated over the past few years. In [15], it was shown that, for a bandlimited signal with a maximum frequency of ω, a sufficient condition for the recovery of the input signal from its modulo samples is that the sampling interval should be no greater than $1/(2\omega e)$, where e is Euler's number. Provided such a sampling condition is met, there exists a minimal difference order, such that the higher-order difference (HOD) of the original signal is equivalent to the modulo of the HOD of the modulo samples. Based on this relationship, the original signal can be recovered (up to a scalar ambiguity) via a higher-order anti-difference operation. This theoretical result represents a breakthrough for data acquisition in the unlimited sensing framework, and has inspired many applications in sparse signal recovery [21], sinusoidal estimation [22], and computational array signal processing [23,24]. The unlimited sensing is also integrated with the one-bit sampling framework [13,25,26], which aims to achieve a high sampling rate, unlimited DR, and less complex, lowpower implementations. In addition, unlimited sensing with non-ideal modulo samples [27], modulo samples with hysteresis [28,29], as well as non-uniformly sampled modulo samples [30] were also studied. Notably, a hardware realization of the modulo ADC was presented in [27] to validate the performance of unlimited sensing with non-ideal circuits. The HOD-based methods, however, are very sensitive to noise. The reason is that the HOD operation shrinks the signal, whereas it has an accumulative effect on the noise, as the noise has a much wider bandwidth than the signal. As a result, even with a moderate difference order, the effective signal-to-noise ratio (SNR) degrades substantially, thus leading to a deteriorated performance. In addition, [17] first applied the modulo ADC into a frequency-modulated continuous wave radar system, demonstrating its advantages for ranging estimation in scenarios where strong and weak targets coexist.

In this paper, we explore the properties of the first-order difference of modulo samples and develop two noise-robust LSE algorithms. Specifically, we prove that, with a sufficiently high sampling rate, the first-order difference of the original signal samples equals that of the modulo samples with a high probability. Based on this theoretical result, our first proposed algorithm recasts LSE as a robust sparse signal recovery problem. The LSE problem with modulo samples is also considered in [22]. The main difference between our work and [22] is that we only utilize the first-order difference of modulo samples, whereas the work [22] depends on the HOD of the modulo samples to extract the spectral parameters. Note that the sparse property was also observed in [27] and utilized to recover the original signal. Specifically, the solution in [27] processes the signal in the Fourier domain. However, it requires knowledge of both the number of impulsive components and the period of the original signal, which is generally unknown in practice. In contrast, our solution is developed using a variational Bayesian approach, which not only eliminates the requirement for this knowledge but also effectively accounts for observation noise. In [31], a data-driven method was proposed to estimate the number of impulsive components. For the second proposed algorithm, we show that, with a sufficiently high sampling rate, the first-order difference of the original signal can be decomposed as a sum of the first-order difference of the modulo samples and a sequence with each element being either zero of $\pm 2\lambda$. This property enables us to cast the problem into a mixed-integer linear program which can be efficiently solved. Simulation results show that both algorithms are robust against noise and achieve significant performance improvement over the HOD-based method.

The remainder of the work is organized as follows. In Section 2, preliminaries about the modulo ADC and higher-order difference are introduced. LSE in the unlimited sensing framework is formulated in Section 3. Section 4 provides an overview of the HOD-based approach to the problem. First-order difference-based methods are then developed in Section 5 and Section 6. Simulation results are presented in Section 7, followed by concluding remarks in Section 8.

2. Preliminaries

The folding operation can be mathematically expressed as the following nonlinear mapping

$$\mathcal{M}_\lambda : f \mapsto 2\lambda \left(\left\langle \frac{f}{2\lambda} + \frac{1}{2} \right\rangle - \frac{1}{2} \right) \tag{1}$$

where $\langle a \rangle \triangleq a - \lfloor a \rfloor$ is the fractional part of a and $\lambda \geq 0$ is the operation range of the modulo ADC. Clearly, such an operation converts an arbitrary value f into the interval $[-\lambda, \lambda]$.

Define the first-order difference matrix as $\boldsymbol{D}_M^1 \in \mathbb{R}^{(M-1)\times M}$, which is given by

$$\boldsymbol{D}_M^1 = \begin{pmatrix} -1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -1 & 1 \end{pmatrix}. \tag{2}$$

Mathematically, the (i,j)th entry of $\boldsymbol{D}_M^1$ is obtained via $\delta[j-i-1]-\delta[j-i]$ where $\delta[\cdot]$ is the Kronecker Delta function. Given a vector $\boldsymbol{x} \in \mathbb{C}^M$, its first-order difference vector $\tilde{\boldsymbol{x}}$ can be obtained as

$$\tilde{\boldsymbol{x}} = \boldsymbol{D}_M^1 \boldsymbol{x} \in \mathbb{C}^{M-1} \tag{3}$$

where $\tilde{\boldsymbol{x}}(i) = \boldsymbol{x}(i+1) - \boldsymbol{x}(i)$. Accordingly, the Nth-order ($N > 1$) difference matrix $\boldsymbol{D}_M^N$ can be recursively constructed as

$$\boldsymbol{D}_M^N = \boldsymbol{D}_{M-N+1}^1 \boldsymbol{D}_M^{N-1}. \tag{4}$$

The Nth-order difference vector of $\boldsymbol{x}$ can be calculated as

$$\tilde{\boldsymbol{x}}^{(N)} = \boldsymbol{D}_M^N \boldsymbol{x} \in \mathbb{C}^{M-N}. \tag{5}$$

Note that we have $\tilde{\boldsymbol{x}} = \tilde{\boldsymbol{x}}^{(1)}$ for simplicity in the following.

3. Problem formulation

Consider the following sum of K frequency components

$$y_m = \sum_{k=1}^{K} \alpha_k e^{-j\omega_k m \Delta T} \tag{6}$$

where m is the sampling index, ΔT is the sampling interval, $\omega_k \in [0, 2\pi)$ and α_k are, respectively, the frequency and the complex amplitude of the kth component. The LSE problem of interest is to obtain an estimate of $\{\alpha_k, \omega_k\}$ from measurements $\{y_m\}_{m=1}^M$. In practical sampling systems, due to the limited DR of conventional ADCs, some of the measurements $\{y_m\}$ may be clipped, leading to severe information loss. To address this difficulty, this work considers the LSE problem with modulo ADCs. For modulo ADCs, when the input signal magnitude exceeds a certain threshold λ, it will reset the signal such that it is folded back to the range $[-\lambda, \lambda]$. Mathematically, it is equivalent to taking the remainder of the division of y_m by λ:

$$z_m = \mathcal{U}_\lambda(y_m) + v_m \tag{7}$$

where $v_m \sim \mathcal{CN}(0, \sigma^2)$ is the complex additive white Gaussian noise, and $\mathcal{U}_\lambda$ is the modulo operation performed on a complex value which is defined as

$$\mathcal{U}_\lambda(a) \triangleq \mathcal{M}_\lambda\big(\mathrm{Re}(a)\big) + j\mathcal{M}_\lambda\big(\mathrm{Im}(a)\big) \tag{8}$$

in which $\mathrm{Re}(a)$ and $\mathrm{Im}(a)$ respectively denote the real and imaginary parts of a complex number a.

To estimate $\{\omega_k, \alpha_k\}$, we discretize the continuous frequency parameter space into a finite set of grid points, say P ($P \gg K$) points in total. Define $\boldsymbol{a}_p \triangleq [e^{-j\omega_p \Delta T} \ \cdots \ e^{-j\omega_p M \Delta T}]^T$, $\boldsymbol{A} \triangleq [\boldsymbol{a}_1 \ \cdots \ \boldsymbol{a}_P]$, $\boldsymbol{\alpha} \triangleq [\alpha_1 \ \cdots \ \alpha_P]^T$, $\boldsymbol{y} \triangleq [y_1 \ \cdots \ y_M]^T$, $\boldsymbol{z} \triangleq [z_1 \ \cdots \ z_M]^T$, and $\boldsymbol{v} \triangleq [v_1 \ \cdots \ v_M]^T$. We can then rewrite (7) in a matrix form as

$$\begin{aligned} \boldsymbol{z} &= \mathcal{U}_\lambda(\boldsymbol{y}) + \boldsymbol{v} \\ &= \mathcal{U}_\lambda(\boldsymbol{A}\boldsymbol{\alpha}) + \boldsymbol{v} \end{aligned} \tag{9}$$

where $\boldsymbol{\alpha}$ is a K-sparse vector. Note that in this work, we assume that the unknown frequency components lie on the discretized grid.

Now the problem becomes estimating the sparse vector $\boldsymbol{\alpha}$ from $\boldsymbol{z}$. Such a problem can be formulated as

$$\begin{aligned} \min_{\boldsymbol{\alpha}} \quad & \|\boldsymbol{\alpha}\|_0 \\ \text{s.t.} \quad & \|\boldsymbol{z} - \mathcal{U}_\lambda(\boldsymbol{A}\boldsymbol{\alpha})\|_2 \le \epsilon \end{aligned} \tag{10}$$

where $\|\cdot\|_0$ is the ℓ_0 norm which counts the number of nonzero components, and ϵ is a user-defined error tolerance. The optimization problem in (10) features both the ℓ_0 norm and a nonlinear constraint. The presence of the ℓ_0 norm renders this optimization problem NP-hard. Additionally, the nonlinear constraint, which arises from the modulo operation, further complicates the problem.

Remark 1. In this work, we focus on scenarios where the frequencies align with a pre-defined discretized grid. However, LSE with off-grid frequencies is commonly encountered in real applications. Devising an algorithm to address the off-grid problem is challenging, even for ideally sampled observations. The complexity further escalates when dealing with modulo observations, which introduce nonlinearity. We transform the LSE with modulo samples into a robust compressive sensing problem using a first-order difference operation. Within the robust compressive sensing framework, several existing approaches, e.g., [32,33], provide valuable insights for advancing off-grid LSE methods to modulo samples.

4. Higher-order difference-based approach

In this section, we introduce a HOD-based approach to address the LSE problem. In [15], it was shown that when the sampling rate is sufficiently high, the HOD (greater than a certain order) of the original signal $\boldsymbol{y}$ can be obtained from the HOD of modulo samples $\boldsymbol{z}$. Based on this result, the original signal can be recovered up to an unknown constant via an anti-difference operation.

Define $\omega \triangleq \max\{\omega_1, \ldots, \omega_K\}$ and $B \triangleq \|\boldsymbol{y}\|_\infty$. The main result in [15] is summarized as follows.

Theorem 1. *Let y_m be samples of $y(t) \in \mathcal{B}_\omega$ obtained with a sampling interval of ΔT, where $\mathcal{B}_\omega$ represents the ω-bandlimited function set. Let e denote Euler's number and*

$$\boldsymbol{z} = \mathcal{U}_\lambda(\boldsymbol{y}) \tag{11}$$

denote the noise-free modulo sample vector. If the sampling interval ΔT satisfies

$$\Delta T \le \frac{1}{2\omega e} \tag{12}$$

and meanwhile the order of difference N is no smaller than

$$N \ge \left\lceil \frac{\log \lambda - \log B}{\log(\Delta T \omega e)} \right\rceil \tag{13}$$

then the Nth-order difference of the original signal $\boldsymbol{y}$ can be obtained from the Nth-order difference of its modulo samples via the following relationship:

$$\boldsymbol{D}_M^N \boldsymbol{y} = \mathcal{U}_\lambda(\boldsymbol{D}_M^N \boldsymbol{z}) \tag{14}$$

where $\boldsymbol{D}_M^N$ is the Nth-order difference matrix.

Remark 2. For finite-length sequences, besides the above requirement on the sampling rate, the number of samples should also satisfy a certain condition in order to ensure that Theorem 1 holds [21]. Specifically, based on the result reported in [21], to recover sinusoidal mixtures, the required number of modulo samples should be no less than $2K + 7B/\lambda$. This requirement is mild and can usually be met in practice.

Utilizing the above theorem and ignoring the noise, the optimization (10) can be reformulated into a conventional compressed sensing problem:

$$\begin{aligned} \min_{\boldsymbol{\alpha}} \quad & \|\boldsymbol{\alpha}\|_0 \\ \text{s.t.} \quad & \mathcal{U}_\lambda(\boldsymbol{D}_M^N \boldsymbol{z}) = \boldsymbol{D}_M^N \boldsymbol{A}\boldsymbol{\alpha} \end{aligned} \tag{15}$$

which can be efficiently solved by off-the-shelf compressed sensing solutions, e.g., the orthogonal matching pursuit method [34]. In practice, the modulo samples $\boldsymbol{z}$ are inevitably corrupted by noise due to quantization errors and thermal noise. In this case, we have

$$\mathcal{U}_\lambda(\boldsymbol{D}_M^N \boldsymbol{z}) = \boldsymbol{D}_M^N \boldsymbol{A}\boldsymbol{\alpha} + \boldsymbol{w} \tag{16}$$

where $\boldsymbol{w}$ represents the error induced by the measurement noise $\boldsymbol{v}$. The problem is that the HOD operation has a shrinkage effect on the signal, whereas it has an accumulative effect on the noise as the noise has a much wider bandwidth than the signal [13,27]. As a result, even with a moderate vale of N, say $N = 3$, the effective signal-to-noise ratio (SNR) degrades substantially, thus leading to a deteriorated recovery performance.

For demonstration purposes, we consider a simple example

$$y(t) = 7\sin(1.6\pi t) \tag{17}$$

$$z_m = \mathcal{U}_\lambda(y_m) + v_m \tag{18}$$

where $\lambda = 1$. In our experiments, we set the sampling interval $\Delta T = 0.0362$ s such that the required conditions in Theorem 1 are satisfied. Fig. 1 shows the histogram of w_m (i.e., the components in $\boldsymbol{w}$) under different choices of N when the SNR is set to 20 dB. We see that the error $\boldsymbol{w}$ induced by the measurement noise becomes significantly larger as N increases.

To overcome this difficulty caused by the higher-order difference, in the following, we explore the properties of the first-order difference of modulo samples, and develop algorithms that are robust against noise.

5. Robust compressive sensing based approach

This proposed approach is based on the observation that, although the first-order difference of the modulo samples is not exactly the same as the first-order difference of the original signal samples, these two indeed coincide with a high probability when the sampling rate is sufficiently high. To illustrate this, consider the following example

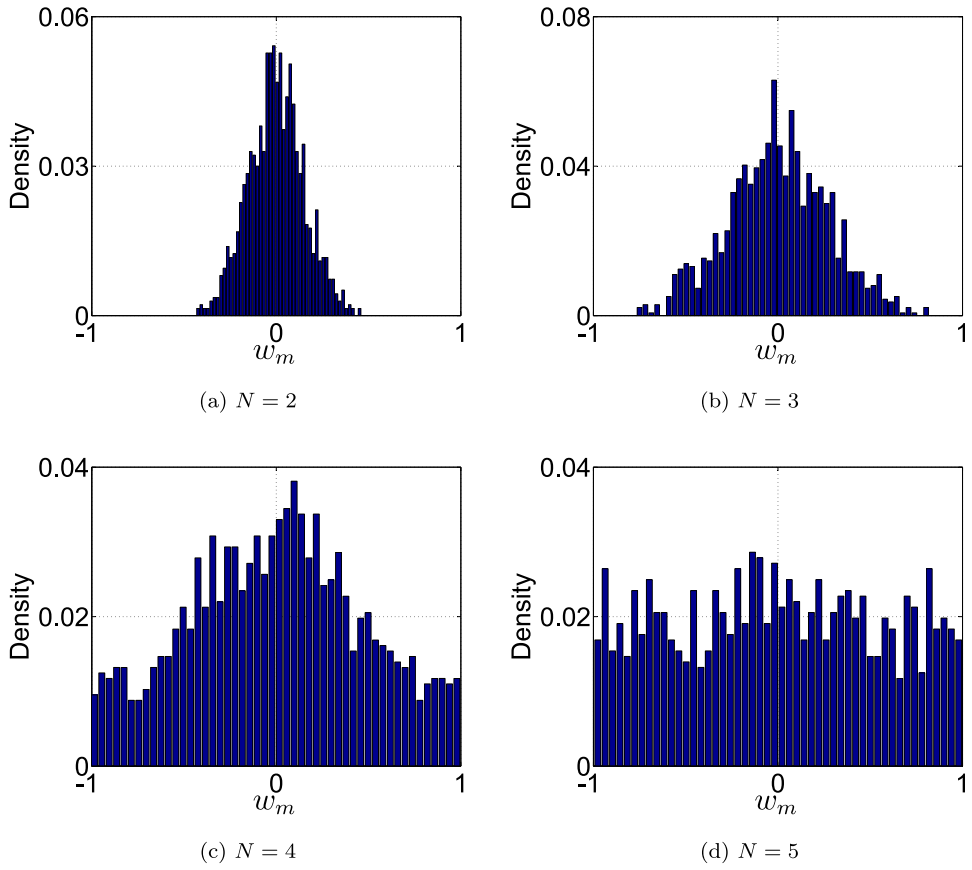
$$y(t) = 4\sin(\pi t/6) - 2\sin(\pi t/3) + 2\sin(2\pi t - 0.2t) \tag{19}$$

Setting the sampling interval to 0.05 s, Fig. 2(a) shows the undistorted original samples, the samples obtained by a conventional ADC, and the modulo samples obtained by a modulo ADC. The measurable ranges of the conventional ADC and modulo ADC are set to $[-1, 1]$. In Fig. 2(b), the first-order difference of the original samples and that of the modulo samples are plotted. We see that most of the entries are identical. An intuitive explanation of this fact is that when the sampling rate is high, the original signal undergoes slow variations between successive samples. As a result, the first-order difference of the original signal samples is highly like to be equal to that of the modulo samples.

In fact, those inconsistent samples can be considered as outliers. Thus the LSE problem can be reformulated as a robust sparse signal recovery problem. First, we formulate the first-order difference of modulo samples as

$$\begin{aligned} \tilde{\boldsymbol{z}} &\triangleq \boldsymbol{D}_M^1 \boldsymbol{z} \\ &= \boldsymbol{D}_M^1 \mathcal{U}_\lambda(\boldsymbol{A}\boldsymbol{\alpha}) + \boldsymbol{D}_M^1 \boldsymbol{v} \\ &= \boldsymbol{D}_M^1 \boldsymbol{A}\boldsymbol{\alpha} + \boldsymbol{s} + \boldsymbol{D}_M^1 \boldsymbol{v} \\ &\triangleq \tilde{\boldsymbol{A}}\boldsymbol{\alpha} + \tilde{\boldsymbol{v}} + \boldsymbol{s} \end{aligned} \tag{20}$$

where $\tilde{\boldsymbol{A}} \triangleq \boldsymbol{D}_M^1 \boldsymbol{A}$, $\tilde{\boldsymbol{v}} \triangleq \boldsymbol{D}_M^1 \boldsymbol{v}$, and $\boldsymbol{s}$ is an unknown, sparse outlier vector. Note that the above equation comes from the fact that when



(a) $N = 2$ (b) $N = 3$

(c) $N = 4$ (d) $N = 5$

Fig. 1. The histogram of $\{w_m\}$ under different N.

Algorithm 1 RCS-based algorithm for LSE problem with modulo samples

Input: $\boldsymbol{z}$ and M.
Output: $\boldsymbol{\alpha}$ and the reconstructed original signal $\hat{\boldsymbol{y}}$.
1) Construct $\boldsymbol{A}$ and $\boldsymbol{D}_M^1$;
2) Compute $\tilde{\boldsymbol{A}} = \boldsymbol{D}_M^1 \boldsymbol{A}$ and $\tilde{\boldsymbol{z}} = \boldsymbol{D}_M^1 \boldsymbol{z}$;
3) Estimate $\boldsymbol{\alpha}$ via a robust compressive sensing method (e.g., [41]);
4) Reconstruct $\hat{\boldsymbol{y}} = \boldsymbol{A}\boldsymbol{\alpha}$.

the sampling rate is sufficiently high, the first-order difference of the modulo samples can be expressed as a sum of the first-order difference of original signal samples and a sparse outlier vector, i.e.

$$\boldsymbol{D}_M^1 \mathcal{U}_\lambda(\boldsymbol{A}\boldsymbol{\alpha}) = \boldsymbol{D}_M^1 \boldsymbol{A}\boldsymbol{\alpha} + \boldsymbol{s} \tag{21}$$

The problem of estimating $\boldsymbol{\alpha}$ from (20) is a classic robust sparse signal recovery problem which has been extensively investigated, e.g. [35–41]. Specifically, a theoretical analysis conducted in [36] shows that under certain conditions, exact recovery of the sparse vector $\boldsymbol{\alpha}$ from sparse outlier-corrupted measurements can be guaranteed. In addition to theoretical justifications, many robust compressed sensing algorithms were developed from either optimization-based or Bayesian inference frameworks. Popular algorithms include the outlier-compensation-based
approach [38,39], and outlier identify-and-reject approach [41]. Algorithm 1 summarizes the proposed robust compressive sensing (RCS) based solution for the LSE problem with modulo samples.

Clearly, exact recovery of the sparse signal $\boldsymbol{\alpha}$ is feasible when most measurements of $\tilde{\boldsymbol{z}}$ are not outliers, in other words, when $\boldsymbol{s}$ is sufficiently sparse. Intuitively, given other parameters fixed, the sparsity of the outlier vector $\boldsymbol{s}$ is related to the sampling rate: the higher the sampling rate, the sparser the outlier vector. In the following, we provide a quantitative analysis which establishes a mathematical relationship between the sampling rate and the sparsity of the outlier vector $\boldsymbol{s}$ obtained via the first-order difference operation.

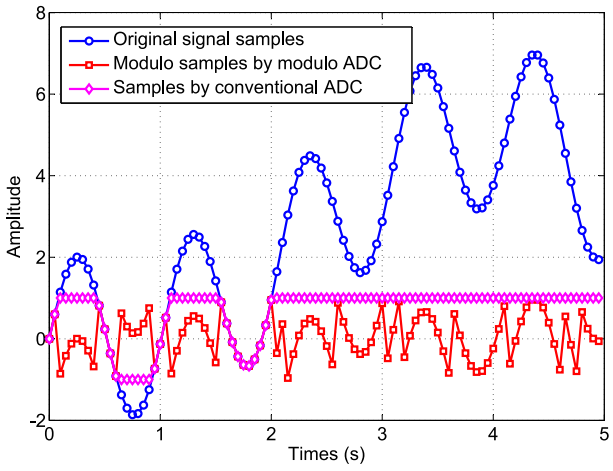
Remark 3. We formulate the line spectral estimation with modulo observations as a robust compressive sensing problem. The characteristics of the corresponding sensing matrix are crucial for sparse signal recovery. The restricted isometry property (RIP) and the mutual coherence are two common metrics for evaluating the sensing matrix's properties. However, verifying RIP for generic dictionaries is typically computationally intractable. Consequently, we focus on mutual coherence to analyze the sensing matrix's properties in (20). According to [42,43], for any dictionary $\boldsymbol{D}$ and a sparse representation of a given signal $\boldsymbol{y}_0$ ($\boldsymbol{y}_0 = \boldsymbol{D}\boldsymbol{d}$), if the inequality

$$\|\boldsymbol{d}\|_0 < \frac{1}{2}\left(1 + \frac{1}{\mu(\boldsymbol{D})}\right), \tag{22}$$

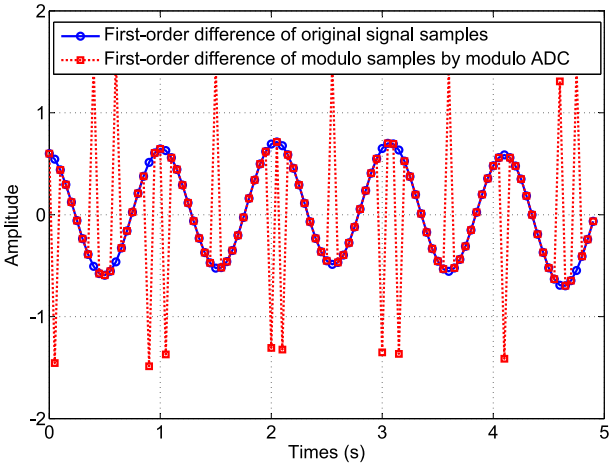
is satisfied where $\mu(\boldsymbol{D})$ is the mutual coherence of $\boldsymbol{D}$, then $\boldsymbol{d}$ is the unique solution to both the ℓ_0 and ℓ_1 minimization problems. Taking $M = 400$ and $P = 20$ as an example (the same setup as in our simulations), we find that when $\Delta T \geq 0.045$, the mutual coherence μ of $\tilde{\boldsymbol{A}}$ is consistently less than 0.1. Under such a circumstance, we can assert that when $\|\boldsymbol{\alpha}\|_0 < 5.5$, $\boldsymbol{\alpha}$ is a unique solution of the optimization problem

$$\begin{aligned} &\min \ \|\boldsymbol{\alpha}\|_0 \ (\text{or } \min \|\boldsymbol{\alpha}\|_1) \\ &\text{s.t.} \quad \tilde{\boldsymbol{z}} = \tilde{\boldsymbol{A}}\boldsymbol{\alpha} + \boldsymbol{s} \end{aligned} \tag{23}$$

provided that the outlier vector $\boldsymbol{s}$ is correctly detected/estimated and the noise term $\tilde{\boldsymbol{v}}$ is absent. In this study, we employ a Bayesian inference approach to address the robust signal recovery problem. The simulation



(a) The original undistorted samples, samples obtained via the conventional ADC and modulo samples obtained via the modulo ADC



(b) The-first order differences of the original samples and modulo samples

Fig. 2. First-order differences of original samples and modulo samples.

results demonstrate that the considered algorithms can effectively estimate the sparse signal $\boldsymbol{\alpha}$ under certain conditions in the presence of both outliers and noises.

5.1. Theoretical analysis

The line spectral model in (6) can be re-expressed as

$$y_m = \sum_{k=1}^{K} |\alpha_k| \left(\cos(\omega_k m\Delta T - \phi_k) - j\sin(\omega_k m\Delta T - \phi_k)\right) \triangleq f_m^r - j f_m^i \tag{24}$$

where $\phi_k = \angle\alpha_k$ denotes the argument of the complex parameter α_k, f_m^r and f_m^i are, respectively, the real and imaginary parts of y_m which are given as

$$f_m^r = \sum_{k=1}^{K} |\alpha_k| \cos(\omega_k m\Delta T - \phi_k) \tag{25}$$

$$f_m^i = \sum_{k=1}^{K} |\alpha_k| \sin(\omega_k m\Delta T - \phi_k) = \sum_{k=1}^{K} |\alpha_k| \cos(\omega_k m\Delta T - \psi_k) \tag{26}$$

where $\psi_k = \phi_k + \pi/2$. We see that the two sinusoidal mixture functions f_m^r and f_m^i share the same frequency components with different phases. To analyze the relationship between the first-order difference of y_m and the first-order difference of its modulo samples, if suffices to examine the first-order difference of either its real component or its imaginary component.

Specifically, consider the following sinusoidal mixture

$$f(t) = \sum_{k=1}^{K} \beta_k \cos(\omega_k t + \varphi_k) \tag{27}$$

where $\varphi_k \in (0, 2\pi]$, $\omega_k \geq 0$ and $\beta_k \in \mathbb{R}$ are, respectively, the phase, frequency and amplitude of the kth component. Define $\omega \triangleq \max\{\omega_1, \ldots, \omega_K\}$, $\bar{\beta} \triangleq \sum_{k=1}^{K} |\beta_k \bar{\omega}_k|$, and $\bar{\omega}_k = \omega_k/\omega$. Let $f_m = f(m\Delta T)$ and $x_m = \mathcal{M}_\lambda(f_m)$ respectively denote the original sample and the modulo sample of $f(t)$, where $\Delta T \in \mathbb{R}^+$ is the sampling interval. Let $\tilde{f}_m \triangleq f_{m+1} - f_m$ denote the first-order difference of the original samples and $\tilde{x}_m \triangleq x_{m+1} - x_m$ denote the first-order difference of the modulo samples.

According to (1), the modulo sample of f_m can be written as

$$x_m = 2\lambda \left(\frac{f_m}{2\lambda} - \left\lfloor \frac{f_m}{2\lambda} + \frac{1}{2} \right\rfloor \right) \tag{28}$$

The above equation can be re-expressed as

$$f_m = 2\lambda \left\lfloor \frac{f_m}{2\lambda} + \frac{1}{2} \right\rfloor + x_m \tag{29}$$

Here f_m is a sum of two terms, where the first term is the quantization output of f_m that is quantized by a step size of 2λ, and the second term x_m can be considered as the quantization noise.

In this paper, we make the following widely-used assumption [see 7, Eq. (4.1.2)], [44–46]:

Assumption 1. The phase of each component in $f(t)$ is a random variable which follows a uniform distribution, i.e.,

$$p(\varphi_k) = \begin{cases} \frac{1}{2\pi}, & \varphi_k \in (-\pi, \pi] \\ 0, & \text{otherwise} \end{cases} \tag{30}$$

Under such an assumption, x_m can be modeled as a uniform random variable when 2λ is sufficiently small:

$$x_m \sim \mathbb{U}(-\lambda, \lambda) \tag{31}$$

A rigorous justification of the uniform distribution of x_m is discussed in Appendix A.

We are now in a position to present our main result, which is summarized as follows.

Theorem 2. *Consider a sinusoidal mixture signal $f(t)$ which has a form of (27), where the phase of each sinusoidal component follows a uniform distribution $\mathbb{U}(-\pi, \pi)$. If the sampling interval satisfies*

$$\Delta T \leq \frac{1}{2\omega} \left(\frac{\bar{\beta}}{2\lambda} \right)^{-1} \tag{32}$$

then the first-order difference of the original samples equals the first-order difference of the modulo samples with probability exceeding

$$Pr(\tilde{f}_m = \tilde{x}_m) \triangleq p \geq 1 - \frac{\Delta T \omega \bar{\beta}}{2\lambda} \tag{33}$$

where $\bar{\beta}$ is defined as

$$\bar{\beta} \triangleq \sum_{k=1}^{K} |\beta_k \bar{\omega}_k|$$

with $\bar{\omega}_k = \omega_k / (\max\{\omega_1, \ldots, \omega_K\})$.

Proof. See Appendix B. □

Theorem 2 indicates that the probability of $\tilde{f}_m = \tilde{x}_m$ can be made arbitrarily close to 1 by increasing the sampling rate $1/\Delta T$. Specifically, to make sure that p is no less than a pre-specified threshold τ, i.e., $p \geq \tau$, the sampling interval needs to satisfy

$$\Delta T \leq (1-\tau)\frac{2\lambda}{\omega\bar{\beta}} \triangleq \frac{1}{2\omega\eta_1} \tag{34}$$

where η_1 is defined as

$$\eta_1 \triangleq \frac{1}{2(1-\tau)}\bar{N} \tag{35}$$

with $\bar{N} \triangleq \bar{\beta}/(2\lambda)$ denoting the effective number of folding times. We see that the sampling condition in (34) has a form similar to the condition in (12), except that they have different scaling factors.

6. Mixed-integer linear programming based recovery approach

6.1. Motivations and algorithm development

It is known that the original sample f_m and the modulo sample x_m satisfy the unique decomposition property: $x_m = f_m + 2\lambda e_m$, where e_m is an integer. Similarly, it can be readily verified that the first-order difference of the original samples and the first-order difference of the modulo samples are related as $\tilde{x}_m = \tilde{f}_m + 2\lambda\tilde{e}_m$, where $\tilde{e}_m$ is an integer. In the previous section, we observed and proved that given a sufficiently high sampling rate, $\{\tilde{e}_m\}$ is a sparse vector with only few nonzero entries. Thus the LSE problem can be cast as a robust sparse signal recovery problem. This formulation, however, fails to utilize the information that $\tilde{e}_m$ is an integer.

In fact, considering the property that e_m is an integer, the LSE problem with noise-free $\boldsymbol{z}$ can be formulated as a combinatorial problem [47,48]

$$\begin{aligned} \min_{\boldsymbol{\alpha},\boldsymbol{e}} \quad & \|\boldsymbol{\alpha}\|_1 \\ \text{s.t.} \quad & \boldsymbol{A}\boldsymbol{\alpha} = \boldsymbol{z} + 2\lambda\boldsymbol{e} \\ & \mathrm{Re}(\boldsymbol{e}) \in \mathbb{Z}^M,\ \mathrm{Im}(\boldsymbol{e}) \in \mathbb{Z}^M \end{aligned} \tag{36}$$

where $\mathbb{Z}^M$ represents the M-dimensional integer space. Such an optimization problem can be transformed into a mixed integer linear program and solved via the branch-and-bound algorithm [49]. This approach, however, is not suitable when the folding number $\lfloor f_m/(2\lambda) + 1/2 \rfloor$ is large as the search space increases exponentially with the folding number.

To reduce the search space, in this section, we take the first-order difference of modulo samples. Specifically, we have

$$\begin{aligned} \tilde{\boldsymbol{z}} &\triangleq \boldsymbol{D}_M^1 \boldsymbol{z} \\ &= \boldsymbol{D}_M^1 (\mathcal{U}_\lambda(\boldsymbol{A}\boldsymbol{\alpha}) + \boldsymbol{v}) \\ &= \boldsymbol{D}_M^1 (\boldsymbol{A}\boldsymbol{\alpha} + 2\lambda\boldsymbol{e} + \boldsymbol{v}) \\ &= \boldsymbol{D}_M^1 \boldsymbol{A}\boldsymbol{\alpha} + 2\lambda\boldsymbol{D}_M^1\boldsymbol{e} + \boldsymbol{D}_M^1\boldsymbol{v} \\ &\triangleq \tilde{\boldsymbol{A}}\boldsymbol{\alpha} + 2\lambda\tilde{\boldsymbol{e}} + \tilde{\boldsymbol{v}} \end{aligned} \tag{37}$$

where $\tilde{\boldsymbol{e}} \triangleq \boldsymbol{D}_M^1\boldsymbol{e}$ is an unknown vector with $\mathrm{Re}(\tilde{\boldsymbol{e}}) \in \mathbb{Z}^{M-1}$ and $\mathrm{Im}(\tilde{\boldsymbol{e}}) \in \mathbb{Z}^{M-1}$. We will show that when the sampling rate exceeds a certain threshold, both the real and imaginary components of entries in $\tilde{\boldsymbol{e}}$ belong to the set $\{0, \pm 1\}$. Thus we can formulate the LSE problem as

$$\begin{aligned} \min_{\boldsymbol{\alpha},\tilde{\boldsymbol{e}}} \quad & \|\boldsymbol{\alpha}\|_1 \\ \text{s.t.} \quad & \|\tilde{\boldsymbol{z}} - \tilde{\boldsymbol{A}}\boldsymbol{\alpha} + 2\lambda\tilde{\boldsymbol{e}}\|_2^2 \leq \epsilon \\ & \mathrm{Re}(\tilde{\boldsymbol{e}}) \in \{0, \pm 1\}^{(M-1)} \\ & \mathrm{Im}(\tilde{\boldsymbol{e}}) \in \{0, \pm 1\}^{(M-1)} \end{aligned} \tag{38}$$

where ϵ is a user-defined parameter. For convenience, we write (38) into a real-valued form, i.e.,

$$\begin{aligned} \min_{\check{\boldsymbol{\alpha}},\check{\boldsymbol{e}}} \quad & \|\check{\boldsymbol{\alpha}}\|_1 \\ \text{s.t.} \quad & \|\check{\boldsymbol{z}} - \check{\boldsymbol{A}}\check{\boldsymbol{\alpha}} + 2\lambda\check{\boldsymbol{e}}\|_2^2 \leq \epsilon \\ & \check{\boldsymbol{e}} \in \{0, \pm 1\}^{2(M-1)} \end{aligned} \tag{39}$$

where $\check{\boldsymbol{\alpha}} \triangleq [\mathrm{Re}(\boldsymbol{\alpha})^T\ \mathrm{Im}(\boldsymbol{\alpha})^T]^T$, $\check{\boldsymbol{e}} \triangleq [\mathrm{Re}(\tilde{\boldsymbol{e}})^T\ \mathrm{Im}(\tilde{\boldsymbol{e}})^T]^T$, $\check{\boldsymbol{z}} \triangleq [\mathrm{Re}(\tilde{\boldsymbol{z}})^T\ \mathrm{Im}(\tilde{\boldsymbol{z}})^T]^T$, and

$$\check{\boldsymbol{A}} \triangleq \begin{bmatrix} \mathrm{Re}(\tilde{\boldsymbol{A}}) & -\mathrm{Im}(\tilde{\boldsymbol{A}}) \\ \mathrm{Im}(\tilde{\boldsymbol{A}}) & \mathrm{Re}(\tilde{\boldsymbol{A}}) \end{bmatrix}$$

For the pth component of $\check{\boldsymbol{\alpha}}$, i.e., $\check{\alpha}_p$, we define two auxiliary variables $\xi_p = \max\{\check{\alpha}_p, 0\}$ and $\zeta_p = \max\{-\check{\alpha}_p, 0\}$ such that

$$\check{\alpha}_p = \xi_p - \zeta_p \tag{40}$$

$$|\check{\alpha}_p| = \xi_p + \zeta_p \tag{41}$$

Therefore we can re-express $\check{\boldsymbol{\alpha}}$ and $\|\check{\boldsymbol{\alpha}}\|_1$ as

$$\check{\boldsymbol{\alpha}} = \boldsymbol{\xi} - \boldsymbol{\zeta} \tag{42}$$

$$\|\check{\boldsymbol{\alpha}}\|_1 = \boldsymbol{1}^T(\boldsymbol{\xi} + \boldsymbol{\zeta}) \tag{43}$$

where $\boldsymbol{\xi} \in \mathbb{R}_+^P$ and $\boldsymbol{\zeta} \in \mathbb{R}_+^P$ are vectors with ξ_p and ζ_p being their pth component respectively, and $\mathbb{R}_+$ denotes the set of non-negative real numbers. Such a representation leads to

$$\begin{aligned} \min_{\boldsymbol{\xi},\boldsymbol{\zeta},\check{\boldsymbol{e}}} \quad & \boldsymbol{1}^T(\boldsymbol{\xi} + \boldsymbol{\zeta}) \\ \text{s.t.} \quad & \|\check{\boldsymbol{z}} - \check{\boldsymbol{A}}(\boldsymbol{\xi} - \boldsymbol{\zeta}) - 2\lambda\check{\boldsymbol{e}}\|_2^2 \leq \epsilon \\ & \check{\boldsymbol{e}} \in \{0, \pm 1\}^{2(M-1)} \\ & \boldsymbol{\xi} \in \mathbb{R}_+^P,\ \boldsymbol{\zeta} \in \mathbb{R}_+^P \end{aligned} \tag{44}$$

Unfortunately, (44) is a mixed-integer quadratic optimization problem, which is in general intractable. To deal with this challenge, we replace the quadratic inequality constraint by a linear constraint, resulting in the following optimization problem

$$\begin{aligned} \arg\min_{\boldsymbol{\xi},\boldsymbol{\zeta},\check{\boldsymbol{e}}} \quad & \boldsymbol{1}^T(\boldsymbol{\xi} + \boldsymbol{\zeta}) \\ \text{s.t.} \quad & -\epsilon'\boldsymbol{1} \preceq \left(\check{\boldsymbol{z}} - \check{\boldsymbol{A}}(\boldsymbol{\xi} - \boldsymbol{\zeta}) - 2\lambda\check{\boldsymbol{e}}\right) \preceq \epsilon'\boldsymbol{1} \\ & \check{\boldsymbol{e}} \in \{0, \pm 1\}^{2(M-1)} \\ & \boldsymbol{\xi} \in \mathbb{R}_+^P,\ \boldsymbol{\zeta} \in \mathbb{R}_+^P \end{aligned} \tag{45}$$

where ϵ' is a user-defined parameter and $\preceq$ denotes the element-wise inequality. The problem (45) is a mixed-integer linear programming that can be solved via the branch-and-bound algorithm (e.g., the off-the-shelf tool *intlinprog* in Matlab). Once we obtain $\boldsymbol{\xi}$, $\boldsymbol{\zeta}$, and $\check{\boldsymbol{e}}$, we can construct $\check{\boldsymbol{\alpha}} = \boldsymbol{\xi} - \boldsymbol{\zeta}$. It should be noted that other advanced solutions for the MILP problem, e.g., [50–52], can also be applied to solve our problem in (45). The mixed-integer linear programming (MILP) based algorithm for the LSE problem with modulo samples is summarized in Algorithm 2.

6.2. Theoretic analysis

In this subsection, we examine the condition under which the integer $\tilde{e}_m = (\tilde{x}_m - \tilde{f}_m)/2\lambda$ is confined to be in the set $\{0, \pm 1\}$. We have the following theorem with respect to the sinusoidal mixture signal defined in (27).

Algorithm 2 MILP-based algorithm for LSE problem with modulo samples

Input: $\boldsymbol{z}$, M, and the error tolerance ϵ.
Output: $\boldsymbol{\alpha}$ and the reconstructed original signal $\hat{\boldsymbol{y}}$.
1) Construct $\boldsymbol{A}$ and $\boldsymbol{D}_M^1$;
2) Compute $\tilde{\boldsymbol{A}} = \boldsymbol{D}_M^1 \boldsymbol{A}$ and $\tilde{\boldsymbol{z}} = \boldsymbol{D}_M^1 \boldsymbol{z}$;
3) Estimate $\boldsymbol{\xi}$ and $\boldsymbol{\zeta}$ by solving the optimization problem in (45) (e.g., via *intlinprog* in Matlab);
4) Calculate $\boldsymbol{\alpha}$ using $\boldsymbol{\xi}$ and $\boldsymbol{\zeta}$;
5) Reconstruct $\hat{\boldsymbol{y}} = \boldsymbol{A}\boldsymbol{\alpha}$.

Theorem 3. *Consider the ensemble of signals $f(t)$ that have a form of (27). If the sampling interval satisfies*

$$\Delta T < \frac{1}{\omega}\left(\frac{\bar{\beta}}{2\lambda}\right)^{-1} \tag{46}$$

then we have $\tilde{f}_m - \tilde{x}_m = 2\lambda\tilde{e}_m$, where $\tilde{e}_m$ is an integer which belongs to the set $\{0, \pm 1\}$.

Proof. See Appendix D. □

Theorem 3 indicates that, when the sampling rate is higher than the threshold in (46), we can guarantee that $\tilde{e}_m = (\tilde{x}_m - \tilde{f}_m)/2\lambda$ is confined to be in the set $\{0, \pm 1\}$. Similarly, we can rewrite the required sampling interval as

$$\Delta T < \frac{1}{\omega}\left(\frac{\bar{\beta}}{2\lambda}\right)^{-1} \triangleq \frac{1}{2\omega\eta_2} \tag{47}$$

where η_2 is given by

$$\eta_2 \triangleq \frac{\bar{N}}{2} \tag{48}$$

in which $\bar{N}$ is the effective folding number. It is clear that η_2 is smaller than η_1 since τ, defined before (34), should be greater than 0.5 in general. Such a relationship between η_1 and η_2 means that the sampling rate condition (47) is less restrictive than condition (34).

Remark 4. The sampling intervals for the higher-order difference based method and our proposed first-order difference based methods (the RCS-based and the MILP based approaches) are given by (12), (34), and (47), respectively. All these sampling intervals share a similar structure, differing only in their scaling factors. Specifically, the scaling factor for the higher-order difference-based method is a constant, whereas those for our methods depend on the effective number of folding times. Notably, the sampling interval requirements for our methods are not necessarily more restrictive than that of the higher-order difference-based method. For instance, consider the MILP-based method. When $\bar{N}/2 < e$ (i.e., $\bar{N} < 2e$), its required sampling interval becomes less restrictive than that of the higher-order difference-based solution.

7. Simulation results

In this section, we provide simulation results to illustrate the performance of the proposed first-order difference-based methods, namely, the RCS-based method and the MILP-based method. The HOD-based method discussed in Section 4 is also included as a benchmark to show the effectiveness of the proposed methods. For the RCS-based method, the sparse Bayesian learning method [53] is employed to solve the formulated robust compressed sensing problem.

7.1. Signal recovery performance

We consider $K = 3$ sinusoidal components with frequencies are $\{0.4\pi, 1.0\pi, 1.8\pi\}$, and the number of discretized grid points, P, is set to 20. For each realization, the complex amplitudes of these three sinusoidal components are set to ensure $B = 13.6$ and $\bar{\beta} = 4$. The sampling interval required by each method can be accordingly determined:

$$\Delta T_{\mathrm{HOD}} \le \frac{1}{2\omega e} = 0.0325 \tag{49}$$

$$\Delta T_{\mathrm{RCS}} \le \frac{1}{2\omega\eta_1} = 0.0884\lambda(1-\tau) \tag{50}$$

$$\Delta T_{\mathrm{MILP}} \le \frac{1}{2\omega\eta_2} = 0.0884\lambda \tag{51}$$

For the HOD based method, the order of difference has to satisfy the following condition

$$N \ge \left\lceil \frac{\log(0.0735\lambda)}{\log(15.3715\Delta T)} \right\rceil \tag{52}$$

where $\Delta T \le \Delta T_{\mathrm{HOD}}$ denotes the sampling interval for the HOD based method.

In our simulations, the initial phase of each sinusoidal component is randomly selected from $(0, 2\pi]$. The signal-to-noise ratio (SNR) is defined as $\mathrm{SNR} = \mathbb{E}(\|\boldsymbol{z}\|^2)/\mathbb{E}(\|\boldsymbol{v}\|^2)$. The range parameter is set to $\lambda = 1$ throughout our simulations. The normalized mean square error (NMSE) of $\boldsymbol{\alpha}$ and the reconstruction SNR (R-SNR) are used as two metrics to evaluate the estimation accuracy of the complex amplitudes and the frequencies. These two metrics are defined respectively as

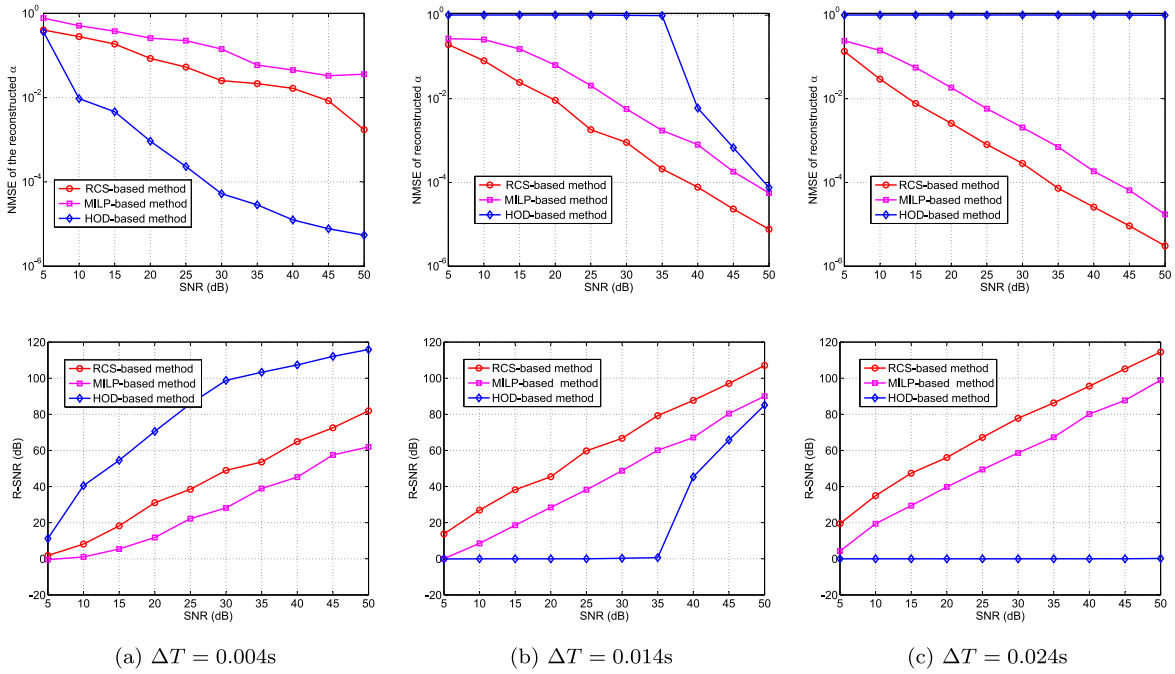
$$\mathrm{NMSE} \triangleq \frac{\mathbb{E}(\|\boldsymbol{\alpha} - \hat{\boldsymbol{\alpha}}\|^2)}{\mathbb{E}(\|\boldsymbol{\alpha}\|^2)} \tag{53}$$

$$\text{R-SNR} \triangleq 20\log_{10}\left(\frac{\|\boldsymbol{y}\|^2}{\|\boldsymbol{y} - \hat{\boldsymbol{y}}\|^2}\right) \tag{54}$$

where $\{\boldsymbol{\alpha}, \boldsymbol{y}\}$ and $\{\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{y}}\}$ denote the true parameters and the estimated ones, respectively.

The performance of respective methods along with the SNR is plotted in Fig. 3, where three different sampling intervals are employed. We see that the HOD-based method achieves the best performance when $\Delta T = 0.004$ s. This is because for the HOD-based method, the required order of difference is $N = 1$ when $\Delta T = 0.004$ s (cf. (52)), in which case the HOD-based method reduces to a first-order difference method. Nevertheless, when the sampling interval increases to $\Delta T = 0.014$ s and $\Delta T = 0.024$ s, the required orders of difference for the HOD-based method are 2 and 3, respectively, in which case the HOD-based method incurs a significant performance degradation. Specifically, when $\Delta T = 0.014$ s, the HOD-based method only works for extremely large SNRs (i.e., SNR $\ge$ 40 dB). In contrast, the proposed first-order difference-based methods achieve satisfactory performance across different sampling intervals, ensuring a reliable estimation even under low SNR conditions.

Fig. 4 depicts the NMSEs and the R-SNR of respective methods as the sampling interval ΔT varies from 0.004 s to 0.084 s with a stepsize of 0.008s, where SNR is set to 30 dB. It can be seen that the HOD-based method works well only when the sampling interval is as low as 0.004 s. This is because the required order of difference is $N = 1$ in this case. When the sampling interval increases to 0.02 s, although it still meets the condition (49), the order of difference required by the HOD-based method is greater than 1 according to (52). Hence the HOD-based method suffers severe performance degradation due to the shrinkage effect on the signal. In contrast, for the proposed first-order difference-based methods, we see that they are able to achieve superior recovery performance even when a smaller sampling rate is employed. In particular, the MILP-based method can provide reliable performance across different values of ΔT as long as the condition in (51) is met. The RCS-based method yields a better estimation accuracy than the MILP-based method when $\Delta T \le 0.052$ s. When ΔT further increases, the performance of the RCS-based method degrades. Specifically, when $\Delta T = 0.052$ s, it can be calculated that the probability of $\Pr(\tilde{f}_m = \tilde{x}_m)$ is no less than 41.2% according to (33). Due to relaxation used in our derivations, this probability is pessimistic. In fact, our experiments suggest that when $\Delta T = 0.052$ s, the probability of $\Pr(\tilde{f}_m = \tilde{x}_m)$ can



(a) $\Delta T = 0.004$s (b) $\Delta T = 0.014$s (c) $\Delta T = 0.024$s

Fig. 3. The NMSE and the R-SNR along with the SNR when $M = 400$.

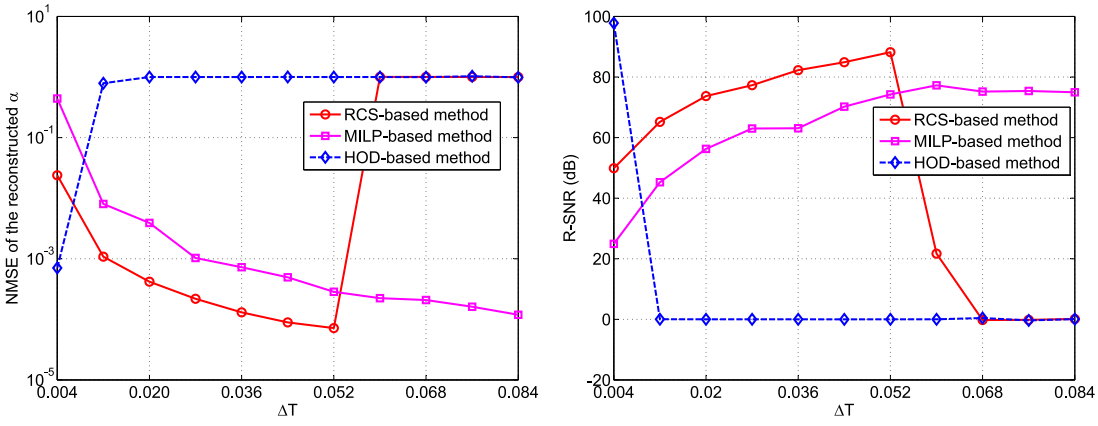


Fig. 4. The NMSE and the R-SNR along with ΔT when $M = 400$ and SNR $= 30$ dB.

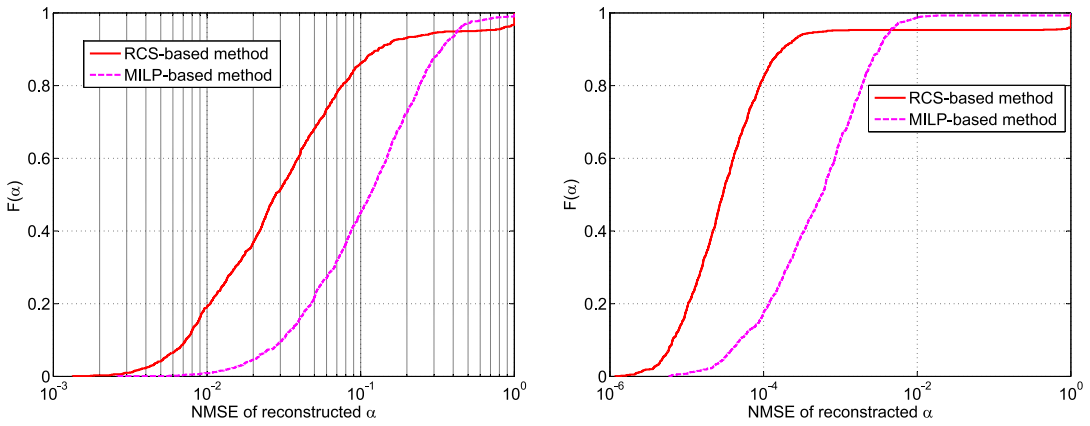
reach up to 90%, which implies a majority of elements in $\boldsymbol{s}$ are zeros. Hence the sparse outlier assumption required by the RCS-based method is well satisfied, and in this case the RCS-based method outperforms the MILP-based method. When ΔT increases to 0.060 s, our experiments suggest that the probability of $\Pr(\tilde{f}_m = \tilde{x}_m)$ drops below 80%. The sparse outlier assumption is barely met in this case. Therefore, when $\Delta T \geq 0.060$ s, the RCS-based method experiences a degraded performance compared to the MILP-based method.

Another interesting observation of Fig. 4 is that, within the region $0.004 \text{ s} \leq \Delta T \leq 0.052$ s, the performance of the RCS-based and MILP-based methods improves as the sampling interval increases. The is because, when the number of samples is fixed, a larger sampling interval usually results in a less coherent sensing matrix with a more favorable restricted isometry condition, which in turn leads to a better recovery performance. This can be verified by checking the mutual coherence of the sensing matrix $\tilde{\boldsymbol{A}}$, which tends to decrease as the sampling interval increases.

In addition, we consider a scenario where the number of sinusoidal components increases to $K = 10$. For each independent Monte Carlo run, the frequencies of these sinusoidal components are selected from the pre-defined grids, while their amplitudes are randomly selected from a Gaussian distribution $\mathcal{N}(0, 1)$. The sampling interval is set to $\Delta T = 0.027$ s and the dynamic range of the modulo ADC is set to $\lambda = 0.5$. We consider two cases, i.e., SNR $= 20$ dB and SNR $= 50$ dB. Results are averaged over 1000 independent realizations. Since the sampling interval is fixed while both the frequencies and the amplitudes are randomly selected, the sampling conditions required for the RCS-based method or/and the MILP-based method may not be satisfied. In Fig. 5, we illustrate the empirical PDF of the NMSE of these two cases, and we can see that the MILP-based method has a higher chance to succeed compared with the RCS-based method. This is because the MILP-based method has a more relaxed sampling condition compared with the RCS-based method, as shown in (34) and (47). This result is consistent with the previous conclusion as seen in Figs. 3 and 4.

7.2. Results with high dynamic ranges

The unlimited sensing framework is a potential solution to deal with a difficulty in wireless communications caused by the near-far



(a) SNR = 20dB (b) SNR = 50dB

Fig. 5. The empirical PDF of the NMSE with $M = 400$ when SNR = 20 dB and SNR = 50 dB.

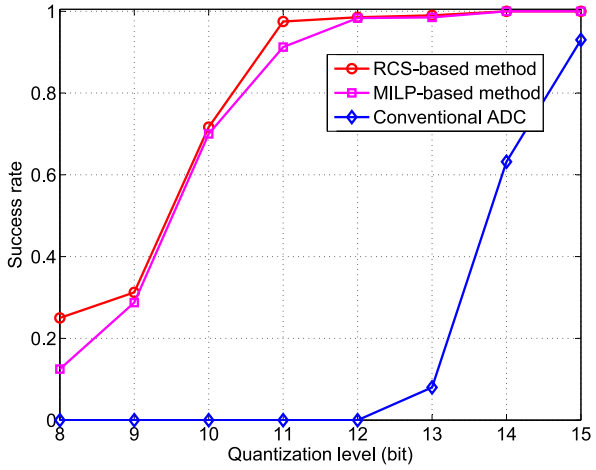


Fig. 6. The success rate with varied quantization levels when $M = 400$.

effect, where due to the limited dynamic range, a weak signal cannot be detected in the presence of a strong background signal. In this experiment, we illustrate the performance of the proposed methods in scenarios with a high dynamic range. To this end, we consider $K = 2$ sinusoidal components, and the corresponding parameters are given as

$$|\alpha_1| = 1000,\ |\alpha_2| = 1,\ \omega_1 = 0.2\pi,\ \omega_2 = 1.8\pi$$

The ratio of the power of the strong signal to the power of the weak signal is 60dB. To perform high-fidelity data-acquisition without saturation, a conventional ADC would need a large number of quantization bits. Otherwise the weak signal will be buried beneath the quantization noise level and cannot be effectively detected. Modulo ADCs do not suffer from this issue. Theoretically, modulo ADCs can achieve an unlimited dynamic range by folding the signal back to its range interval. Here we also consider the quantization noise for modulo ADCs. Unlike conventional ADCs, modulo ADCs quantize its modulo samples instead of the original input samples. In our simulations, we set $\lambda = 10$ and $\Delta T = 0.01$s.

To evaluate the performance of detecting the weak signal amidst the strong background signal, we use the success rate as a metric: a trial is considered successful if both of the following conditions are met:

1. The support set of the sparse signal $\boldsymbol{\alpha}$ is correctly found. In general, most components of the estimated $\boldsymbol{\alpha}$ will not exactly equal 0 due to the quantization noise. We set those components to zero if their amplitude is smaller than 0.1.
2. The estimation error of each non-zero element, defined as $|\hat{\boldsymbol{\alpha}}(i) - \boldsymbol{\alpha}(i)|/|\boldsymbol{\alpha}(i)|$, is no greater than 0.15.

Fig. 6 plots the success rates of respective methods as a function of the number of quantization bits. We see that the success rates of both proposed methods increase when more quantization bits are used. Specifically, a success rate of 1 can be achieved when 11 bits are employed to quantize the modulo samples. In contrast, the conventional ADC needs 15 quantization bits to achieve a decent success rate. In Fig. 7, we plot the estimated $\boldsymbol{\alpha}$ for a particular realization when $M = 600$ and the number of quantization bits is set to 10. It can be seen that while the proposed modulo ADC based methods have no difficulty in recovering both the strong and weak signals, the conventional ADC cannot distinguish the weak signal from many false signals it produces.

7.3. Hardware test

To further illustrate the performance of the proposed method, we conduct hardware test, including the development of a simple demo system for modulo sampling. The schematic diagram of the demo system is shown in Fig. 8. This architecture supports input signals with amplitudes up to at least triple the reference voltage, enabling reliable and scalable signal folding for high dynamic range applications. Specifically, the input signal is first compared against predefined thresholds, denoted as λ and $-\lambda$. If the signal exceeds λ, the comparator outputs a high voltage; if it falls below λ, a low voltage output is generated. Bipolar junction transistors are employed as switches to manage the reference voltage in response to the comparator's output. A combination of a differential amplifier and an adder is used in the subsequent stage to modulate the input signal, ensuring it remains within the defined range of $[-\lambda, \lambda]$. When the input signal exceeds the thresholds multiple times, the design accommodates this by integrating additional modulation stages in series. This cascading configuration allows for repeated folding of the signal, preserving its integrity and ensuring accurate processing within the modulo ADC framework. To accurately detect whether the input signal is within the threshold range, the ADA4522 comparator is employed due to its ultra-low offset voltage, minimal drift, and excellent temperature stability. Its rail-to-rail output, low noise characteristics, and integrated EMI filters make it ideal for precision signal processing in demanding environments, such as data acquisition systems and industrial instrumentation.

We consider two sinusoidal components as the input signal. Specifically, the frequencies are 95 Hz and 100 Hz, while the associated amplitudes are 3.0 and 1.5. The sampling rate of the demo system is

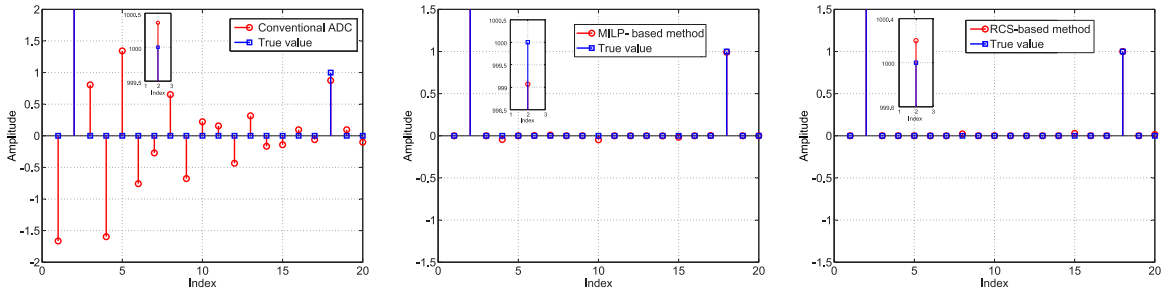


Fig. 7. The estimated α of different methods in one specific run when $M = 600$ and the quantization level is 10bit.

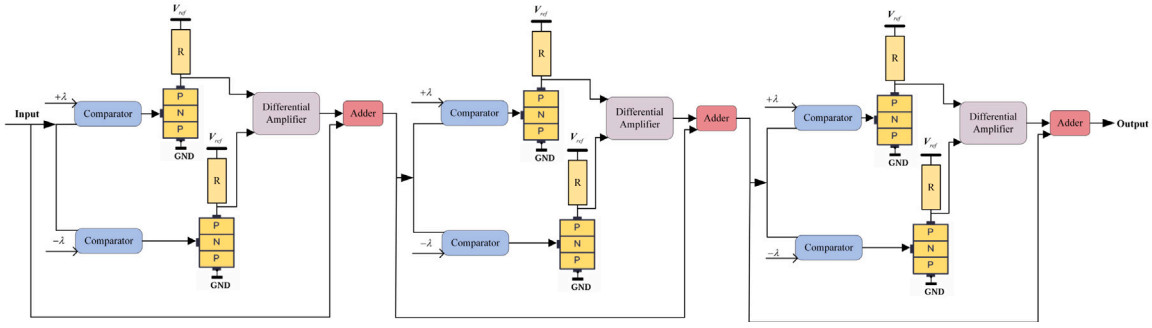


Fig. 8. The schematic diagram of the demo system for modulo sampling.

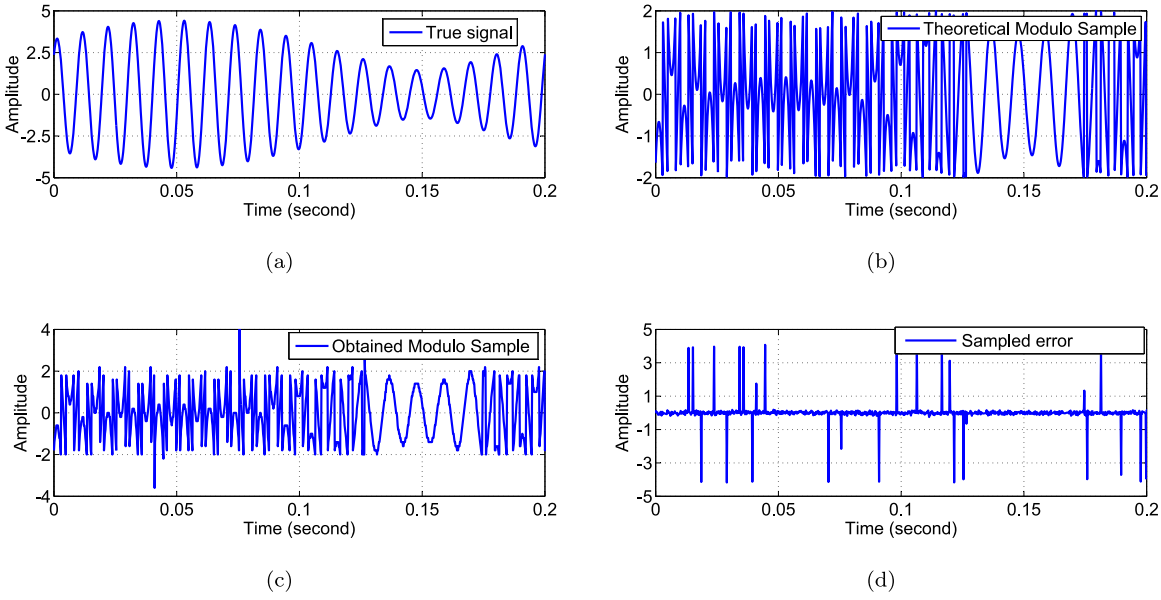


Fig. 9. The estimated α of different methods in one specific run when $M = 600$ and the quantization level is 10bit.

set to 5000 Hz and we take 2500 samples in total. Partial data (1000 samples) are illustrated in Fig. 9, and it can be observed that the sampled error contains two parts. The first one is outliers, which may be caused by the hardware limitations such as the hysteresis; the other part can be considered as the noise, which may caused by quantization noises.

Only partial samples are randomly selected from the original dataset as the tested samples. In addition, to evaluate the proposed method's performance across varying sampling intervals, we adjust the intervals by directly extracting samples from the original dataset. We compared our RCS-based approach with the HOD-based method but excluded the MILP-based method's results. This exclusion is because hardware

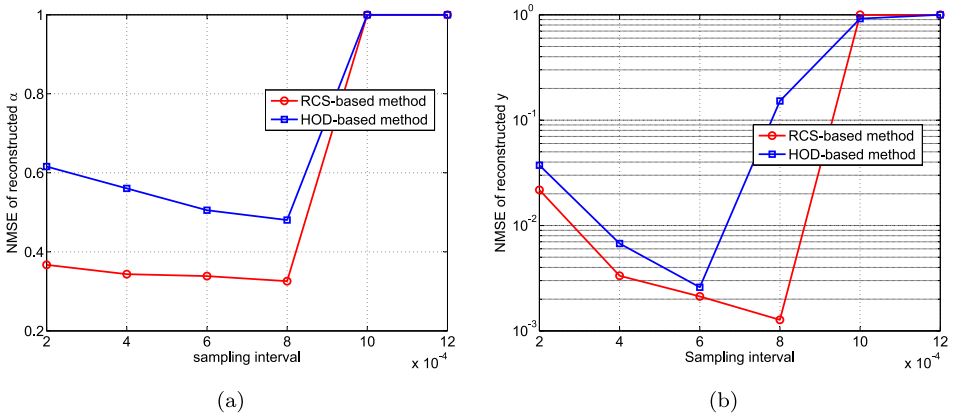


Fig. 10. The NMSEs of the constructed α and y along with ΔT when $M = 300$.

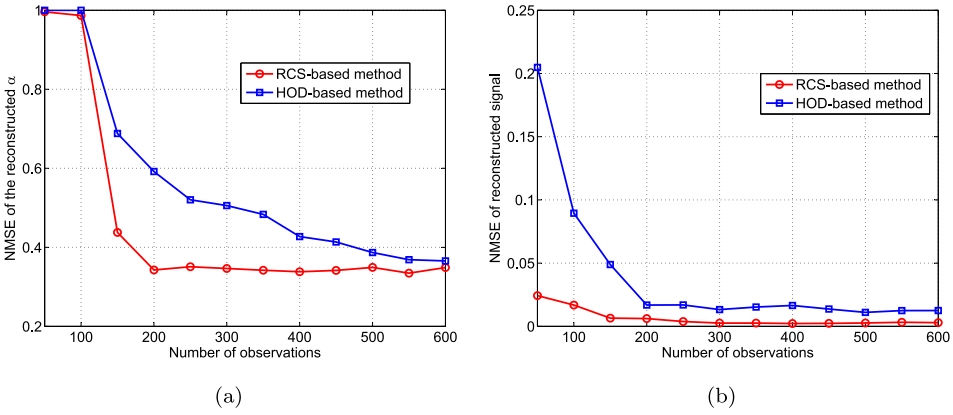


Fig. 11. The NMSEs of the constructed α and y along with M when $\Delta T = 4 \times 10^{-4}$.

non-ideality introduces impulsive noise, which prevents the residue between the first-order modulo and original samples from being confined to the set $\{0, \pm 1\}$.

The NMSEs of the reconstructed $\boldsymbol{\alpha}$ and y relative to sampling intervals and the number of modulo samples are, respectively, displayed in Figs. 10 and 11. The performance of both methods improves with the increase of ΔT until ΔT exceeds 10^{-3}, which is consistent with the trends observed in our numerical results shown in Fig. 4. In addition, as the number of observations increase, the NMSEs of these two algorithms decline and stabilize at a certain level. Overall, the RCS-based method consistently outperforms the HOD-based method in robustness and accuracy across both scenarios.

8. Conclusions

In this work, we studied the problem of LSE in an unlimited sensing framework, which employs modulo ADCs to nonlinearly map the received signal to modulo samples to avoid clipping or saturation problems. We first introduced a HOD based approach, which is sensitive to noise and lacks satisfactory performance in the presence of noise. To overcome this difficulty, we investigated the properties of the first-order difference of modulo samples, and developed two first-order difference-based methods, namely, the RCS-based and the MILP-based methods. Simulation results show that both methods are robust against noise and achieve a significant performance improvement over the HOD-based method. In practice, due to implementation imperfections of modulo ADCs such as hysteresis effects [29] and impulsive noises [54], the folding numbers might not be integers. How to extend our methods to deal with these nonlinearities is an important topic for future investigation. In addition, LSE with off-grid frequencies is commonly encountered in real applications. Devising an off-grid LSE methods to modulo samples in the robust compressive sensing is another key direction for future research.

CRediT authorship contribution statement

Hongwei Wang: Writing – original draft, Validation, Software, Methodology, Conceptualization. **Jun Fang:** Supervision, Conceptualization. **Hongbin Li:** Writing – review & editing, Validation, Conceptualization. **Geert Leus:** Writing – review & editing, Validation. **Ruixiang Zhu:** Data curation. **Lu Gan:** Writing – review & editing, Data curation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Discussion on (31)

According to [55], for an input signal x to be quantized, if the characteristic function (CF) of the random variable x, denoted as $\boldsymbol{\Phi}_x(u)$, is "bandlimited", i.e., it satisfies

$$\boldsymbol{\Phi}_x(u) = 0 \quad |u| \geq \frac{2\pi}{2\lambda} \tag{A.1}$$

where 2λ is the quantization step size, then the quantization noise follows a uniform distribution. This condition is restrictive, as most frequently encountered signals such as Gaussian signals do not have perfectly bandlimited CFs. Therefore, an approximate condition is widely

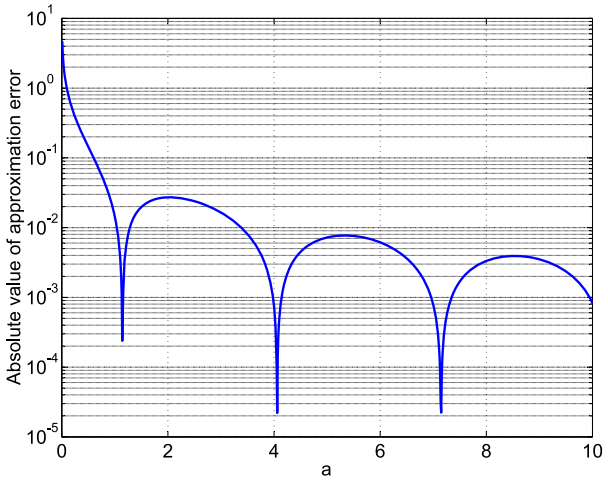


Fig. A.12. The absolute value of the approximation error of (A.7) as a function of a.

accepted. Specifically, if $\boldsymbol{\Phi}_x(u)$ is *approximately bandlimited* for some quantization step size, i.e.,

$$\boldsymbol{\Phi}_x(u) \approx 0 \qquad |u| \geq \frac{2\pi}{2\lambda} \tag{A.2}$$

then the quantization noise can be approximately modeled as a uniform distribution.

In our problem, we assume that the phase of the kth component, i.e., $f_k \triangleq \beta_k \cos(\omega_k t + \varphi_k)$, is uniform in the interval $(-\pi, \pi)$. Therefore, the probability density function of f_k is given by [44–46]

$$p_k(f_k) = \begin{cases} \frac{1}{\pi\sqrt{\beta_k^2 - f_k^2}}, & |f_k| \leq |\beta_k| \\ 0, & |f_k| > |\beta_k| \end{cases} \tag{A.3}$$

Accordingly, the CF of the random variable f_k is given by [44]

$$\boldsymbol{\Phi}_k(u) = \mathcal{J}_0\big(|\beta_k| u\big) \tag{A.4}$$

where $\mathcal{J}_0(\cdot)$ denotes the Bessel function of the first kind. Note that the CF of the sum of K independent random variables is the product of individual CFs associated with these random variables. Therefore, the CF of $f(t)$ can be calculated as

$$\boldsymbol{\Phi}_f(u) = \prod_{k=1}^{K} \boldsymbol{\Phi}_k(u) = \prod_{k=1}^{K} \mathcal{J}_0\big(|\beta_k| u\big) \tag{A.5}$$

Note that the Bessel function $|\mathcal{J}_0(a)| < 1$ when $a \neq 0$. Hence we have

$$\Big|\boldsymbol{\Phi}_f(u)\Big| < \Big|\mathcal{J}_0\big(|\beta_k| u\big)\Big|, \ \forall k \in \{1, \ldots, K\} \tag{A.6}$$

where $u \neq 0$. In addition, the Bessel function $\mathcal{J}_0(a)$ is an even function and looks like a decaying sinusoid that decays proportionally to $1/\sqrt{|a|}$. Specifically, for $|a| \gg 1/4$, $\mathcal{J}_0(a)$ can be approximated as [56]

$$\mathcal{J}_0(a) \approx \sqrt{\frac{2}{\pi|a|}} \cos(|a| - \frac{\pi}{4}) \tag{A.7}$$

The absolute value of the approximation error in (A.7) is shown in Fig. A.12, where we can see that, when $a \geq 2$, the absolute error is less than 0.03. In the following, we suppose that (A.7) holds exactly (i.e., the approximation error is negligible) when $|a| \geq 2$.

Setting $\tilde{\beta} \triangleq \max\{|\beta_k|\}$, $\mathcal{J}_0(\tilde{\beta}u)$ has a faster decaying rate than $\mathcal{J}_0(|\beta_k u|)$. When $|\tilde{\beta}u| \geq 2$, we have

$$\begin{aligned} \Big|\boldsymbol{\Phi}_f(u)\Big| &\overset{(a)}{<} \Big|\mathcal{J}_0(\tilde{\beta}u)\Big| \\ &\overset{(b)}{\approx} \left|\sqrt{\frac{2}{\pi|\tilde{\beta}u|}} \cos(|\tilde{\beta}u| - \frac{\pi}{4})\right| \leq \sqrt{\frac{2}{\pi\tilde{\beta}|u|}} \end{aligned} \tag{A.8}$$

where (a) is directly from (A.6), and in (b) we employ the approximation in (A.7). From (A.8) we know that if

$$|u| \geq \max\left\{\frac{2}{\pi\tilde{\beta}\varsigma^2}, \frac{2}{\tilde{\beta}}\right\} \tag{A.9}$$

then the following inequality holds

$$\Big|\boldsymbol{\Phi}_f(u)\Big| \leq \varsigma \tag{A.10}$$

where ς is a small positive value. Generally, ς is set sufficiently small such that $\pi\varsigma^2 \leq 1$ (specifically, ς is required to be no greater than 0.564). Therefore, the condition in (A.9) is equivalent to

$$|u| \geq \frac{2}{\pi\tilde{\beta}\varsigma^2} \tag{A.11}$$

Based on the above result, to ensure that the approximately bandlimited condition (A.2) holds, the following condition should be satisfied

$$|u| \geq \frac{2\pi}{2\lambda} \geq \frac{2}{\pi\tilde{\beta}\varsigma^2} \tag{A.12}$$

The above condition implies

$$2\lambda \leq (\pi\varsigma)^2\tilde{\beta} \tag{A.13}$$

Considering the fact that $B \triangleq \sum_{k=1}^{K} |\beta_k|$ and $\tilde{\beta} \triangleq \max\{|\beta_k|\}$, we have

$$2\lambda \leq (\pi\varsigma)^2 \frac{B}{K} \tag{A.14}$$

to ensure (A.12).

From (A.14), we know that the operation range of the modulo ADC, i.e., 2λ, should be smaller than $(\pi\varsigma)^2 B/K$, such that x_m can be approximately modeled as a uniform distribution. This is equivalent to the condition that the maximum folding number, defined by $\mathcal{K} \triangleq B/(2\lambda)$, should satisfy the following inequality:

$$\mathcal{K} \geq \frac{K}{(\pi\varsigma)^2} \tag{A.15}$$

If we set $K = 4$ and $\varsigma = 0.1$, $\mathcal{K}$ should be no less than 40. This condition is usually met considering the fact that the unlimited sensing framework aims to deal with high dynamic range problems.

Appendix B. Proof of Theorem 2

By the modulo decomposition property [15], we have

$$x_{m+1} = f_{m+1} + 2\lambda e_{m+1}, \ x_m = f_m + 2\lambda e_m$$

where $e_{m+1}, e_m \in \mathbb{Z}$. Therefore, we have

$$\tilde{x}_m = \tilde{f}_m - 2\lambda(e_m - e_{m+1}) \triangleq \tilde{f}_m - 2\lambda\tilde{e}_m \tag{B.1}$$

where $\tilde{e}_m \triangleq (e_m - e_{m+1}) \in \mathbb{Z}$. Therefore, $\tilde{x}_m$ can be uniquely represented by a sum of $\tilde{f}_m$ and an unknown constant. In scenarios where this constant equals 0, $\tilde{x}_m$ is equivalent to $\tilde{f}_m$. Therefore, we have

$$\Pr(\tilde{f}_m = \tilde{x}_m) \Leftrightarrow \Pr(\tilde{e}_m = 0) \tag{B.2}$$

By the definition of the modulo operation in (1), we have

$$x_{m+1} = f_{m+1} - 2\lambda\left\lfloor\frac{f_{m+1}}{2\lambda} + \frac{1}{2}\right\rfloor \tag{B.3}$$

which indicates that

$$e_{m+1} = -\left\lfloor\frac{f_{m+1}}{2\lambda} + \frac{1}{2}\right\rfloor \tag{B.4}$$

Similarly, e_m can be expressed as

$$e_m = -\left\lfloor\frac{f_m}{2\lambda} + \frac{1}{2}\right\rfloor \tag{B.5}$$

Combining (B.4) and (B.5) results in

$$\begin{aligned} \tilde{e}_m &\triangleq e_m - e_{m+1} \\ &= \left(\left\lfloor\frac{f_{m+1}}{2\lambda} + \frac{1}{2}\right\rfloor - \left\lfloor\frac{f_m}{2\lambda} + \frac{1}{2}\right\rfloor\right) \end{aligned}$$

$$= \left(\left\lfloor \frac{f_m}{2\lambda} + \frac{1}{2} + \frac{f_{m+1} - f_m}{2\lambda} \right\rfloor - \left\lfloor \frac{f_m}{2\lambda} + \frac{1}{2} \right\rfloor \right) \tag{B.6}$$

Since $f(t)$ is continuous and differentiable on the interval $[(t-1)\Delta T, t\Delta T]$, based on the mean value theorem, there exists a value $\zeta \in [(t-1)\Delta T, t\Delta T]$ such that

$$f'(\zeta) = \frac{f_{m+1} - f_m}{\Delta T} \tag{B.7}$$

where $f'(t)$ is the derivative of $f(t)$. Substituting (B.7) into (B.6), we arrive at

$$\tilde{e}_m = \lfloor b \rfloor - \lfloor a \rfloor \tag{B.8}$$

where a and b are, respectively, defined as

$$a \triangleq \frac{f_m}{2\lambda} + \frac{1}{2},\ b \triangleq \frac{f_m}{2\lambda} + \frac{1}{2} + \frac{f'(\zeta)\Delta T}{2\lambda}$$

By these definitions we have

$$b = a + \frac{f'(\zeta)\Delta T}{2\lambda} \triangleq a + c \tag{B.9}$$

where c is defined as

$$c \triangleq \frac{f'(\zeta)\Delta T}{2\lambda} \tag{B.10}$$

The derivative of $f(t)$ is given by

$$f'(t) = -\sum_{k=1}^{K} \beta_k \omega_k \sin(\omega_k t + \varphi_k) \tag{B.11}$$

which leads to the following bound

$$|c| \le \frac{\Delta T \sum_{k=1}^{K} |\beta_k \omega_k|}{2\lambda} = \frac{\Delta T \omega \bar{\beta}}{2\lambda} \triangleq \delta \tag{B.12}$$

As in (29), f_m can be expressed as

$$f_m = 2\lambda \left\lfloor \frac{f_m}{2\lambda} + \frac{1}{2} \right\rfloor + x_m = 2\lambda \lfloor a \rfloor + x_m \tag{B.13}$$

where x_m can be modeled as a uniform distribution:

$$x_m \sim \mathbb{U}(-\lambda, \lambda) \tag{B.14}$$

In addition, from the quantization theory in [55], x_m is statically independent of f_m. Taking a simple algebraic operation on (B.13) results in

$$\frac{f_m}{2\lambda} + \frac{1}{2} = \lfloor a \rfloor + \tilde{q}_m \tag{B.15}$$

where $\tilde{q}_m \triangleq x_m/(2\lambda) + 1/2$ also follows a uniform distribution, i.e., $\tilde{q}_m \sim \mathbb{U}(0,1)$. Recalling the definition of a, we have

$$a = \lfloor a \rfloor + \tilde{q}_m \tag{B.16}$$

Therefore, a can be assumed to follow a uniform distribution over the interval $(\lfloor a \rfloor, \lfloor a \rfloor + 1)$:

$$p_A(a) \sim \mathbb{U}\Big(\lfloor a \rfloor, \lfloor a \rfloor + 1 \Big) \tag{B.17}$$

For c, it can be modeled as a random variable and its distribution $p_C(c)$ is given as (as we have $|c| \le \delta$)

$$p_C(c) = \begin{cases} 0, & c < -\delta \text{ or } c > \delta \\ p_C(c), & -\delta \le c \le \delta \end{cases} \tag{B.18}$$

where $p_C(c)$ is an arbitrary distribution that satisfies

$$\int_{-\delta}^{\delta} p_C(c)dc = 1 \tag{B.19}$$

Furthermore, it is reasonable to assume that a and c are independent of each other. This is because the quantization noise, $\tilde{q}_m$, is independent of $f(t)$, and c is only related to the derivative of $f(t)$. In this case, the probability density function (PDF) of b is given by the convolution of $p_A(a)$ and $p_C(c)$, i.e.,

$$p_B(b) = \int_{-\infty}^{\infty} p_C(b-a) p_A(a) da \tag{B.20}$$

Recall that the sampling interval ΔT satisfies

$$\Delta T \le \frac{1}{2\omega} \left(\frac{\bar{\beta}}{2\lambda} \right)^{-1} \tag{B.21}$$

Hence we have

$$\delta = \frac{\Delta T \omega \bar{\beta}}{2\lambda} \le \frac{1}{2} \tag{B.22}$$

and consequently, $\lfloor a \rfloor + 1 - \delta \ge \lfloor a \rfloor + \delta$. Therefore, $p_B(b)$ can be calculated as

$$p_B(b) = \begin{cases} 0, & b < \lfloor a \rfloor - \delta \\ F_C(b - \lfloor a \rfloor), & \lfloor a \rfloor - \delta \le b < \lfloor a \rfloor + \delta \\ 1, & \lfloor a \rfloor + \delta \le b < \lfloor a \rfloor + 1 - \delta \\ 1 - F_C(b - \lfloor a \rfloor - 1), & \lfloor a \rfloor + 1 - \delta \le b < \lfloor a \rfloor + 1 + \delta \\ 0, & b > \lfloor a \rfloor + 1 + \delta \end{cases} \tag{B.23}$$

where $F_C(\cdot)$ is the cumulative distribution function (CDF) of c. The derivation of $p_B(b)$ can be found in Appendix C. The desired probability is then computed as

$$\begin{aligned} &\Pr(\tilde{e}_m = 0) \Leftrightarrow \Pr(\lfloor b \rfloor = \lfloor a \rfloor) \\ &= \Pr(\lfloor a \rfloor \le b < \lfloor a \rfloor + 1) \\ &= \int_{\lfloor a \rfloor}^{\lfloor a \rfloor + 1} p_B(b) db \\ &\triangleq \kappa_1 + \kappa_2 + \kappa_3 \end{aligned} \tag{B.24}$$

where κ_1, κ_2 and κ_3 are respectively calculated by

$$\kappa_1 = \int_{\lfloor a \rfloor}^{\lfloor a \rfloor + \delta} F_c(b - \lfloor a \rfloor) db = \int_0^{\delta} F_C(c) dc \tag{B.25}$$

$$\kappa_2 = \int_{\lfloor a \rfloor + \delta}^{\lfloor a \rfloor + 1 - \delta} (1) db = 1 - 2\delta \tag{B.26}$$

$$\kappa_3 = \int_{\lfloor a \rfloor + 1 - \delta}^{\lfloor a \rfloor + 1} \big(1 - F_C(b - \lfloor a \rfloor - 1)\big) db = \delta - \int_{-\delta}^{0} F_C(c) dc \tag{B.27}$$

Therefore, we have

$$\begin{aligned} \Pr(\tilde{e}_m = 0) &= \kappa_1 + \kappa_2 + \kappa_3 \\ &= 1 - \delta + \int_0^{\delta} F_C(c) dc - \int_{-\delta}^{0} F_C(c) dc \end{aligned} \tag{B.28}$$

Due to the property of CDF, $F_C(c)$ is non-negative and non-decreasing over the interval $[-\delta, \delta]$. Therefore, the following holds:

$$\int_0^{\delta} F_C(c) dc - \int_{-\delta}^{0} F_C(c) dc \ge 0 \tag{B.29}$$

Combining (B.28) and (B.29), we arrive at our main result

$$\Pr(\tilde{f}_m = \tilde{x}_m) \Leftrightarrow \Pr(\tilde{e}_m = 0) \ge 1 - \delta = 1 - \frac{\Delta T \omega \bar{\beta}}{2\lambda} \tag{B.30}$$

This completes our proof. □

Appendix C. Derivation of $p_B(b)$

The PDF of b can be expressed as

$$\begin{aligned} p_B(b) &= \int_{-\infty}^{\infty} p_C(b-a) p_A(a) da \\ &= \int_{\lfloor a \rfloor}^{\lfloor a \rfloor + 1} p_C(b-a) da \\ &\overset{(a)}{=} \int_{b - \lfloor a \rfloor - 1}^{b - \lfloor a \rfloor} p_C(c) dc \end{aligned} \tag{C.1}$$

where in (a) we utilize the definition $c \triangleq b - a$. Recall that $p_C(c)$ is nonzero only within the interval $[-\delta, \delta]$. Therefore we have $p_B(b) = 0$ if $b - \lfloor a \rfloor < -\delta$ or $b - \lfloor a \rfloor - 1 > \delta$ (i.e., $b < \lfloor a \rfloor - \delta$ or $b > \lfloor a \rfloor + 1 + \delta$).

On the other hand, when $-\delta \le b - \lfloor a \rfloor - 1 < \delta$, i.e., $-\delta + \lfloor a \rfloor + 1 \le b < \lfloor a \rfloor + 1 + \delta$, we have

$$p_B(b) = \int_{b - \lfloor a \rfloor - 1}^{\delta} p_C(c) dc = 1 - F_C(b - \lfloor a \rfloor - 1) \tag{C.2}$$

Moreover, when $b - \lfloor a \rfloor - 1 < -\delta$ and $b - \lfloor a \rfloor \geq \delta$, i.e., $\lfloor a \rfloor + \delta \leq b < \lfloor a \rfloor + 1 - \delta$, we have

$$p_B(b) = \int_{-\delta}^{\delta} p_C(c)dc = 1 \tag{C.3}$$

Also, when $-\delta \leq b - \lfloor a \rfloor < \delta$, i.e., $\lfloor a \rfloor - \delta \leq b < \lfloor a \rfloor + \delta$, we have

$$p_B(b) = \int_{-\delta}^{b-\lfloor a \rfloor} p_C(c)dc = F_C(b - \lfloor a \rfloor) \tag{C.4}$$

Summarizing the above results, we arrive at (B.20).

Appendix D. Proof of Theorem 3

From (B.8), we know that $\tilde{e}_m$ can be expressed by

$$\tilde{e}_m = \lfloor b \rfloor - \lfloor a \rfloor \tag{D.1}$$

where $b = a + c$. Based on (B.12), we know that

$$|c| \leq \delta = \frac{\Delta T \omega \bar{\beta}}{2\lambda} \tag{D.2}$$

Recalling (46), we have

$$\|c\|_\infty < \frac{\omega\bar{\beta}}{2\lambda}\frac{1}{\omega}\left(\frac{\bar{\beta}}{2\lambda}\right)^{-1} = 1 \tag{D.3}$$

which indicates that c belongs to the set

$$-1 < c < 1 \tag{D.4}$$

On the other hand, we know that

$$\lfloor a \rfloor \leq a < \lfloor a \rfloor + 1 \tag{D.5}$$

Combining (D.4) and (D.5) we obtain that

$$\lfloor a \rfloor - 1 < b < \lfloor a \rfloor + 2 \tag{D.6}$$

Therefore $\lfloor b \rfloor$ belongs to the set $\{\lfloor a \rfloor, \lfloor a \rfloor \pm 1\}$, which implies that $\tilde{e}_m \in \{0, \pm 1\}$. This completes the proof. □

Data availability

Data will be made available on request.

References

[1] B. Liao, S.-C. Chan, L. Huang, C. Guo, Iterative methods for subspace and DOA estimation in nonuniform noise, IEEE Trans. Signal Process. 64 (12) (2016) 3008–3020.

[2] Y. Tan, K. Wang, L. Wang, H. Wen, Efficient FFT based multi source DOA estimation for ULA, Signal Process. 189 (2021) 108284.

[3] J. Yang, Y. Yang, J. Lu, L. Yang, Iterative methods for DOA estimation of correlated sources in spatially colored noise fields, Signal Process. 185 (2021) 108100.

[4] R. Carriere, R.L. Moses, High resolution radar target modeling using a modified Prony estimator, IEEE Trans. Antennas Propag. 40 (1) (1992) 13–18.

[5] Y. Li, S. Chang, Z. Liu, W. Ren, Q. Liu, Range ambiguity suppression under high-resolution estimation using the MUSIC-AP algorithm for pulse-doppler radar, Signal Process. 214 (2024) 109237.

[6] J.L. Flanagan, Speech Analysis Synthesis and Perception, vol. 3, Springer Science & Business Media, 2013.

[7] P. Stoica, R.L. Moses, et al., Spectral Analysis of Signals, vol. 452, Pearson Prentice Hall Upper Saddle River, NJ, 2005.

[8] J.F. Hauer, C. Demeure, L. Scharf, Initial results in Prony analysis of power system response signals, IEEE Trans. Power Syst. 5 (1) (1990) 80–89.

[9] P. Stoica, A. Eriksson, MUSIC estimation of real-valued sine-wave frequencies, Signal Process. 42 (2) (1995) 139–146.

[10] M.H. Gruber, Statistical digital signal processing and modeling, 1997.

[11] R. Roy, T. Kailath, ESPRIT-estimation of signal parameters via rotational invariance techniques, IEEE Trans. Acoust. Speech Signal Process. 37 (7) (1989) 984–995.

[12] J. Fang, F. Wang, Y. Shen, H. Li, R.S. Blum, Super-resolution compressed sensing for line spectral estimation: An iterative reweighted approach, IEEE Trans. Signal Process. 64 (18) (2016) 4649–4662.

[13] A. Eamaz, K.V. Mishra, F. Yeganegi, M. Soltanalian, Uno: Unlimited sampling meets one-bit quantization, IEEE Trans. Signal Process. 72 (2024) 997–1014.

[14] A. Bhandari, F. Krahmer, R. Raskar, On unlimited sampling, in: 2017 International Conference on Sampling Theory and Applications, SampTA, IEEE, 2017, pp. 31–35.

[15] A. Bhandari, F. Krahmer, R. Raskar, On unlimited sampling and reconstruction, IEEE Trans. Signal Process. 69 (2020) 3827–3839.

[16] Q. Zhang, J. Zhu, F. Qu, Z. Zhu, D.W. Soh, On the identifiability from modulo measurements under dft sensing matrix, 2023, arXiv preprint arXiv:2401.00194.

[17] Q. Zhang, J. Zhu, F. Qu, et al., Line spectral estimation via unlimited sampling, IEEE Trans. Aerosp. Electron. Syst. (2024).

[18] G. Shtendel, A. Bhandari, HDR-ToF: HDR time-of-flight imaging via modulo acquisition, in: 2022 IEEE International Conference on Image Processing, ICIP, IEEE, 2022, pp. 3808–3812.

[19] Z. Liu, A. Bhandari, B. Clerckx, λ–MIMO: Massive MIMO via modulo sampling, IEEE Trans. Commun. 71 (11) (2023) 6301–6315.

[20] T. Feuillen, M. Alaee-Kerahroodi, A. Bhandari, B. Ottersten, et al., Unlimited sampling for FMCW radars: A proof of concept, in: 2022 IEEE Radar Conference, RadarConf22, IEEE, 2022, pp. 1–5.

[21] A. Bhandari, F. Krahmer, R. Raskar, Unlimited sampling of sparse signals, in: 2018 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP, IEEE, 2018, pp. 4569–4573.

[22] A. Bhandari, F. Krahmer, R. Raskar, Unlimited sampling of sparse sinusoidal mixtures, in: 2018 IEEE International Symposium on Information Theory, ISIT, IEEE, 2018, pp. 336–340.

[23] S. Fernández-Menduina, F. Krahmer, G. Leus, A. Bhandari, DOA estimation via unlimited sensing, in: 2020 28th European Signal Processing Conference, EUSIPCO, IEEE, 2021, pp. 1866–1870.

[24] S. Fernández-Menduiña, F. Krahmer, G. Leus, A. Bhandari, Computational array signal processing via modulo non-linearities, IEEE Trans. Signal Process. 70 (2022) 2168–2179.

[25] O. Graf, A. Bhandari, F. Krahmer, One-bit unlimited sampling, in: ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP, IEEE, 2019, pp. 5102–5106.

[26] A. Eamaz, F. Yeganegi, M. Soltanalian, Streamlining uno: A generalized sampling approach to optimal one-bit modulo sensing, in: ICASSP 2025-2025 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP, IEEE, 2025, pp. 1–5.

[27] A. Bhandari, F. Krahmer, T. Poskitt, Unlimited sampling from theory to practice: Fourier-prony recovery and prototype ADC, IEEE Trans. Signal Process. 70 (2021) 1131–1141.

[28] D. Florescu, F. Krahmer, A. Bhandari, Unlimited sampling with hysteresis, in: 2021 55th Asilomar Conference on Signals, Systems, and Computers, IEEE, 2021, pp. 831–835.

[29] D. Florescu, F. Krahmer, A. Bhandari, The surprising benefits of hysteresis in unlimited sampling: Theory, algorithms and experiments, IEEE Trans. Signal Process. 70 (2022) 616–630.

[30] D. Florescu, F. Krahmer, A. Bhandari, Event-driven modulo sampling, in: ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP, IEEE, 2021, pp. 5435–5439.

[31] A. Bhandari, Back in the US-SR: Unlimited sampling and sparse super-resolution with its hardware validation, IEEE Signal Process. Lett. 29 (2022) 1047–1051.

[32] G. Tang, P. Shah, B.N. Bhaskar, B. Recht, Robust line spectral estimation, in: 2014 48th Asilomar Conference on Signals, Systems and Computers, IEEE, 2014, pp. 301–305.

[33] J. Dai, H.C. So, Sparse Bayesian learning approach for outlier-resistant direction-of-arrival estimation, IEEE Trans. Signal Process. 66 (3) (2017) 744–756.

[34] J.A. Tropp, A.C. Gilbert, Signal recovery from random measurements via orthogonal matching pursuit, IEEE Trans. Inform. Theory 53 (12) (2007) 4655–4666.

[35] J. Wright, A.Y. Yang, A. Ganesh, S.S. Sastry, Y. Ma, Robust face recognition via sparse representation, IEEE Trans. Pattern Anal. Mach. Intell. 31 (2) (2008) 210–227.

[36] N.H. Nguyen, T.D. Tran, Robust lasso with missing and grossly corrupted observations, IEEE Trans. Inform. Theory 59 (4) (2012) 2036–2058.

[37] N.H. Nguyen, T.D. Tran, Exact recoverability from dense corrupted observations via ℓ_1-minimization, IEEE Trans. Inform. Theory 59 (4) (2013) 2017–2035.

[38] R.E. Carrillo, K.E. Barner, Lorentzian iterative hard thresholding: Robust compressed sensing with prior information, IEEE Trans. Signal Process. 61 (19) (2013) 4822–4833.

[39] E. Ollila, H.-J. Kim, V. Koivunen, Robust iterative hard thresholding for compressed sensing, in: 2014 6th International Symposium on Communications, Control and Signal Processing, ISCCSP, IEEE, 2014, pp. 226–229.

[40] J.T. Parker, P. Schniter, V. Cevher, Bilinear generalized approximate message passing—part I: Derivation, IEEE Trans. Signal Process. 62 (22) (2014) 5839–5853.

[41] Q. Wan, H. Duan, J. Fang, H. Li, Z. Xing, Robust Bayesian compressed sensing with outliers, Signal Process. 140 (2017) 104–109.

[42] R. Gribonval, M. Nielsen, Sparse representations in unions of bases, IEEE Trans. Inform. Theory 49 (12) (2003) 3320–3325.

[43] M. Elad, Optimized projections for compressed sensing, IEEE Trans. Signal Process. 55 (12) (2007) 5695–5702.
[44] R. Barakat, First-order statistics of combined random sinusoidal waves with applications to laser speckle patterns, Opt. Acta 21 (11) (1974) 903–921.
[45] R. Barakat, J. Cole III, Statistical properties of N random sinusoidal waves in additive Gaussian noise, J. Sound Vib. 62 (3) (1979) 365–377.
[46] A. Papoulis, S. Unnikrishna Pillai, Probability, Random Variables and Stochastic Processes, fourth ed., McGraw-Hill, New York, NY, USA, 2002.
[47] D. Prasanna, C. Sriram, C.R. Murthy, On the identifiability of sparse vectors from modulo compressed sensing measurements, IEEE Signal Process. Lett. 28 (2020) 131–134.
[48] O. Musa, P. Jung, N. Goertz, Generalized approximate message passing for unlimited sampling of sparse signals, in: 2018 IEEE Global Conference on Signal and Information Processing, GlobalSIP, IEEE, 2018, pp. 336–340.
[49] J. Clausen, Branch and Bound Algorithms-Principles and Examples, Department of Computer Science, University of Copenhagen, 1999, pp. 1–30.
[50] H. Keshmiri, H. Bakhshi, A new 2-phase optimization-based guaranteed connected target coverage for wireless sensor networks, IEEE Sensors J. 20 (13) (2020) 7472–7486.
[51] T.R. Fernandes, B. Venkatesh, M.C. de Almeida, Distribution system topology identification via efficient milp-based wlav state estimation, IEEE Trans. Power Syst. 38 (1) (2022) 75–84.
[52] L. Scavuzzo, K. Aardal, N. Yorke-Smith, Learning optimal objective values for milp, 2024, arXiv preprint arXiv:2411.18321.
[53] D.P. Wipf, B.D. Rao, Sparse Bayesian learning for basis selection, IEEE Trans. Signal Process. 52 (8) (2004) 2153–2164.
[54] A. Bhandari, Unlimited sampling with sparse outliers: Experiments with impulsive and jump or reset noise, in: ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP, 2022, pp. 5403–5407.
[55] B. Widrow, I. Kollar, M.-C. Liu, Statistical theory of quantization, IEEE Trans. Instrum. Meas. 45 (2) (1996) 353–361.
[56] M. Abramowitz, I.A. Stegun, et al., Handbook of mathematical functions, vol. 55, Dover, New York, 1964.