Computational Array Signal Processing via Modulo Non-Linearities

Fernandez-Menduina, Samuel; Krahmer, Felix; Leus, Geert; Bhandari, Ayush

DOI
10.1109/TSP.2021.3101437

Publication date
2022

Document Version
Final published version

Published in
IEEE Transactions on Signal Processing

Computational Array Signal Processing via Modulo Non-Linearities

Samuel Fernández-Menduiña, Felix Krahmer, Geert Leus, and Ayush Bhandari

***Abstract*—Conventional literature on array signal processing (ASP) is based on the "capture first, process later" philosophy and to this end, signal processing algorithms are typically decoupled from the hardware. This poses fundamental limitations because if the sensors result in information loss, the algorithms may no longer be able to achieve their guaranteed performance. In this paper, our goal is to overcome the barrier of information loss via sensor saturation and clipping. This is a significant problem in application areas including physiological monitoring and extra-terrestrial exploration where the amplitudes may be unknown or larger than the dynamic range of the sensor. To overcome this fundamental bottleneck, we propose "computational arrays" which are based on a co-design approach so that a collaboration between the sensor array hardware and algorithms can be harnessed. Our work is inspired by the recently introduced unlimited sensing framework. In this context, our computational arrays encode the high-dynamic-range information by folding the signal amplitudes, thus introducing a new form of information loss in terms of the modulo measurements. On the decoding front, we develop mathematically guaranteed recovery algorithms for spatio-temporal array signal processing tasks that include DoA estimation, beamforming and signal reconstruction. Numerical examples corroborate the applicability of our approach and pave a path for the development of novel computational arrays for ASP.**

***Index Terms*—Array signal processing, direction of arrival (DoA) estimation, multi-channel sampling, non-linear sensing.**

I. Introduction

IN THE context of his pioneering apparatus for transatlantic wireless communication, G. Marconi also invented heuristics for direction-of-arrival (DoA) estimation [5]. Over the course of the century, these accomplishments were catalyzed by different breakthroughs in the fields of physics, mathematics and engineering and culminated in to what is known as *array signal processing* (ASP). At the heart of ASP [6] is the idea that spatio-temporal signal acquisition using multiple sensors offers unprecedented advantages when combined with high-resolution algorithms. Fast-forward to the last decade, several research contributions have led to advancements when it comes to exploring new array geometries [7]–[9] and developing high-resolution algorithms [10], [11]. That said, many of these advances are pivoted towards the development of the algorithmic machinery. Despite the remarkable progress on this front, algorithms are often *decoupled* from the sensor array hardware. In contrast, in this paper, we propose *computational arrays* where a part of the algorithmic intelligence is built-in to the hardware, which results in non-linearly encoded measurements. For the decoding step, we propose an algorithm that is referred to as *Unlimited Sensing based Array Signal Processing* or US-ASP. As we will see, the rationale behind this approach is that it allows us to recover signals that are much larger in amplitude than the dynamic range of the conventional sensor array.

A fundamental bottleneck in Array Signal Processing. In sensor arrays, digitization of the impinging signals is carried out by analog-to-digital converters (ADCs). This is what has revolutionized the field of ASP because digital signal processing algorithms can be employed for ASP related tasks such as DoA estimation, beamforming and denoising. In practice, however, a signal's amplitude range may be unknown and possibly, much larger than the dynamic range of the ADC. These scenarios result in clipping and saturation at the sensor array. In such cases it is natural that the algorithms, that are typically decoupled from the hardware, would completely deviate from their theoretically guaranteed performance. This is illustrated by the following two examples.

- Scientific probes in space exploration rely on sensor arrays for source localization and sub-surface mapping [12]. Nonetheless, in foreign environments, the range of signal amplitudes is unknown, which is typically handled by employing automatic gain control (AGC) systems. In this sense, NASA's Apollo Mission report [13] highlights the sensor saturation problem (cf. pg 43, [13]) and elaborates on the omnipresent use of AGCs. Even if the ADCs (equipped with AGCs) are calibrated, bursts and spikes [14] can saturate the sensor array, resulting in clipped measurements and permanent information loss.
- Another example of sensor array saturation stems from the *near-far* problem. Consider the case when two emitters are present; one of them being much closer to the receiver than

Manuscript received August 19, 2020; revised March 22, 2021; accepted July 22, 2021. Date of publication March 1, 2022; date of current version May 5, 2022. This work was supported in part by the U.K. Research and Innovation Council's Future Leaders Fellowship program "Sensing Beyond Barriers" (MRC Fellowship MR/S034897/1) and the European Partners Fund. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Nicolas Dobigeon. A part of this work was presented at the 28th European Signal Processing Conference, EUSIPCO 2020. [DOI: 10.23919/Eusipco47968.2020.9287595]*(Corresponding author: Ayush Bhandari.)*

Samuel Fernández-Menduiña and Ayush Bhandari are with the Department of Electrical and Electronic Engineering, Imperial College London, London SW7 2AZ, U.K. (e-mail: sf219@ic.ac.uk; ayush@alum.mit.edu).

Felix Krahmer is with the Department of Mathematics, Technical University of Munich, 85748 Garching/Munich, Germany (e-mail: felix.krahmer@tum.de).

Geert Leus is with the Faculty of Electrical Engineering, Mathematics and Computer Science, Delft University of Technology, 2826 CD Delft, The Netherlands (e-mail: g.j.t.leus@tudelft.nl).

Digital Object Identifier 10.1109/TSP.2021.3101437

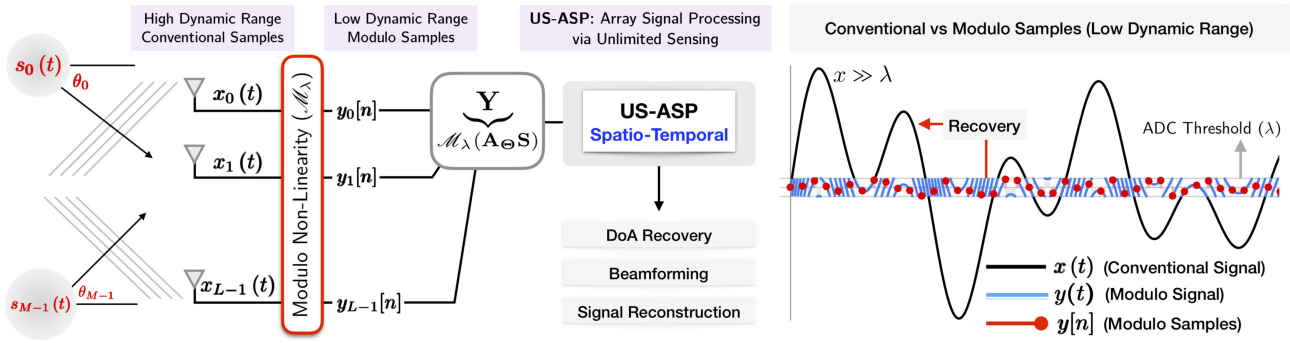

Fig. 1. Computational array signal processing setup using the unlimited sensing architecture [2]–[4]. Modulo non-linearity maps high-dynamic-range, *sensor array* samples into low-dynamic-range *folded* samples. While the modulo operation prevents the sensor saturation problem, it leads to a new form of information loss which can be efficiently handled by capitalizing on the joint application of unlimited sampling and array signal processing techniques. This unified approach allows to carry out the estimation of the directions of arrival, beamforming and sample reconstruction from the folded samples.

the other. Then, the sensor can either focus on the near-field emitter, drowning the far-field emitter in quantization noise, or aim at retrieving the information of the far-field emitter, clipping the samples of the near-field emitter [15]. A general trend in the recent years— especially in wireless communications—has been to use ADCs which can work with *wideband* receivers. It is well established that this requires a higher dynamic range [16], [17]. For instance, in GSM systems, to be able to cope up with the *near-far* problem, ADCs are required to digitize signals in the interval $[-104, -13]$ dBm, that is 91 dB dynamic range, thus posing serious hardware constraints.

Despite the pervasiveness of the sensor saturation problem and the significant advancements in ASP, very few papers tackle the problem of dynamic range constraints. In fact, existing approaches are mainly focused on the idea of one-bit ADC based DoA estimation [18]. However, due to the loss of crucial information, signal recovery and beamforming are hindered, affecting the channel capacity of the system [19]. This class of low-complexity solutions may be aptly justified for localization purposes, but in many applications, signal recovery is of utmost importance. For example, in space exploration, experiments entail exorbitant costs and are non-repeatable. Similarly, in bio-medical [20] and ultrasound [21] applications, the signal itself encodes important information.

A. Towards Computational Arrays via Unlimited Sensing

Our *computational arrays* overcome the limitations in the existing literature by relying on the *co-design* of the sensor array hardware and the algorithms, which is inspired by the *unlimited sensing* approach [2]–[4]. Instead of working with conventional, point-wise samples which may be clipped, we propose computational arrays that are capable of folding amplitudes in the interval $[-\lambda, \lambda]$. This is accomplished by injecting a modulo non-linearity in the sensing process that results in folded measurements, as shown in Fig. 1. Mathematically, this can be written as a *folding operator*,

$$\mathscr{F}_\lambda : x \mapsto 2\lambda \left(\left[\!\!\left[\frac{x}{2\lambda} + \frac{1}{2} \right]\!\!\right] - \frac{1}{2} \right),$$

where $\mathscr{F}_\lambda(\cdot)$ is the conventional, centered modulo operation, and $[\![x]\!] \stackrel{\text{def}}{=} x - \lfloor x \rfloor$ denotes the fractional part of x. Thanks to the recent advances in ADC design technology, such non-linearities can be implemented by re-purposing folding or self-reset ADCs [22], [23]. Recently, a modulo sampling ADC was reported in [24] which also provides a first validation of the unlimited sensing approach. Even when the amplitude of a bandlimited signal significantly exceeds the modulo threshold, the following sampling theorem guarantees recovery of the signal from modulo samples via an appropriate algorithm.

Theorem 1 (Unlimited Sampling Theorem [2], [3]): *Let $x(t)$ be a continuous-time, finite-energy, bandlimited function with maximum frequency Ω and let $y[n] = \mathscr{F}_\lambda(x(nT))$ be its modulo samples with sampling rate T. Then, a sufficient condition for recovery of $x(t)$ from its modulo samples (up to an additive multiple of 2λ) is $T \leqslant 1/2\Omega e$.*

A distinct feature of this sampling theorem is that only depends on the bandwidth and is independent of λ, retaining the simplicity of the Nyquist-Shannon sampling theorem. In this work, we capitalize on the modulo architecture and the underpinning sampling theorem to develop theory and algorithms for *array signal processing*. In particular, we study spatio-temporal recovery conditions with respect to modulo samples that allow for a general solution to the DoA estimation problem and beamforming. Our setup is flexible in that it covers the case of arbitrary array geometries, including linear, non-linear and sparse arrays [7].

B. Advantages of Computational–Array Signal Processing

Our novel computational-array based approach enjoys several advantages, both at the encoding and the decoding fronts.

1) Encoding via Computational-Arrays: This novel approach allows us to encode high-dynamic-range samples as low-dynamic-range measurements in the interval $[-\lambda, \lambda]$, thus circumventing the saturation or clipping problem.

2) Decoding via US-ASP: Our approach tailored for ASP, henceforth US-ASP, has the following advantages,

- *Decoding without Unfolding.* We are able to perform ASP tasks such as DoA estimation and beamforming without having to recover the signal from folded samples.

- *Backwards compatibility with conventional ASP.* Our recovery approach can be combined with any known ASP technique. Coprime arrays are an example of this property.

C. Contributions

The overarching goal of this paper is to formalize the inverse problem of modulo sensing based computational arrays. Our work develops a theoretical framework for mathematically guaranteed recovery algorithms. We study both *spatial* and *temporal* ASP methods comprising of three components: (1) DoA estimation, (2) beamforming and (3) signal reconstruction. Our key contributions are as follows:

C_1) **Recovery guarantees:** In Theorem 2 and Theorem 3, we establish recovery guarantees for temporal and spatial domains, respectively, so that conventional ASP methods can be readily applied. An interesting aspect of our work is that the spatial and temporal sampling rates are independent of the modulo threshold λ.

C_2) **Recovery algorithms:** We develop an efficient and empirically stable algorithmic framework for DoA estimation, beamforming and signal recovery from modulo samples, relying on the properties of the array response.

C_3) **Extension to sparse arrays:** We consider sparse arrays showing that our methods are backwards compatible with existing array signal processing approaches.

II. Preliminaries

Notation: We use $\mathbb{R}, \mathbb{Z}, \mathbb{C}$ and $\mathbb{N}$ to denote the set of reals, integers, complex and natural numbers, respectively. For $x \in \mathbb{R}$, we define the floor and the ceiling operations by $\lfloor x \rfloor = \sup\{n \in \mathbb{Z} : n \leqslant x\}$ and $\lceil x \rceil = \inf\{n \in \mathbb{Z} : n \geqslant x\}$, respectively. For $I \in \mathbb{R}$, we use $\lfloor \mathbb{I} \rceil$ to denote the non-negative integers until I, that is, $\{n : n \in \mathbb{Z} \cap [0, I)\}$. The cardinality of a set $\mathbb{I}$ is denoted by $\#\mathbb{I}$. We write the complex number $x \in \mathbb{C}$ as $x = \mathsf{Re}(x) + \jmath\mathsf{Im}(x)$, and the complex conjugate is denoted as x^*. Continuous functions and discrete sequences are represented by $x(t), t \in \mathbb{R}$ and $x[n], n \in \mathbb{Z}$, respectively. Matrices and vectors are written in capital and small boldface fonts, respectively. Vectors are assumed to be arranged in columns. Matrix element of $\mathbf{X}$ at index (m, n) is denoted as $[\mathbf{X}]_{m,n}$ and similarly, $[\mathbf{x}]_m$ for vector $\mathbf{x}$. We use $\mathbf{X}^{\mathrm{H}}$ and $\mathbf{X}^\top$ to denote conjugate-transpose and transpose of matrix $\mathbf{X}$, respectively. We use $\mathrm{span}(\mathbf{X})$ to denote the column span of a matrix. The maximum value of $\mathbf{X}$ is denoted by, $\|\mathbf{X}\| = \max|[\mathbf{X}]_{m,n}|$. For complex-valued matrices, we also define, $\|\mathbf{X}\|_{\mathsf{Re}} = \max|[\mathsf{Re}(\mathbf{X})]_{m,n}|$ and $\|\mathbf{X}\|_{\mathsf{Im}} = \max|[\mathsf{Im}(\mathbf{X})]_{m,n}|$ for the real and imaginary parts of $\mathbf{X}$, respectively. The Kronecker product between vectors is denoted as $\otimes$, and the Hadamard product is denoted as $\circ$. The covariance of a random matrix $\mathbf{X}$ is written as $\mathcal{R}(\mathbf{X}) = \mathcal{E}(\mathbf{X}\mathbf{X}^H)$, where $\mathcal{E}(\cdot)$ denotes the expectation operator. Function and sequence spaces are denoted by L^p and ℓ^p, respectively and the corresponding norms are defined by $\|\cdot\|_{L_p(\mathbb{R})}$ and $\|\cdot\|_{\ell_p(\mathbb{R})}$. When $p = \infty$, the norms denote the max-norm. A function x bandlimited to maximum frequency Ω is denoted by $x \in \mathcal{B}_\Omega$ while $x \in \mathsf{PW}_\Omega$ denotes a function in the Paley–Wiener class of bandlimited and square-integrable functions, that is, $x \in \mathcal{B}_\Omega \cap L^2$. Let $\mathbf{C}^K$ denote the space of functions with K continuous derivatives. The K^{th} order derivative of $x(t) \in \mathbf{C}^K$ is denoted as $x^{(K)}(t)$. Similarly, the K^{th} order finite difference of the sequence $x[n]$ is recursively defined as, $(\Delta^K x)[n] = \Delta^{K-1}(\Delta x)[n]$, $K > 1$ and for $K = 1$, $(\Delta x)[n] = x[n+1] - x[n]$. More directly, $(\Delta^K x)[n] \stackrel{\text{def}}{=} \sum_{k=0}^{K} (-1)^{K-k} \binom{K}{k} x[n+k]$ where $\binom{K}{k}$ is the binomial coefficient. For implementing differences using matrices, we define the right difference matrix $\mathbf{D}_N^1 \in \mathbb{R}^{N \times (N-1)}$ with matrix element, $[\mathbf{D}_N^1]_{m,n} = \delta[m-n-1] - \delta[m-n]$ where $\delta[n]$ is the Kronecker Delta symbol. Higher order difference matrix of order K is obtained via the recursion,

$$\mathbf{D}_N^K = \mathbf{D}_N^{K-1} \mathbf{D}_{N-K+1}^1 \in \mathbb{R}^{N \times (N-K)}. \tag{1}$$

For a matrix $\mathbf{X} \in \mathbb{C}^{L \times N}$, the K^{th} order finite-difference over its rows and columns is given by, $\mathbf{X}\mathbf{D}_N^K$ and $(\mathbf{D}_L^K)^\top \mathbf{X}$, respectively. For $\alpha \in (0,1)$ and $\beta > \lambda$, we define a quantity we refer to as the *packing quotient*,

$$\mathcal{L}_\lambda(\alpha, \beta) \stackrel{\text{def}}{=} \left\lceil \frac{\log(\lambda) - \log(\beta)}{\log(\alpha)} \right\rceil. \tag{2}$$

Problem Setup for ASP: Mathematical models for ASP are well investigated. Typically, the l^{th} sensor input is modeled as a linear combination of M narrow band sources, $\{s_m(t)\}_{m \in \lfloor \mathbb{M} \rceil}$, with wavelength $\nu = (2\pi c)/\omega$. More precisely, one assumes that

$$x_l(t) = \sum_{m \in \lfloor \mathbb{M} \rceil} a_l(\theta_m, \phi_m)\, s_m(t) \tag{3}$$

where the elevation $\mathbf{\Phi} = \{\phi_m\}_{m \in \lfloor \mathbb{M} \rceil}$ and the azimuth $\mathbf{\Theta} = \{\theta_m\}_{m \in \lfloor \mathbb{M} \rceil}, \theta_m \neq \phi_m$, are the unknown parameters that encode the DoAs and the array steering vector $\mathbf{a} = [a_0, \ldots, a_{L-1}]^\top$ is defined below. We assume ideal sampling, that is, $x[n] = x(nT)$ where $T > 0$ is the sampling interval. For M sources, L sensors and N time-domain samples this assumption gives rise to the matrix equation

$$\underbrace{\mathbf{X}}_{\mathbb{C}^{L \times N}} = \underbrace{\mathbf{A}}_{\mathbb{C}^{L \times M}} \underbrace{\mathbf{S}}_{\mathbb{C}^{M \times N}}, \tag{4}$$

where

1) $x_l[n] = x_l(nT), n \in \lfloor \mathbb{N} \rceil$, is the sampled waveform stacked in the *data matrix* $\mathbf{X}$ as $[\mathbf{X}]_{l,n} = x_l[n]$.
2) $a_l(\theta_m, \phi_m) = [\mathbf{a}(\theta_m, \phi_m)]_l$ is the l^{th} entry of the *array steering vector* which is defined using the convention,

$$\mathbf{a}(\theta_m, \phi_m) = \left[e^{\jmath \mathbf{d}_0^\top \mathbf{u}(\theta_m, \phi_m)} \ \ldots \ e^{\jmath \mathbf{d}_{L-1}^\top \mathbf{u}(\theta_m, \phi_m)} \right]^\top$$

where $\mathbf{u}(\theta, \phi) = (2\pi/\nu)[\sin\theta\cos\phi \quad \cos\theta\cos\phi \quad \sin\phi]^\top$ and $\mathbf{d}_l = [d_{l,x} \quad d_{l,y} \quad d_{l,z}]^\top$, $l \in \lfloor \mathbb{L} \rceil$ denote the 3D coordinates. The resulting matrix is,

$$\mathbf{A} = \left[\mathbf{a}(\theta_0, \phi_0) \ \cdots \ \mathbf{a}(\theta_{M-1}, \phi_{M-1}) \right]. \tag{5}$$

3) $s_m[n] = s_m(nT)$, $m \in \lfloor \mathbb{N} \rceil$, is the narrowband waveform stacked in the *source matrix* $\mathbf{S}$ as $[\mathbf{S}]_{m,n} = s_m[n]$.

• **Linear arrays:** When working with *linear arrays*, elements are placed along the x-axis *i.e.,* $\forall l, d_{l,y} = d_{l,z} = 0$. In case of Uniform Linear Arrays (ULAs), we will assume that the elements are equally spaced, *i.e.,* $d_{l,x} = ld$, with d the constant spacing. Furthermore, when working with linear arrays we will

assume that the impinging signals have zero elevation *i.e.,* $\forall m$ we have $\phi_m = 0$. This results in the simplified notation $\mathbf{A} = \mathbf{A}_{\boldsymbol{\Theta}}$,

$$[\mathbf{a}(\theta_m)]_l = \xi_m^{2l} \ \forall l \in \lfloor \mathbb{L} \rceil, \text{ where } \xi_m = e^{\jmath \frac{\pi d}{v} \sin(\theta_m)} \quad (6)$$

$$\mathbf{A}_{\boldsymbol{\Theta}} = \left[\mathbf{a}(\theta_0) \ \cdots \ \mathbf{a}(\theta_{M-1}) \right]. \quad (7)$$

In the conventional ASP setup (4), the goals are to,

1) Estimate the DoAs $\boldsymbol{\Theta}$ (and $\boldsymbol{\Phi}$ for non-linear arrays) from the data matrix $\mathbf{X}$.
2) Reduce interference by means of *beamforming* defined by,

$$\mathcal{B} : \mathbf{X} \in \mathbb{C}^{L \times N} \mapsto \mathbf{W}^{\mathsf{H}} \mathbf{X} \in \mathbb{C}^{L' \times N}, \quad (8)$$

where L' is the number of desired sources.

• **Sparse arrays:** In the context of our US-ASP, we will also consider *sparse arrays*—linear arrays whose inter-element spacing is larger than $\nu/2$. Namely, we will focus on *coprime arrays* [7] and two level *nested arrays* [9]. In the coprime case, we will follow the *difference coarray approach*. Let L_a and L_b be two coprime integers with $L_a < L_b$ and let d be the inter-sensor spacing. Let $\mathbb{L}$ be the set with sensor positions, $\mathbb{L} \stackrel{\text{def}}{=} \{dL_a n : n \in \lfloor \mathbb{L}_b \rceil\} \cup \{dL_b m : m \in \lfloor \mathbb{L}_a \rceil\}$. Letting $l_p \in \mathbb{L}$ for all $p \in [0, L_a + L_b - 1]$, the response is given by $[\mathbf{A}_{\boldsymbol{\Theta}}]_{p,m} = e^{\jmath l_p / \nu \sin(\theta_m)}$, for all $m \in \lfloor \mathbb{M} \rceil$, and $\mathbf{X} = \mathbf{A}_{\boldsymbol{\Theta}} \mathbf{S}$. Then, the entry at the position (i, k) of the covariance matrix $\mathcal{R}(\mathbf{X})$ is phased as in an array with an element in $l_i - l_k$. Thus, we can access a virtual array with elements at the difference set, which allows us to extend the number of *degrees of freedom* or DoF, henceforth.

Two level nested arrays [25] represent the union of a dense ULA with L_a elements, and a sparse ULA with L_b elements, such that $L = L_a + L_b - 1$. The spacing of the sensors in the sparse subarray equals the aperture of the dense subarray. Hence, the difference coarray has elements at $\mathbb{P} = \{nd_1\}_{|n| \leqslant L_T}$, with $L_T = L_b(L_a + 1) - 1$.

For these array constructions, vectorizing the covariance shows the link between data and the DoAs [26],

$$\mathbf{z} = \text{vec}\left(\mathcal{R}(\mathbf{X})\right) = \widetilde{\mathbf{A}}_{\boldsymbol{\Theta}} \mathbf{b}, \ \widetilde{\mathbf{A}}_{\boldsymbol{\Theta}} = \left[\widetilde{\mathbf{a}}(\theta_0) \ \cdots \ \widetilde{\mathbf{a}}(\theta_{M-1}) \right] \quad (9)$$

where, $\widetilde{\mathbf{a}}(\theta) = \mathbf{a}^*(\theta) \otimes \mathbf{a}(\theta)$, and $[\mathbf{b}]_m = \mathcal{E}(|s_m[n]|^2)$ for all $m \in \lfloor \mathbb{M} \rceil$. Given $\mathbf{z}$, methods such as spatial smoothing [9] allow conventional algorithms to retrieve the DoAs.

III. Towards Unlimited Sensing Based ASP

Departing from the conventional ASP setup, here we consider modulo measurements that are based on the acquisition model shown in Fig. 1. Mathematically, we have

$$y_l[n] = \mathscr{M}_\lambda(x_l(nT)), \quad T > 0, \quad (10)$$

where $\mathscr{M}_\lambda(\cdot)$ is the complex–valued variant of the modulo operation in [3] which is defined by the non-linear mapping,

$$\mathscr{M}_\lambda(z) : z \in \mathbb{C} \to \mathscr{F}_\lambda(\mathsf{Re}(z)) + \jmath \mathscr{F}_\lambda(\mathsf{Im}(z)). \quad (11)$$

The mapping $\mathscr{M}_\lambda(\cdot)$ in (10) ensures that the resulting modulo samples $\|y_l\|_{\ell_\infty(\mathbb{C})} \leq \lambda$. In our case, the inverse problem boils down to efficiently recovering the received signal and some or all of the transmitted sequences, in its sampled ($\mathbf{X}$ and $\mathbf{S}$) or continuous-time versions, from modulo measurements,

$$\mathbf{Y} \stackrel{\text{def}}{=} \mathscr{M}_\lambda(\mathbf{X}) = \mathscr{M}_\lambda(\mathbf{AS}). \quad (12)$$

In what follows, we will develop mathematically guaranteed techniques that allow for recovery of $\mathbf{X}$ and parameters thereof, so that conventional, spatio-temporal, Array Signal Processing techniques can be applied.

TABLE I
US–ASP: Model Setup

	Temporal Approach	Spatial Approach
Sources	$s_m \in \mathsf{PW}_\Omega$, $\forall m$.	Bounded s_m, $\forall m$.
Array	Arbitrary array, $\mathbf{A}$.	ULA, $\mathbf{A}_{\boldsymbol{\Theta}}$.
Azimuth $\boldsymbol{\Theta}$	$\theta_m \in [-\pi, \pi]$, $\forall m$.	$\theta_m \in [-\frac{\pi}{2}, \frac{\pi}{2}]$, $\forall m$.
Elevation $\boldsymbol{\Phi}$	$\phi_m \in [0, \pi]$, $\forall m$.	$\phi_m = 0$, $\forall m$.
Difference	$\mathbf{X}_K = \mathbf{X}\mathbf{D}_N^K$.	$\mathbf{X}_K = \left(\mathbf{D}_L^K\right)^\top \mathbf{X}$.

A. Overview of Our Approach

Here, we briefly explain the mechanistic principles that allow us to access the signal or its subspaces from the data matrix $\mathbf{Y}$ in (12), despite the loss due to the modulo operation. This is crucial for the subsequent ASP tasks to be meaningful. We proceed along similar lines as in [3], but adapted to matrices, and state the following observation.

Proposition 1 (Modulo Decomposition Property): Let $\mathbf{X}$ be an arbitrary matrix, $\mathbf{Y} = \mathscr{M}_\lambda(\mathbf{X})$ where λ is a fixed, non-zero constant. Then, $\mathbf{X}$ admits a unique decomposition,

$$\mathbf{X} = \mathscr{M}_\lambda(\mathbf{X}) + \Pi(\mathbf{X}) = \mathbf{Y} + \Pi(\mathbf{X}) \quad (13)$$

where $\Pi(\mathbf{X})$ is a matrix such that, $[\Pi(\mathbf{X})]_{m,n} \in 2\lambda\mathbb{Z}$.

There on, our strategy is to apply a discrete differentiation operator (denoted by $\mathcal{D}(\cdot)$), followed by a modulo operation exploiting that $\mathbf{Z} = \mathscr{M}_\lambda(\mathcal{D}(\mathbf{Y}))$ has the same column space or row space (depending on the nature of the difference operator and assuming that $N > L$) as the original signal $\mathbf{X}$ but *not necessarily* agrees with $\mathbf{X}$ itself. In contrast to [3], we do not proceed by reconstructing $\mathbf{X}$ but rather apply ASP tools and techniques directly to $\mathbf{Z}$, most notably *DoA estimation* and *beamforming*. The intuition behind the subspace preservation is the following. As $\mathbf{X}$ is a smooth signal, its columns (temporal ASP) and rows (spatial ASP) are highly correlated. This implies that the finite differences between its samples (column-wise or row-wise) are going to result in a shrinkage of amplitudes. The extent of shrinkage depends on the normalized sampling rate (cf. Lemma 1). In particular, as shown in [3], applying a finite difference of order K shrinks the amplitude by T^K. This allows for the finite difference and the modulo operator to commute in a certain sense. As a result of this commutativity, we can directly access the row or column space of the signal. The special structure of the matrix $\mathbf{A}$ in (4) allows for the recovery of the DoAs just from this subspace information without first recovering $\mathbf{X}$. In what follows, we will formally develop our results for the temporal and spatial cases in Sections III-B and III-C, respectively. The data model for the spatio-temporal approaches is summarized in Table I.

Algorithm 1: Temporal US-DoA Estimation.

Data: Matrix of modulo samples $\mathbf{Y}$ in (12).
Number of sources M.
An estimate on the supreme,
$\mathsf{B}_x \geq \|\{\mathbf{x}_l\}_{l=0}^{L-1}\|_{L_\infty(\mathbb{R})}$.
Result: $\boldsymbol{\Theta} = \{\theta_m\}_{m=0}^{M-1}$, $\boldsymbol{\Phi} = \{\phi_m\}_{m=0}^{M-1}$.
Step 1 Compute $K \geq \left\lceil \frac{\log\lambda - \log \mathsf{B}_x}{\log(T\Omega e)} \right\rceil$.
Step 2 Compute $\mathbf{X}_K \stackrel{\text{def}}{=} \mathbf{X}\mathbf{D}_N^K = \mathscr{M}_\lambda(\mathbf{Y}_K)$.
Step 3 Obtain the covariance matrix $\mathcal{R}(\mathbf{X}_K) = \mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^{\mathsf{H}}$.
Step 4 Let the eigenvalues be sorted in decreasing order and $\mathbf{V} = [\mathbf{v}_1 \quad \mathbf{v}_2 \quad \dots \quad \mathbf{v}_L]$.
Step 5 Compute

$$P(\theta,\phi) = \left(\sum\nolimits_{l=L-M}^{L} \mathbf{v}_l^{\mathsf{H}} \mathbf{a}(\theta,\phi) \right)^{-1}.$$

Step 6 Set $\boldsymbol{\Theta}$ and $\boldsymbol{\Phi}$ to the M sets of azimuth and elevation angles that maximize $P(\theta,\phi)$.

B. Temporal Approach to US-ASP

In this section, we will consider finite differences along the temporal dimension of the sensor array. The model setup is depicted in Table I.

1) Recovery Conditions: The starting point for formalizing recovery guarantees is the following lemma that relates the max-norm of continuous derivatives ($x^{(K)}$) to the max-norm of finite-differences ($\Delta^K x$) in terms of sampling rate T. Here, we will consider a real-valued signal $x \in \mathbb{R}$ and later, the same approach is adapted to complex-valued $\mathbf{X}$ in (4).

Lemma 1 (Difference-Derivative Inequality [2]): For any $x(t) \in \mathbf{C}^K(\mathbb{R}) \cap L_\infty(\mathbb{R})$, its samples $x[n] = x(nT)$ satisfy, $\|\Delta^K x\|_{\ell_\infty(\mathbb{R})} \leqslant (Te)^K \|x^{(K)}(t)\|_{L_\infty(\mathbb{R})}$.

This result shows that oversampling $(Te < 1)$ shrinks the upper bound on $\|\Delta^K x\|_{\ell_\infty(\mathbb{R})}$. For practical purposes, estimating the maximum value of a function is easier than accessing its derivatives. To this end, we bound the max-norm of $x^{(K)}(t)$ by invoking the Bernšteĭn's inequality,

$$x \in \mathcal{B}_\Omega, \qquad \|x^{(K)}\|_{L_\infty(\mathbb{R})} \leqslant \Omega^K \|x\|_{L_\infty(\mathbb{R})}. \tag{14}$$

This result when combined with Lemma 1 yields [3],

$$\|\Delta^K x\|_{\ell_\infty(\mathbb{R})} \leqslant (T\Omega e)^K \|x\|_{L_\infty(\mathbb{R})}. \tag{15}$$

Let B_x, an upper bound on $|x(t)|$, be known. Choosing $T\Omega e < 1$ implies that $(T\Omega e)^K \mathsf{B}_x \leqslant \lambda$, for K given by

$$K \geq \left\lceil \frac{\log\lambda - \log \mathsf{B}_x}{\log(T\Omega e)} \right\rceil = \mathcal{L}_\lambda\left(T\Omega e, \mathsf{B}_x\right). \tag{16}$$

Therefore, with $T < 1/\Omega e$ and K in (16), it follows that,

$$\|\Delta^K x\|_{\ell_\infty(\mathbb{R})} \leqslant \lambda \Rightarrow \Delta^K x = \mathscr{M}_\lambda\left(\Delta^K x\right). \tag{17}$$

To relate higher order differences of x with the measurements y, we use the following proposition [3].

Proposition 2 Let $x[n] = x(nT)$ be samples of a bounded, bandlimited function $x \in \mathcal{B}_\Omega$ with sampling rate $T < 1/\Omega e$. Then, $\mathscr{M}_\lambda(\Delta^K x) = \mathscr{M}_\lambda(\Delta^K(\mathscr{M}_\lambda(x)))$ with K in (16).

Our objective now is to generalize this theory to the array scenario. Since the time-domain samples are arranged along the row-dimension of the $L \times N$ matrix $[\mathbf{X}]_{l,n} = x_l(nT)$, the difference operator acts over each row independently,

$$\left(\Delta^K x_l\right)(nT) = [\mathbf{X}_K]_{l,n} = \left[\mathbf{X}\mathbf{D}_N^K\right]_{l,n} \tag{18}$$

where $\mathbf{D}_N^K$ is the difference matrix defined in (1). Combining the result of Proposition 2 and (18), we obtain the link between higher order differences and the modulo samples,

$$\Delta^K x = \mathscr{M}_\lambda\left(\Delta^K y\right) \stackrel{(18)}{\longleftrightarrow} \mathbf{X}_K = \mathscr{M}_\lambda\left(\mathbf{Y}_K\right),\ \mathbf{Y}_K = \mathbf{Y}\mathbf{D}_N^K. \tag{19}$$

By combining the key ingredients of modulo sampling theory for the case of sensor arrays, we arrive at the following recovery guarantee that will be at the heart of the ASP tasks, in particular, multi-channel signal recovery.

Theorem 2 (Temporal US–ASP): *Let $N > K + M$ and let $s_m[n] = s_m(nT)$, $n \in \lfloor \mathbb{N} \rceil$, be the samples of $\{s_m\}_{m\in\lfloor \mathbb{M} \rceil} \in \mathcal{B}_\Omega$ with sampling rate T. Also, let $\mathbf{Y} = \mathscr{M}_\lambda(\mathbf{AS})$ in (12) be the modulo samples with $\mathbf{A}$ defined in (5). Provided that $T \leqslant 1/2\Omega e$, choosing $K \geqslant \mathcal{L}_\lambda(T\Omega e, \mathsf{B}_x)$ guarantees that $\mathbf{X}_K = \mathscr{M}_\lambda(\mathbf{Y}_K)$.*

Combined with existing results on unlimited sensing, this theorem directly yields a first recovery guarantee. As here, unlike the infinite-dimensional scenario in [3], we are working with finite-dimensional vectors, we will apply a local version of the unlimited sampling theorem as derived in [27], [28] in the context of recovery of sparse and parametric signals. More precisely, Theorem 2 in [27] establishes that

$$T \leqslant \frac{1}{2\Omega e} \qquad \text{and} \qquad \widetilde{N} \geqslant N + 7\frac{\mathsf{B}_x}{\lambda} \tag{20}$$

guarantees recovery of N, finitely many, contiguous samples of $x \in \mathcal{B}_\Omega$ from $\widetilde{N}$ modulo samples. Combining this result with Theorem 2 yields the following recovery condition.

Corollary 1: *Let $\{s_m(t)\}_{m\in\lfloor \mathbb{M} \rceil} \in \mathsf{PW}_\Omega$ with $[\mathbf{S}]_{m,n} = s_m(nT), T > 0$, $n \in \lfloor \mathbb{N} \rceil$. Then, a sufficient condition for recovery of $\{x_l(nT)\}_{l\in\lfloor \mathbb{L} \rceil}$, as in (4), from its modulo samples (up to an additive multiple of 2λ) is given by*

$$T \leq \frac{1}{2\Omega e} \qquad \text{and} \qquad \widetilde{N} > K + M + 7\frac{\mathsf{B}_x}{\lambda} \tag{21}$$

Here, in the spirit of Theorem 1, the sampling rate is still independent of λ, however, the number of modulo samples $(\widetilde{N})$ depends on the global dynamic range (B_x/λ).

2) DoA Estimation: To use Theorem 2 for DoA estimation, we build on the observation that the discrete derivative of the samples, as obtained by (19) gives rise to a DoA estimation problem with the same array response matrix, regardless of the geometry of the array. Indeed, we can write

$$\mathbf{X}_K \stackrel{\text{def}}{=} \mathscr{M}_\lambda\left(\mathbf{Y}_K\right) = \mathbf{X}\mathbf{D}_N^K = \mathbf{A}\left(\mathbf{S}\mathbf{D}_N^K\right) = \mathbf{A}\mathbf{S}_K, \tag{22}$$

and $\mathcal{R}(\mathbf{X}_K) = \mathbf{A}\mathcal{R}(\mathbf{S}_K)\mathbf{A}^H$, where $\mathbf{S}_K \stackrel{\text{def}}{=} \mathbf{S}\mathbf{D}_N^K$. Thus, conventional approaches to DoA estimation can be directly applied to $\mathbf{X}\mathbf{D}_N^K$ *without* recovering $\mathbf{X}$, saving computations in the reconstruction step. This will yield meaningful DoA estimates provided that $k \leq K$, $\operatorname{span}(\mathcal{R}(\mathbf{X})) = \operatorname{span}(\mathcal{R}(\mathbf{X}_k))$, which is

the case when both $\mathbf{S}$ and $\mathbf{S}_k = \mathbf{S}\mathbf{D}_N^k$ have full rank. A necessary condition to satisfy this property is $N - K > M$, which puts an additional constraint on the finite order difference $K < N - M$. As an example, the procedure to estimate the DoAs from modulo samples is shown in Algorithm 1 for the case of the MUSIC algorithm. In a similar fashion, other competitive algorithms are an equally viable option provided that specific conditions over M, N and L are satisfied [5]. For instance, for both MUSIC and ESPRIT algorithms, one needs that $L > M$. We remind the reader that when Theorem 2 is applicable, the sampling rate is independent of λ and hence, a constant-factor oversampling suffices for DoA estimation (directly from the folded measurements).

3) Beamforming and Recovery: Having estimated the DoAs, beamforming ($\mathcal{B}$) can be performed. Instead of recovering data matrix $\mathbf{X}$ from $\mathbf{X}_K$ (as in [27]), we can directly apply beamforming and recover the filtered signal. Provided that the conditions of Theorem 2 hold, the linearity of the beamforming operator allows us to write,

$$\mathcal{B}\left(\mathbf{X}\right) = \mathcal{U}\left(\mathcal{B}\left(\mathscr{M}_\lambda\left(\mathbf{Y}_K\right)\right)\right), \tag{23}$$

where $\mathcal{U}(\cdot)$ is some properly chosen reconstruction operator such that $\mathcal{U}(\mathbf{X}_K) = \mathbf{X}$ and $\mathcal{B}$ is the *beamformer* defined in (8). Despite the reconstruction operator may be highly non-linear, the property $\mathcal{U}(\mathbf{X}_K) = \mathbf{X}$ allows us to swap both operators without loss of information. The recovery approach in [29] is an example of such an operator when the signals are bandlimited. This result guarantees that the beamformed signal can be directly retrieved from the folded samples, saving computational time as compared to the full recovery method based on unfolding summarized in Corollary 1 above. On the other hand, the full recovery approach has the advantage that none of the rich signal features beyond the beamformed signal are lost in the algorithmic reconstruction.

4) Sparse Arrays: To extend our theory to sparse arrays, we first seek the conditions under which estimation from the covariance matrix of $\mathbf{X}_K$ is possible. To this end, consider the following result.

Corollary 2: *Let $\mathbf{X} = \mathbf{A}_{\boldsymbol{\Theta}}\mathbf{S}$, where $\mathbf{S}$ comprises of sampled bandlimited vectors $\{s_m\}_{m\in\lfloor\mathbb{M}\rceil}$ and where $\mathbf{A}_{\boldsymbol{\Theta}}$ is the sparse array response. Then, $\forall m \in \lfloor\mathbb{M}\rceil$, $\mathcal{E}(|\Delta^K s_m[n]|^2) \neq 0$ is a sufficient condition guaranteeing $\mathscr{E}(\mathbf{X}) = \mathscr{E}(\mathbf{X}_K)$ where $\mathscr{E}(\cdot)$ represents a DoA estimator for sparse arrays.*

This corollary is aptly justified by taking into account the expression for the coarray approach in (9), but considering $\Delta^K s_m[n]$ as the source signals. The above result in conjunction with Theorem 2 guarantees that the coarray approach can be applied to the folded measurements as well, leading to the following equation

$$\{s_m \in \mathcal{B}_\Omega\}_{m\in\lfloor\mathbb{M}\rceil}, \quad \mathscr{E}\left(\mathbf{X}\right) = \mathscr{E}\left(\mathscr{M}_\lambda\left(\mathbf{Y}_K\right)\right), \tag{24}$$

provided that $T < 1/2\Omega e$. This result allows to perform DoA estimation using any sparse array from the folded samples. Hence, sparsity-based DoA methods (including the ones based on compressed sensing) can straightforwardly be adapted to our new model since not only the signal subspace is retained but also the sparsity pattern revealing the DoAs.

C. Spatial Approach to US-ASP

In this section, we consider the application of finite differences along the spatial dimension of the signal.

1) Recovery conditions: The first key point of this approach is preservation of the array response under the difference operator. The result is stated in the following lemma.

Lemma 2: Let $\xi_m = \exp(\jmath\frac{\pi d}{\nu}\sin\theta_m)$ and let $\mathbf{A}_{\boldsymbol{\Theta}}$ denote the array response matrix with elements $[\mathbf{A}_{\boldsymbol{\Theta}}]_{l,m} = \xi_m^{2l}$. Then, for all $K \in \mathbb{N}$ such that $K < L$, we have,

$$\left(\mathbf{D}_L^K\right)^\top \mathbf{A}_{\boldsymbol{\Theta}} = \mathbf{I}_K \mathbf{A}_{\boldsymbol{\Theta}} \mathbf{E}^K \tag{25}$$

where, $\mathbf{I}_K \in \mathbb{R}^{(L-K)\times L}$ with $[\mathbf{I}_K]_{k,l} = \delta[k-l]$ and $\mathbf{E} \in \mathbb{C}^{M\times M}$ is a diagonal matrix, its diagonal elements are given by $[\mathbf{E}]_{m,m} = (-\jmath 2\mathsf{Im}(\xi_m))\xi_m$.

Proof: See Section VI. ∎

The implication of Lemma 2 is that we are able to write,

$$\mathbf{X}_K = \left(\mathbf{D}_L^K\right)^\top \mathbf{X} = \mathbf{I}_K \mathbf{A}_{\boldsymbol{\Theta}} \mathbf{E}^K \mathbf{S}. \tag{26}$$

This result is instrumental in ensuring that $\mathbf{X}_K$ can be upper bounded by λ or, $\|\mathbf{X}_K\|_{\mathsf{Re}} < \lambda$ and $\|\mathbf{X}_K\|_{\mathsf{Im}} < \lambda$ for $l \in \lfloor\mathbb{L} - \mathbb{K}\rceil$ and $n \in \lfloor\mathbb{N}\rceil$. The formal statement is as follows.

Lemma 3: Let $\mathbf{X} = \mathbf{A}_{\boldsymbol{\Theta}}\mathbf{S}$ where $\mathbf{S}$ is an arbitrary matrix and $\mathbf{A}_{\boldsymbol{\Theta}}$ a matrix with $[\mathbf{A}_{\boldsymbol{\Theta}}]_{l,m} = \xi_m^{2l}$, where $\xi_m = \exp(\jmath\frac{\pi d}{\nu}\sin\theta_m)$. Then, $\mathbf{X}_K$ in (26) admits the upper-bound,

$$K < L, \;\; \|\mathbf{X}_K\| \leqslant \sqrt{2}M\left(2\sin\left(\tfrac{\pi d}{\nu}\right)\right)^K \|\mathbf{S}\|.$$

Proof: See Section VI. ∎

The implication of Lemma 3 is that by suitably choosing the inter-element spacing d and difference order K, we can arbitrarily shrink $\|\mathbf{X}_K\|_{\mathsf{Re}}$ and $\|\mathbf{X}_K\|_{\mathsf{Im}}$. We would like to upper-bound them with λ. To this end,

$$2\sin\left(\pi d/\nu\right) < 1 \Rightarrow d < \nu/6 \tag{27}$$

and this provides an upper-bound on d, the *spatial sampling frequency*. Now, suppose that $\mathsf{B}_s \geq \max_{m\in\lfloor\mathbb{M}\rceil}\|s_m\|_{L_\infty(\mathbb{R})}$, is known. Then, using d in (27) and K given by,

$$K \geq \left\lceil \frac{\log(\lambda) - \log(M\mathsf{B}_s)}{\log\left(2\sin\left(\pi d/\nu\right)\right)} \right\rceil = \mathcal{L}_\lambda\left(2\sin\left(\frac{\pi d}{\nu}\right), M\mathsf{B}_s\right) \tag{28}$$

guarantees that $\|\mathbf{X}_K\|_{\mathsf{Re}} \leqslant \lambda$ and $\|\mathbf{X}_K\|_{\mathsf{Im}} \leqslant \lambda$. We can summarize our result in the following theorem.

Theorem 3 (Spatial US-ASP): *Let $\mathbf{X} = \mathbf{A}_{\boldsymbol{\Theta}}\mathbf{S}$, where,*

- *$\mathbf{A}_{\boldsymbol{\Theta}} \in \mathbb{C}^{L\times M}$ is a matrix with elements $[\mathbf{A}_{\boldsymbol{\Theta}}]_{l,m} = \xi_m^{2l}$, where $\xi_m = \exp(\jmath\frac{\pi d}{\nu}\sin\theta_m)$.*
- *$\mathbf{S} \in \mathbb{C}^{M\times N}$ is an arbitrary matrix, such that $\mathsf{B}_s \geq \|\mathbf{S}\|$.*

Furthermore, let $\mathbf{Y} = \mathscr{M}_\lambda(\mathbf{X})$. Then, provided that $d < \nu/6$ and choosing, $L > K \geq \mathcal{L}_\lambda(2\sin(\pi d/\nu), M\mathsf{B}_s)$ results in $\mathscr{M}_\lambda(\mathbf{Y}_K) = \mathbf{X}_K$, with $\mathbf{Y}_K = (\mathbf{D}_L^K)^\top\mathbf{Y}$.

Proof: See Section VI. ∎

Theorem 3 in conjunction with Lemma 2 yields,

$$\mathscr{M}_\lambda\left(\mathbf{Y}_K\right) \overset{(26)}{=} \mathbf{I}_K \mathbf{A}_{\boldsymbol{\Theta}} \mathbf{E}^K \mathbf{S}. \tag{29}$$

This gives a direct relationship between the folded samples $\mathbf{Y}$ and a scaled version of the transmitted symbols $\mathbf{E}^K\mathbf{S}$. It is worth highlighting that no row of the matrix $\mathbf{E}^K\mathbf{S}$ will be zero as long

as $[\mathbf{S}]_{m,n} \neq 0$ for all $n \in \lfloor \mathbb{N} \rceil$ and the DoAs are different than 0. Moreover, the array response is preserved, saving for the loss of the last K rows.

Remark on Array Spacing Criterion: With reference to the conventional half-wavelength spacing criterion ($d < \nu/2$), our condition $d < \nu/6$ in Theorem 3 requires spatial oversampling. In application areas such as bio-signal processing, seismic array processing and phased arrays, the recent trend has been to leverage spatial oversampling. This has been catalysed by advances in sensor design and availability of cheaper, more capable and efficient hardware. Concrete examples include, (a) Electroencephalography wherein Robinson *et al.* [30] use "super-Nyquist" density sensor placement and report its unique advantages. Similarly, "high density" electrode placement has been used in ECG acquisition [31]. (b) Seismic Imaging where "dense arrays" have enabled new applications [32]–[34] and (c) Phased Arrays where spatial oversampling [35] is used to exploit degrees of freedom while relaxing the circuit constraints.

2) DoA estimation: From our result above in (29), we can conclude that the information about the directions of arrival in $\mathbf{X}$ is still accessible from $\mathbf{Y}$. By defining $\widehat{\mathbf{A}}_{\boldsymbol{\Theta}} \stackrel{\text{def}}{=} \mathbf{I}_K \mathbf{A}_{\boldsymbol{\Theta}}$, we can model $\mathscr{M}_\lambda(\mathbf{Y}_K)$ as an equivalent system, where weighted versions of the original transmitted signals impinge on a ULA with the same inter-element separation and response $\widehat{\mathbf{A}}_{\boldsymbol{\Theta}}$, counting $L-K$ rows. Therefore, the DoAs can be estimated from the difference samples (29), as long as the corresponding relation between the parameters $L-K$ and M holds for the considered DoA estimation method. In this sense, our method is backwards compatible with any existing DoA estimation technique. For example, methods as MUSIC and ESPRIT can be readily applied if the condition $L > M + K$ holds.

3) Different approaches to recovery: Here, we leverage our previously estimated DoAs to address the recovery process for two different scenarios, namely, (a) beamformed version of $\mathbf{X}$ or (b) $\mathbf{X}$ itself. As will be shown, the latter is important due to stability issues related with the former.

Recovery via beamforming: Having estimated the DoAs, provided that $\forall m_l \in \lfloor \mathbb{M} \rceil$, $\theta_{m_1} \neq \theta_{m_2} \neq 0$, we can recover the samples via beamforming. In this setting, the structure of $\mathbf{A}_{\boldsymbol{\Theta}}$ allows us for a *non-iterative* reconstruction from the modulo samples. However, operations have to be performed on the modified data model $\mathbf{A}_K\mathbf{S}$, where $\mathbf{A}_K \stackrel{\text{def}}{=} \mathbf{I}_K \mathbf{A}_{\boldsymbol{\Theta}} \mathbf{E}^K$. Then, given a matrix beamformer constructed for $\mathbf{A}_K$, call it $\mathbf{W}_{\mathbf{E}}$, we can obtain the beamformed signal via

$$\mathbf{W}_{\mathbf{E}}^{\mathsf{H}} \mathscr{M}_\lambda\left(\mathbf{Y}_K\right) = \mathbf{W}_{\mathbf{E}}^{\mathsf{H}} \mathbf{A}_K \mathbf{S}. \tag{30}$$

In practice, one option is to consider a zero-forcing beamformer and hence compute $\mathbf{W}_{\mathbf{E}}^{\mathsf{H}}$ as the pseudo-inverse of $\mathbf{I}_K \mathbf{A}_{\boldsymbol{\Theta}} \mathbf{E}^K$. In this case, we can write

$$\mathbf{W}_{\mathbf{E}}^{\mathsf{H}}\, \mathscr{M}_\lambda\left(\mathbf{Y}_K\right) = \mathbf{E}^{-K} \left(\mathbf{I}_K \mathbf{A}_{\boldsymbol{\Theta}}\right)^{\dagger} \mathscr{M}_\lambda\left(\mathbf{Y}_K\right) \stackrel{(26)}{=} \mathbf{S}. \tag{31}$$

The construction of more sophisticated beamformers, that can handle the effects of the noise on modulo samples, is an open problem that we plan to tackle in the future.

Recovery via unfolding: Recovery in (30) relies on inverting $\mathbf{E}^K$. According to (32), the diagonal elements in $\mathbf{E}^{-K}$ contain $\left(2\sin(\pi d/\nu \sin(\theta_m))\right)^{-K}$ which may be ill-conditioned for higher difference orders required for processing of modulo samples. This problem can be circumvented by resorting to an alternative recovery approach that can handle finite-dimensional vectors [28] as long as the DoAs are non-zero. To this end, we again use the *local reconstruction theorem* stated in [27], [28], but applied to the spatial sampling case (unlike (20)). Having obtained $\mathbf{X}_K$ from $\mathbf{Y}$, we can recover L' unfolded spatial samples of $\mathbf{X}$, provided that

$$d < \nu/6 \quad \text{and} \quad L > L' + 7\mathsf{B}_x/\lambda.$$

This allows us to recover L' spatial samples of $\mathbf{X}$ before applying any beamformer, leading to a more stable recovery algorithm because the inversion avoids computation of the ill-posed matrix $\mathbf{E}^{-K}$ and uses the recovery method in [3].

Algorithm 2: Matrix Unfolding (Closed-Form).

Data: Matrix of modulo samples $\mathbf{Y}$.
Number of sources M.
Packing quotient K.
Difference matrix $\mathbf{X}_K$.
Step 1: Obtain the covariance matrix $\mathcal{R}(\mathbf{X}_{\mathbf{K}}) = \mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^{\mathsf{H}}$.
Step 2: Estimate the number of sources $M' = \text{rank}(\boldsymbol{\Lambda})$.
Step 3: Estimate the M' DoAs $\boldsymbol{\Theta}$.
Step 4: Recover the samples, $\widehat{\mathbf{X}}$, as in (30), using M'.
Step 5: If $M' = M$, then $\mathbf{X} \leftarrow \widehat{\mathbf{X}}$ and halt.
Step 6: Otherwise, let $\mathbf{Q} \leftarrow \mathbf{Y} - \widehat{\mathbf{X}}$.
Step 7: Take a decision in the 2λ grid about $\mathbf{Q}$ to obtain $\mathbf{G}$.
Step 8: Set $\mathbf{X} \leftarrow \widehat{\mathbf{X}} + \mathbf{G}$.

Zero recovery: Given $M, \mathbf{X}_K$, we can estimate the number of sources, say $M' \in \mathbb{N}$, using $\mathbf{X}_K$. Since $M = M' \Rightarrow \theta_m \neq 0$ and hence the process can be performed normally. In case $M' < M$, one of the angles is zero. Then, we can follow the same algorithm, but considering only M' sources, thus retrieving $\mathbf{X}' \in \mathbb{C}^{L \times N}$. When $\theta_m = 0$, the array response is an all-ones vector and the term $\mathbf{Q} = \mathbf{Y} - \mathbf{X}' = (\mathbf{X} - \mathbf{X}') + \Pi(\mathbf{X})$ has constant columns. Then, by making a decision in the value of each column of $\mathbf{Q}$, we obtain $\mathbf{G}$ and the unfolded matrix $\mathbf{X} = \mathbf{X}' + \mathbf{G}$ can be recovered up to additive integer multiples of 2λ. The complete recovery process is summarized in Algorithm 2. This prior knowledge of row/column structure of $\mathbf{Q}$ allows us to enhance algorithmic capability by using techniques such as projection over convex sets (POCS) [36].

Recovery of bandlimited signals: Theorem 3 provides all the necessary tools for guaranteeing the recovery of continuous bandlimited waveforms from folded samples, provided that the *temporal sampling rate* satisfies the conventional Nyquist criterion. This yields the following result.

Corollary 3: *Let* $\{s_m(t)\}_{m \in \lfloor \mathbb{M} \rceil} \in \mathsf{PW}_\Omega$ *with* $\mathsf{B}_\mathsf{s} \geq \max_{m \in \lfloor \mathbb{M} \rceil} \|s_m\|_{L_\infty(\mathbb{R})}$ *and also let* $[\mathbf{S}]_{m,n} = s_m(nT), T > 0$. *Under the conditions of Theorem 3, a sufficient condition for recovery of* $\{x(nT)\}_{l \in \lfloor \mathbb{L} \rceil}$, *as in (3), from its modulo samples (up to an additive multiple of* 2λ*) is that* $T \leqslant \pi/\Omega$.

TABLE II
US–ASP: UNLIMITED SENSING BASED ARRAY SIGNAL PROCESSING AND OVERVIEW OF MAIN RESULTS

	Temporal Approach (Section III-B)	**Spatial Approach** (Section III-C)
Sampling Rate	$T < 1/2\Omega e$	$d < \nu/6$
Definition of K	$K \geqslant \left\lceil \frac{\log \lambda - \log \mathsf{B}_x}{\log(T\Omega e)} \right\rceil$.	$K \geqslant \left\lceil \frac{\log \lambda - \log(M\mathsf{B}_s)}{\log(2\sin(\pi d/\nu))} \right\rceil$.
DoA Estimation	(22). Compatible with any technique given M, L.	(29). Compatible with any technique given $M, L-K$.
Beamformer	$\mathcal{B}(\mathbf{X}) \overset{(23)}{=} \mathcal{U}(\mathcal{B}(\mathscr{M}_\lambda(\mathbf{Y}_K)))$.	$\mathbf{W}_{\mathbf{E}}^{\mathsf{H}} \mathscr{M}_\lambda(\mathbf{Y}_K) \overset{(30)}{=} \mathbf{W}_{\mathbf{E}}^{\mathsf{H}} \mathbf{A}_K \mathbf{S}$.
Sparse arrays	*Corollary 2*: Compatible with sparse arrays using L physical sensors as long as the sources satisfy $\mathcal{E}\left(\lvert\Delta^K s_m[n]\rvert^2\right) \neq 0$ for all $m \in \lfloor \mathbb{M} \rceil$.	*Corollary 4*: Compatible with sparse arrays using $(K+1)L$ physical sensors. US nested arrays require only $L_1 + (K+1)L_2 - K$ physical sensors.

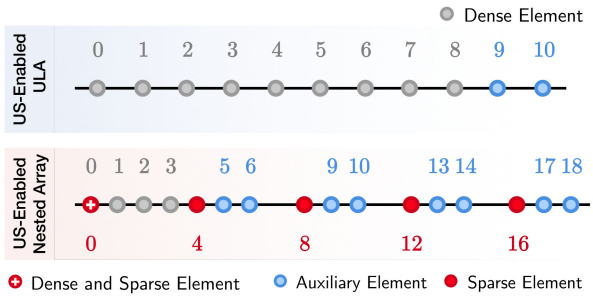

Fig. 2. Architectures for US-ASP using ULAs and nested arrays, for $K = 2$. The spacing between sensors is normalized by d. The information provided by the auxiliary sensors (sensors lost when performing the spatial difference) is employed to access the unfolded samples. (Top) ULA array with $L = 11$ physical sensors and 9 effective elements. (Bottom) Nested Array with 4 effective elements in the first level and 4 effective elements in the second level. As it comes clear from the figure, the auxiliar elements are only required in the sparse subarray.

Remarkably, both the spatial and the temporal sampling rate required in this result are agnostic to λ. Only the recovery procedure will crucially rely on the knowledge of λ.

Besides, as the functions in the Paley-Wiener class constitute a subspace, this result can be applied to any beamformed version of the rows of $\mathbf{X}$ or $\mathbf{X}_K$.

4) Sparse Arrays: Despite being stated for ULAs, Theorem 3 can be generalized to be backwards compatible with any array geometry. The following result develops this idea.

Corollary 4 (Arbitrary array geometry): *Consider a sensor geometry formed by L sensors at positions $\mathbf{d}_l$, $l \in \lfloor \mathbb{L} \rceil$, with array response $\widetilde{\mathbf{A}}$. Then, L ULAs counting with $(K+1)$ physical sensors each, with $K \geqslant \mathcal{L}_\lambda(2\sin(\frac{\pi d}{\nu}), M\mathsf{B}_s)$, inter-element spacing $d < \nu/6$ and array response $\mathbf{A}_l$, are sufficient for obtaining $\widetilde{\mathbf{X}} = \widetilde{\mathbf{A}}\widetilde{\mathbf{S}}$, where $\widetilde{\mathbf{S}}$ is a version of $\mathbf{S}$ where the rows are scaled, from the L sets of λ-folded samples $\{\mathbf{Y}_l = \mathscr{M}_\lambda(\mathbf{A}_l\mathbf{S})\}_{l\in\lfloor\mathbb{L}\rceil}$.*

The main drawback of this approach is the loss of KL degrees of freedom with respect to a ULA with the same number of physical sensors. However, a US-enabled nested array only needs $(K+1)(L_2-1)+L_1$ physical sensors, instead of $(K+1)(L_1+L_2)$, due to its special geometry. As explained in Fig. 2, we only require the auxiliary subarray in the sparse section of the nested array, simplifying the architecture. US enabled nested arrays have the advantage of enhancing the number of degrees-of-freedom retaining the flavor of conventional nested arrays.

For reader's convenience, we summarize our temporal and spatial domain results in Table II.

IV. NUMERICAL DEMONSTRATION

In this section, we provide examples of DoA estimation that confirm our theoretical results. Furthermore, we consider the case of noise, showing the empirical stability of our approach in the presence of perturbations. We mainly focus on subspace-based methods to illustrate the proposed ideas, which can be extended to compressed sensing based methods.

Experimental protocol: We consider 800 temporal samples. The bounds on the amplitude fulfill $\|\mathbf{X}\|_{\mathsf{Im}} = \|\mathbf{X}\|_{\mathsf{Re}} = 2$ for the time domain approach and $\|\mathbf{S}\|_{\mathsf{Im}} = \|\mathbf{S}\|_{\mathsf{Re}} = 2$ for the spatial approach. The maximum frequency is set to $\Omega = \pi$ rads/s. We list the experimental parameters in Table III. For each experiment, we start from the modulo data matrix $\mathbf{Y}$. Using the mathematically guaranteed methods presented in our discussion, we extract the DoA information from $\mathbf{Y}$. We compare our approach with the case when the conventional data matrix $\mathbf{X}$ is available.

A. Time Domain US-ASP

Experiment A: non-linear arrays: We consider a Uniform Circular Array with $M = 4$ sources arriving with non-zero azimuth and elevation. The result of MUSIC-based recovery using $\mathbf{X}$ and $\mathbf{Y}$ is shown in Fig. 3(A). The resulting spectrums were smoothed with a Gaussian kernel (3 samples width) to ease the visibility. These results confirm that temporal US-ASP is able to find sources in the 2D space using non-linear arrays.

Experiment B: empirical stability in presence of noise: An interesting aspect of our work is that without any algorithmic modifications, our approach is empirically stable in the presence of noise. We demonstrate this by comparing the mean squared error (MSE) between the ground-truth and the estimated values using $\mathbf{Y} = \mathscr{M}_\lambda(\mathbf{X}) + \mathbf{N}$, where $\mathbf{N}$ is additive white Gaussian noise. We perform the experiments with two different values of λ. The results using the ESPRIT algorithm are shown in Fig. 3(B). For low SNRs, we observe the "threshold effect" [38] in computational arrays. Because our algorithm is not yet designed to tackle noise, this effect arises because (16) does not hold for any value of K. However, the thresholding point is more favorable to larger λ: while for $\lambda = 0.2$ we note an improvement from SNR $= 20$ dB onward, for $\lambda = 0.5$ it lies closer to SNR

TABLE III

EXPERIMENTAL SETUP. UCA STANDS FOR UNIFORM CIRCULAR ARRAY. WHEN WORKING WITH LINEAR ARRAYS (ULA, NESTED, COPRIME), THE ELEVATION ANGLE IS ASSUMED TO BE ZERO. FOR SPARSE ARRAYS, THE NUMBER OF ELEMENTS IS SHOWN IN FORMAT L_1/L_2. FOR US ENABLED ULAS, THE NUMBER OF SENSORS, INCLUDING THE NUMBER OF ELEMENTS LOST WHEN PERFORMING THE SPATIAL DIFFERENCE OPERATION, K, IS SHOWN AS L (K). FOR US-ENABLED NESTED ARRAYS, WE USE THE NOTATION $L_1/(L_2(K+1))$ (K) TO DENOTE THE NUMBER OF ELEMENTS. TA STANDS FOR *TEMPORAL APPROACH* AND SA DENOTES *SPATIAL APPROACH*. THE TEMPORAL SAMPLING RATE IS SET TO BE $1/2\Omega e - 1/100 = 0.0485$ s

Exp.	M	**Array**	L	d/ν	**Angles** (degrees)	λ	T (s)	**Algorithm**
A	4	UCA	5	1/2	$\boldsymbol{\Theta} = 0.1 \times \{-2, -1, 1, 2\}$, $\boldsymbol{\Phi} = 5 + \{0, 0.1, 0.2, 0.3\}$.	1/5	0.0485	`TA-MUSIC` [6]
B	2	ULA	5	1/2	$\boldsymbol{\Theta} = \{3, 15\}$.	1/5, 1/2	0.0485	`TA-ESPRIT` [6]
C	14	Coprime	7/5	1/2	Uniformly spaced in $[-50, 50]$.	1/2	0.0485	`TA-MUSIC+Smooth` [36]
D	4	ULA	11 (3)	1/10	$\boldsymbol{\Theta} = \{-60, -30, 40, 70\}$.	1/5	1	`SA-MUSIC` [6]
E	1	ULA	11 (3)	1/10	Ranges from -90 to 90.	1/5	1	`SA-MUSIC` [6]
F	4	Nested	5/6 (1)	1/25	$\boldsymbol{\Theta} = \{-60, -30, 40, 70\}$.	3/4	1	`SA-Direct MUSIC` [24]

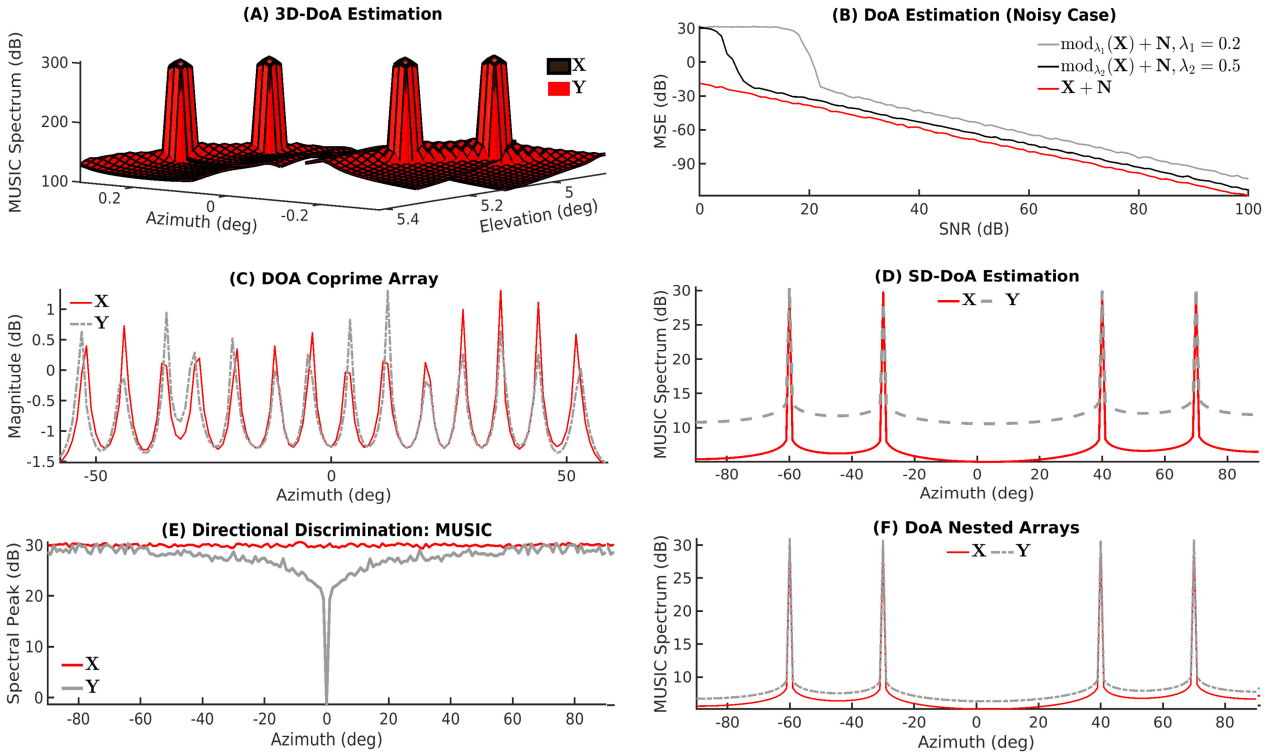

Fig. 3. Unlimited Sensing based Array Signal Processing (US–ASP). Temporal and spatial approaches are shown in subfigures (A, B, C) and (D, E, F), respectively. (A) Smoothed MUSIC pseudo-spectrum for DoA estimation in the 2D azimuth-elevation plane based on $\mathbf{X}$ and $\mathbf{Y}$. The difference between both estimators is almost negligible. (B) MSE vs SNR for DoA estimation using the ESPRIT algorithm with $\lambda = 0.2$ and $\lambda = 0.5$. For low SNRs, we observe the "threshold effect" [38] in computational arrays. (C) MUSIC pseudo-spectrum for the coprime scenario, using the covariance of $\mathbf{X}$ and $\mathbf{Y}$. A slight difference using US-ASP appears due to the distortion introduced by the difference operator in the power of the sources. (D) MUSIC pseudo-spectrum for an 11 element ULA with $K = 3$, obtained using $\mathbf{X}$ and $\mathbf{Y}$, which shows that both methods reach the same results, but the loss of receiving elements increases the noise-floor, and the effect of the difference operator reduces the amplitude of the peaks. (E) Maximum value of the MUSIC pseudo-spectrum as a function of the azimuth using $\mathbf{X}$ and $\mathbf{Y}$. The notching effect of the difference filter in the space domain is visible which almost vanishes for $|\theta_m| > 20$ deg. (F) MUSIC pseudo-spectrum using a US-enabled nested array of 11 elements; 5 in the dense part, 3 in the sparse part and $K = 1$ additional sensors for each element in the sparse subarray, hence using 3 sensors more, which will be used to obtain the differences, as shown in Fig. 2, for both $\mathbf{X}$ and $\mathbf{Y}$. The results are close to the pseudo-spectrum of the conventional approach, as the sparse section enhances the performance of the algorithm.

$= 10$ dB. For large SNRs, the remaining gap, due to the noise amplification via Δ^K operator, is also larger for smaller values of λ.

Experiment C: compatibility with coprime arrays: As mentioned, our approach can be integrated with coprime arrays. The upshot being, we can resolve more sources than physical sensors. To this end, the MUSIC pseudo-spectrum is shown in Fig. 3(C), where the difference between the results obtained by using $\mathbf{X}$ and $\mathbf{Y}$ stem from the weaker power of the signal part in $\mathbf{X}\mathbf{D}_N^K$, which play a key role in forming the difference coarray. As we can see, 14 sources can be resolved with only 12 sensors.

B. Spatial Domain US-ASP

Experiment D: uniform linear arrays: For the spatial domain approach, we compare the MUSIC pseudo-spectrum for modulo samples with that of $\mathbf{X}$ as benchmark. As shown in Fig. 3(D),

the MUSIC pseudo-spectrum of $\mathscr{M}_\lambda(\mathbf{Y}_K)$ matches that of $\mathbf{X}$ as the two subspaces contain the same information. That said, there is a scaling factor due to the effect of $\mathbf{E}$, which will be explored in the next experiment.

Experiment E: directional discrimination: As shown in (29), the application of difference filter scales the magnitude of the spectrum by $|\sin(\pi d/\nu \sin\theta_m)|^K$ creating a notch at $\theta = 0$. In Fig. 3(E), we show the MUSIC pseudo-spectrum as a function of the azimuth obtained from $\mathbf{X}$, and processing $\mathbf{Y}$, for $K = 3$. The gray curve shows a notch or blind spot at $\theta = 0$. However, this degradation is localized (for instance, spectral components above 20 degrees are left almost unaffected), and depends on K.

Experiment F: nested arrays: With nested arrays, we focus on mitigating the performance degradation arising from the inter-spacing of sensors. To this end, we consider a 11 sensor array forming a two level nested array, as in Fig. 2, using 5 elements in the dense part, 3 elements in the sparse part and with $K = 1$ additional sensors for each element in the sparse subarray, hence employing 3 sensors more, which will be used to obtain the differences, as shown in Fig. 2. We perform DoA estimation using the MUSIC algorithm [25]. The pseudo-spectrum is shown in Fig. 3(F). Note that spatial US-ASP (cf. Fig. 3(D)) entails Δ^K operation along the spatial dimension. This implies that we are left with $L - K$ samples. However, with sparse subarrays, this loss in samples can be compensated, resulting in superior performance.

V. Conclusion

In this paper, we introduced the idea of "computational arrays" that harness a synergistic tandem between hardware and algorithms. Our core novelty is the conceptualization of a sensor array which records modulo encoded measurements. This acquisition protocol circumvents the sensor saturation. By leveraging insights from the unlimited sensing framework, we develop mathematically guaranteed recovery algorithms for ASP tasks such as (a) DoA estimation, (b) beamforming and (c) signal reconstruction. In effect, our work allows for a high-dynamic-range sensing strategy and this is particularly well suited for scenarios where dynamic range is unknown or there are strong interferers in the sensing environment. An interesting feature of our work is that our method is backwards compatible with conventional ASP techniques.

Our paper raises a number of interesting questions around the topic and inspires new hardware architectures for ASP. We briefly cover a few directions that will bring our work closer to practice.

1) Stability and robustness in presence of noise remains an important topic of investigation. For instance, performance analysis in terms of the Cramér-Rao bounds for DoA estimation will provide guidelines for practical implementation. This will also clarify the role of λ when observing the "threshold effect" [38].
2) Our current approach relies on row-wise or column-wise operations. We believe that designing algorithms that jointly exploit spatio-temporal information may lead to robust solutions for ASP. The problem of detecting the number of sources or the compatibility with off-the-grid methods [10] also remain to be explored.
3) The temporal sampling condition $T \leq 1/2\Omega e$ directly follows from our earlier result in [3] where we also observed that (a) reconstruction is possible with sampling intervals larger than $1/2\Omega e$, and (b) modulo samples are injective just above the Nyquist rate. For the latter part, provided that a subset of unfolded samples is known a priori, a recovery method for functions on the real-line has been presented in [39]. Exploring tight sampling bounds and corresponding recovery algorithms, specially in the context of finite sample regime (cf. (21)) is a topic that will benefit practical implementations.
4) Low-complexity DoA estimation based on one-bit ADCs [18] has seen a surge of recent interest in the past decade. As shown in [40], one-bit approaches can also be integrated with unlimited sensing, resulting in new architectures for ASP in the future.

VI. Proofs

Proof of Lemma 2: To prove this result, it suffices to evaluate the differences with respect to the spatial dimension, $\Delta^K \xi_m^{2l}$. Since Δ^K is a convolution operator and the complex-exponentials $\{\xi_m\}_{m\in\lfloor\mathbb{M}\rceil}$ are eigen-functions of linear systems, we have, $\Delta^K \xi_m^{2l} = (\exp(\jmath\frac{2\pi d}{\nu}\sin\theta_m) - 1)^K \xi_m^{2l}$. The eigenvalues $(\xi_m^2 - 1)^K$ can be simplified by using the factorization,

$$m \in \lfloor\mathbb{M}\rceil,\ \left(\xi_m^2 - 1\right) = -\jmath 2\xi_m \mathsf{Im}\left(\xi_m\right) = [\mathbf{E}]_{m,m}, \quad (32)$$

where $\mathsf{Im}(\cdot)$ denotes the imaginary part. Consequently, for $l = 0, \ldots, L-K-1$,

$$[\left(\mathbf{D}_L^K\right)^\top \mathbf{A}_{\boldsymbol{\Theta}}]_{l,m} = [\mathbf{A}_{\boldsymbol{\Theta}}]_{l,m}(-\jmath 2\xi_m \mathsf{Im}\left(\xi_m\right))^K \Leftrightarrow \mathbf{I}_K \mathbf{A}_{\boldsymbol{\Theta}} \mathbf{E}^K,$$

which proves our result. ■

Proof of Lemma 3: According to (26), we have, $\mathbf{X}_K = \mathbf{I}_K \mathbf{A}_{\boldsymbol{\Theta}} \mathbf{E}^K \mathbf{S}$. We will upper-bound $\mathbf{X}_K$ element-wise and for this purpose, for any $K < L$, let us define, M–dimensional vectors for all l, n, $\mathbf{c}_{l,n} \in \mathbb{C}^M$ and $\mathbf{d}_K \in \mathbb{R}^M$ given by,

$$[\mathbf{c}_{l,n}]_m = -\jmath^K \xi_m^{2l+K} [\mathbf{S}]_{m,n} \quad \text{and} \quad [\mathbf{d}_K]_m = (2\mathsf{Im}(\xi_m))^K$$

respectively. This allows us to write,

$$[\mathbf{X}_K]_{l,n} = \sum_{m\in\lfloor\mathbb{M}\rceil} [\mathbf{A}_{\boldsymbol{\Theta}}]_{l,m} \left[\mathbf{E}^K\right]_{m,m} [\mathbf{S}]_{m,n} = \mathbf{d}_K^\top \mathbf{c}_{l,n}.$$

Now since, $[\mathsf{Re}(\mathbf{X}_K)]_{l,n} = \mathbf{d}_K^\top \mathsf{Re}(\mathbf{c}_{l,n})$, we can bound its maximum by writing,

$$\left|[\mathsf{Re}(\mathbf{X}_K)]_{l,n}\right| = \left|\mathbf{d}_K^\top \mathsf{Re}\left(\mathbf{c}_{l,n}\right)\right| \leq ||\mathbf{d}_K||_\infty \|\mathsf{Re}\left(\mathbf{c}_{l,n}\right)\|_{\ell_1(\mathbb{R})}$$

where the upper-bound is a result of applying Hölder's inequality. To bound $\|\mathsf{Re}(\mathbf{c}_{l,n})\|_{\ell_1(\mathbb{R})}$, we write $[\mathbf{S}]_{m,n} = |[\mathbf{S}]_{m,n}| e^{\jmath\angle[\mathbf{S}]_{m,n}}$ and notice that $\mathbf{c}_{l,n}$ is obtained by phase-modulation of $\mathbf{S}$ or,

$$K < L,\ [\mathbf{c}_{l,n}]_m = \underbrace{-\jmath^K e^{+\jmath\left(\frac{\pi d}{\nu}(2l+K)\sin\theta_m + \angle[\mathbf{S}]_{m,n}\right)}}_{\text{Phase Modulation}} \left|[\mathbf{S}]_{m,n}\right|$$

and it follows that $\|\mathsf{Re}(\mathbf{c}_{l,n})\|_{\ell_1(\mathbb{R})} \leq M|||\mathbf{S}|||$. Furthermore, $[\mathbf{d}_K]_m = 2^K \sin^K(\frac{\pi d}{\nu}\sin\theta_m)$ with $|\theta_m| \leq \pi/2$ and hence, the upper bound, $||\mathbf{d}_K||_\infty \leq 2^K \sin^K(\frac{\pi d}{\nu})$. By combining the respective upper-bounds on $||\mathbf{d}_K||_\infty$ and $\|\mathsf{Re}(\mathbf{c}_{l,n})\|_{\ell_1(\mathbb{R})}$, we obtain, an element-wise upper bound and hence,

$$|||\mathbf{X}_K|||_{\mathsf{Re}} \leq M\left(2\sin\left(\tfrac{\pi d}{\nu}\right)\right)^K |||\mathbf{S}|||. \tag{33}$$

The same applies to $|[\mathsf{Im}(\mathbf{X}_K)]_{l,n}|$ and the result follows. ■

Proof of Theorem 3: Let $\Pi_K(\mathbf{X}) = (\mathbf{D}_L^K)^\top \Pi(\mathbf{X})$. According to Proposition 1, (13) holds and, $\mathbf{Y}_K = \mathbf{X}_K + \Pi_K(\mathbf{X})$. Since $\mathscr{M}_\lambda(a+b) = \mathscr{M}_\lambda(\mathscr{M}_\lambda(a) + \mathscr{M}_\lambda(b)), \forall a, b \in \mathbb{C}$, we have, $\mathscr{M}_\lambda(\mathbf{Y}_K) = \mathscr{M}_\lambda(\mathscr{M}_\lambda(\mathbf{X}_K) + \mathscr{M}_\lambda(\Pi_K(\mathbf{X})))$. Thus, setting $d < \nu/6$ and choosing K using (28) ensures that $|||\mathbf{X}_K|||_{\mathsf{Re}} \leqslant \lambda$ and $|||\mathbf{X}_K|||_{\mathsf{Im}} \leqslant \lambda$ and hence, $\mathscr{M}_\lambda(\mathbf{X}_K) = \mathbf{X}_K$. Moreover, the integer-valued coefficients in $\mathbf{D}_L^N$ preserve the ring structure or $\Pi_K(\mathbf{X}) \in 2\lambda\mathbb{Z}$ and therefore, $\mathscr{M}_\lambda(\Pi_K(\mathbf{X})) = 0$. This yields $\mathscr{M}_\lambda(\mathbf{Y}_K) = \mathbf{X}_K$, which proves the result. ■

REFERENCES

[1] S. Fernandez-Menduina, F. Krahmer, G. Leus, and A. Bhandari, "DoA estimation via unlimited sensing," in *Proc. 28th Eur. Signal Process. Conf.*, 2021, pp. 1866–1870. doi: 10.23919/Eusipco47968.2020.9287595

[2] A. Bhandari, F. Krahmer, and R. Raskar, "On unlimited sampling," in *Proc. 12th Int. Conf. Sampling Theory Appl.*, Jul. 2017, pp. 31–35.

[3] A. Bhandari, F. Krahmer, and R. Raskar, "On unlimited sampling and reconstruction," *IEEE Trans. Signal Process.*, vol. 69, pp. 3827–3839, Dec. 2020, doi: 10.1109/TSP.2020.3041955.

[4] A. Bhandari, F. Krahmer, and R. Raskar, "Methods and apparatus for modulo sampling and recovery," U.S. Pat. US20 190 103 876A1, 2020.

[5] T. E. Tuncer and B. Friedlander, Eds., *Classical and Modern Direction-of-Arrival Estimation*. Academic Press, 2009.

[6] H. Krim and M. Viberg, "Two decades of array signal processing research: The parametric approach," *IEEE Signal Process. Mag.*, vol. 13, no. 4, pp. 67–94, Jul. 1996.

[7] P. P. Vaidyanathan and P. Pal, "Sparse sensing with co-prime samplers and arrays," *IEEE Trans. Signal Process.*, vol. 59, no. 2, pp. 573–586, Feb. 2011.

[8] D. Romero, D. D. Ariananda, Z. Tian, and G. Leus, "Compressive covariance sensing: Structure-based compressive sensing beyond sparsity," *IEEE Signal Process. Mag.*, vol. 33, no. 1, pp. 78–93, Jan. 2016.

[9] P. Pal and P. P. Vaidyanathan, "Nested arrays: A novel approach to array processing with enhanced degrees of freedom," *IEEE Trans. Signal Process.*, vol. 58, no. 8, pp. 4167–4181, Aug. 2010.

[10] G. Tang, B. N. Bhaskar, P. Shah, and B. Recht, "Compressed sensing off the grid," *IEEE Trans. Inf. Theory*, vol. 59, no. 11, pp. 7465–7490, Nov. 2013.

[11] D. Malioutov, M. Cetin, and A. S. Willsky, "A sparse signal reconstruction perspective for source localization with sensor arrays," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 3010–3022, Aug. 2005.

[12] J. Lai, Y. Xu, X. Zhang, and Z. Tang, "Structural analysis of lunar subsurface with chang E-3 lunar penetrating radar," *Planet. Space Sci.*, vol. 120, pp. 96–102, Jan. 2016.

[13] R. H. Dietz, D. E. Rhoades, and L. J. Davidson, "Apollo experience report-lunar module communications system," *Nat. Aeronaut. Space Admin.*, Tech. Rep. TN D-6974, Sep. 1972.

[14] R. Cao, S. Earp, S. A. L. de Ridder, A. Curtis, and E. Galetti, "Near-real-time near-surface 3D seismic velocity and uncertainty models by wavefield gradiometry and neural network inversion of ambient seismic noise," *Geophysics*, vol. 85, no. 1, pp. KS 13–KS27, Nov. 2019.

[15] B. Wang, Y. Zhao, and J. Liu, "Mixed-order MUSIC algorithm for localization of far-field and near-field sources," *IEEE Signal Process. Lett.*, vol. 20, no. 4, pp. 311–314, Apr. 2013.

[16] B. Brannon, "Wideband radios need wide dynamic range converters," *Analog Dialogue*, vol. 29, no. 2, pp. 11–12, 1995.

[17] W. Oberhammer and B. Li, "Dynamic range extension of wideband receiver," U.S. Pat. US6 333 707B1, 2001.

[18] O. Bar-Shalom and A. Weiss, "DOA estimation using one-bit quantized measurements," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 38, no. 3, pp. 868–884, Jul. 2002.

[19] J. Mo and R. W. Heath, "Capacity analysis of one-bit quantized MIMO systems with transmitter channel state information," *IEEE Trans. Signal Process.*, vol. 63, no. 20, pp. 5498–5512, Oct. 2015.

[20] P. J. Soh *et al.*, "A smart wearable textile array system for biomedical telemetry applications," *IEEE Trans. Microw. Theory Techn.*, vol. 61, no. 5, pp. 2253–2261, May 2013.

[21] M. Allam and J. Greenleaf, "Isomorphism between pulsed-wave doppler ultrasound and direction-of-arrival estimation. I. Basic principles," *IEEE Trans. Ultrason., Ferroelect., Freq. Control*, vol. 43, no. 5, pp. 911–922, Sep. 1996.

[22] J. Rhee and Y. Joo, "Wide dynamic range CMOS image sensor with pixel level ADC," *Electron. Lett.*, vol. 39, no. 4, pp. 360–361, 2003.

[23] S. Hirsch, M. Strobel, W. Klingler, J. D. S. Spüntrup, Z. Yu, and J. N. Burghartz, "Realization and opto-electronic characterization of linear self-reset pixel cells for a high dynamic CMOS image sensor," *Adv. Radio Sci.*, vol. 17, pp. 239–247, Sep. 2019.

[24] A. Bhandari, F. Krahmer, and T. Poskitt, "Unlimited sampling from theory to practice: Fourier-Prony recovery and prototype ADC," *IEEE Trans. Sig. Proc.*, vol. 70, pp. 1131–1141, Sep. 2021, doi: 10.1109/TSP.2021.3113497.

[25] P. P. Vaidyanathan and P. Pal, "Direct-MUSIC on sparse arrays," in *Proc. Int. Conf. Signal Process. Commun.*, 2012, pp. 1–5.

[26] P. Pal and P. P. Vaidyanathan, "Coprime sampling and the MUSIC algorithm," in *Proc. Digit. Signal Process. Signal Process. Educ. Meeting*, 2011, pp. 289–294.

[27] A. Bhandari, F. Krahmer, and R. Raskar, "Unlimited sampling of sparse sinusoidal mixtures," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2018, pp. 336–340.

[28] A. Bhandari, F. Krahmer, and R. Raskar, "Unlimited sampling of sparse signals," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal. Process.*, Apr. 2018, pp. 4569–4573.

[29] A. Bhandari and F. Krahmer, "On identifiability in unlimited sampling," in *Proc. Int. Conf. Sampling Theory Appl.*, Jul. 2019, pp. 1–4.

[30] A. K. Robinson, P. Venkatesh, M. J. Boring, M. J. Tarr, P. Grover, and M. Behrmann, "Very high density EEG elucidates spatiotemporal aspects of early visual processing," *Sci. Rep.*, vol. 7, no. 1, pp. 1–11, Nov. 2017.

[31] Y. Zhu *et al.*, "Analyzing high-density ECG signals using ICA," *IEEE Trans. Biomed. Eng.*, vol. 55, no. 11, pp. 2528–2537, Nov. 2008.

[32] P. Goldstein and R. J. Archuleta, "Deterministic frequency-wavenumber methods and direct measurements of rupture propagation during earthquakes using a dense array: Theory and methods," *J. Geophysical Res.: Solid Earth*, vol. 96, no. B4, pp. 6173–6185, Apr. 1991.

[33] D. A. Quiros *et al.*, "Reflection imaging with earthquake sources and dense arrays," *J. Geophysical Res.: Solid Earth*, vol. 122, no. 4, pp. 3076–3098, Apr. 2017.

[34] C. W. Johnson, D. Kilb, A. Baltay, and F. Vernon, "Peak ground velocity spatial variability revealed by dense seismic array in southern California," *J. Geophysical Res.: Solid Earth*, vol. 125, no. 6, Jun. 2020, Art. no. e2019JB019157.

[35] C.-P. Yeang, G. W. Wornell, L. Zheng, and J. Krieger, "Dense transmit and receive phased arrays," in *Proc. IEEE Int. Symp. Phased Array Syst. Technol.*, Oct. 2010, pp. 934–939.

[36] C. Byrne, "Iterative oblique projection onto convex sets and the split feasibility problem," *Inverse Problems*, vol. 18, no. 2, pp. 441–453, 2002.

[37] C. Liu and P. P. Vaidyanathan, "Remarks on the spatial smoothing step in coarray MUSIC," *IEEE Signal Process. Lett.*, vol. 22, no. 9, pp. 1438–1442, Sep. 2015.

[38] D. Rife and R. Boorstyn, "Single tone parameter estimation from discrete-time observations," *IEEE Trans. Inf. Theory*, vol. 20, no. 5, pp. 591–598, Sep. 1974.

[39] E. Romanov and O. Ordentlich, "Above the nyquist rate, modulo folding does not hurt," *IEEE Signal Process. Lett.*, vol. 26, no. 8, pp. 1167–1171, Aug. 2019.

[40] O. Graf, A. Bhandari, and F. Krahmer, "One-bit unlimited sampling," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, May 2019, pp. 5102–5106.

Samuel Fernández Menduiña received the B.Sc. degree in telecommunications engineering from the University of Vigo, Vigo, Spain, in 2019 and the M.Sc. degree in communications and signal processing from Imperial College London, London, U.K., in 2020. Since October 2020, he has been a Research Engineer with AtlanTTic Research Center. His research interests include computational sensing, multimedia forensics, and statistical signal processing. He was awarded the Extraordinary B.Sc. Prize in 2020 by the University of Vigo and the Xunta de Galicia. For the M.Sc. work and thesis, he was awarded the Outstanding Achievement Award in Communications and Signal Processing by Imperial College London.

Felix Krahmer received the B.Sc. degree in mathematics from International University Bremen, Bremen, Germany, and the M.Sc. and the Ph.D. degrees in mathematics from New York University, New York City, under the supervision of Percy Deift and Sinan Güntürk. From 2009 to 2012, he was a HCM Postdoc with the Group of Holger Rauhut, University of Bonn, Bonn, Germany. In 2012, he moved to the University of Göttingen, Göttingen, Germany as an Assistant Professor for mathematical data analysis, where he was awarded an Emmy Noether Junior Research Group. In 2015, he joined the Department of Mathematics, Technical University of Munich, Munich, Germany, as a tenure-track Assistant Professor for optimization and data analysis, he was tenured and promoted to the level of Associate Professor in 2021. His research interests include mathematical signal and data processing, in particular the theory of computational sensing, applied harmonic analysis, and random matrix theory.

Geert Leus received the M.Sc. and Ph.D. degree in electrical engineering from the KU Leuven, Leuven, Belgium, in June 1996 and May 2000, respectively. He is currently a Full Professor with the Faculty of Electrical Engineering, Mathematics, and Computer Science, Delft University of Technology, Delft, The Netherlands. He was the recipient of the 2021 EURASIP Individual Technical Achievement Award, a 2005 IEEE Signal Processing Society Best Paper Award, and a 2002 IEEE Signal Processing Society Young Author Best Paper Award. He is a Fellow of the EURASIP. He was a Member-at-Large of the Board of Governors of the IEEE Signal Processing Society, the Chair of the IEEE Signal Processing for Communications and Networking Technical Committee, and the Editor in Chief of the EURASIP JOURNAL ON ADVANCES IN SIGNAL PROCESSING. He is currently the Chair of the EURASIP Technical Area Committee on Signal Processing for Multisensor Systems and the Editor-in-Chief of the *EURASIP Signal Processing*.

Ayush Bhandari received the Ph.D. degree from the Massachusetts Institute of Technology (MIT), Cambridge, MA, USA, in 2018, for his work on computational sensing and imaging which is being shaped as a forthcoming, coauthored book in MIT Press (2022). He is currently a Faculty Member with the Department of Electrical and Electronic Engineering, Imperial College London, London, U.K. He has held research positions with INRIA, Rennes, France, Nanyang Technological University, Singapore, the Chinese University of Hong Kong, Hong Kong, and Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland among other institutes. In 2019, he was appointed the August–Wilhelm Scheer Visiting Professor with the Department of Mathematics, Technical University of Munich, Munich, Germany. His research interests include mathematical signal processing, inverse problems, and computational sensing and imaging. He was a tutorial speaker at various venues, including the ACM Siggraph in 2014 and 2015 and the IEEE ICCV in 2015. Some aspects of his work have led to new sensing and imaging modalities which have been widely covered in press and media including *BBC news*. Applied aspects of his research have led to more than ten US patents. His scientific contributions have led to numerous prizes, most recently, the Best Paper Award at IEEE ICCP 2020 (International Conference on Computational Photography) and the Best Student Paper Award (senior coauthor) at IEEE ICASSP 2019 (International Conference on Acoustics, Speech and Signal Processing). His doctoral work was awarded the 2020 IEEE Best Ph.D. Dissertation Award from the Signal Processing Society. In the same year, he was awarded the UKRI Future Leaders Fellowship. In 2021, he was the recipient of the President's Medal for Outstanding Early Career Researcher at Imperial College London.