Line Spectral Estimation via Unlimited Sampling

QI ZHANG
Singapore University of Technology and Design, Singapore

JIANG ZHU, Senior Member, IEEE
FENGZHONG QU, Senior Member, IEEE
Zhejiang University, Zhoushan, China

DE WEN SOH
Singapore University of Technology and Design, Singapore

Frequency-modulated continuous wave (FMCW) radar has been widely applied in automotive anticollision systems, automatic cruise control, and indoor monitoring. However, conventional analog-to-digital converters (ADCs) can suffer from significant information loss when strong and weak targets coexist in ranging applications. To address this issue, the unlimited sampling (US) strategy was proposed, which applies a modulo operator prior to sampling. In this article, we investigate the range-estimation problem using FMCW radar in the context of US, which can be formulated as a 1-D line spectral estimation (LSE) via US. By exploiting the oversampling property and proving that the leakage onto a certain frequency band can be controlled, we establish an integer optimization problem in the Fourier and first-order difference domain. We then propose a dynamic programming (DP)-based algorithm followed by the orthogonal matching pursuit (OMP) method to solve it. In addition, a two-stage US LSE (USLSE) is proposed, where the line spectral signal is first recovered and then the parameters are estimated by applying a state-of-the-art LSE algorithm. Substantial numerical simulations and real experiments demonstrate that the proposed algorithm, USLSE, outperforms existing algorithms.

I. INTRODUCTION

Frequency-modulated continuous wave (FMCW) radar has been widely applied in automotive anticollision systems, automatic cruise control, and indoor monitoring due to its high ranging precision and resolution, outstanding interference resistance, and simple structure and compact size [1], [2], [3], [4], [5]. In ranging applications, the presence of the inverse-square law, where the intensity of a target is inversely proportional to the square of the distance from the target to the source, and variations in the reflective properties of different materials often result in scenarios where strong and weak targets coexist [6], [7]. Such scenarios include detecting humans near trucks, humans across a broad range, and indoor monitoring amid significant clutter [8], [9].

For the conventional analog-to-digital converter (ADC) used in FMCW radar systems, the information loss is inevitable in the scenario where strong and weak targets coexist due to the following dilemma [10]: On one hand, due to the existence of strong targets, the amplitude of the signal may exceed the dynamic range of the ADC, which will lead to saturation or clipping, resulting in the loss of information from the most significant bits (MSBs). On the other hand, if we adjust the dynamic range of the ADC to avoid clipping, the weak signal may be submerged in quantization noise, causing the loss of information from the least significant bits (LSBs). To address this problem, self-reset ADCs are implemented in which a modulo operation is applied before sampling [11], [12], [13], [14]. In these architectures, side information like the amount of folding for each sample or the folding sign is stored for perfect reconstruction, which requires complex electronic circuitry.

Recently, signal recovery from modulo samples only has been studied in depth. The theory that there is a one-to-one mapping between modulo samples and unfolded measurements has been established when the sampling rate is slightly above the critical Nyquist rate for bandlimited and finite energy signals [15], [16]. In addition, if the number of samples per interval is equal to or exceeds $2L + 1$ and $2L + 1$ is a prime number, the $L$-order periodic bandlimited signal can be uniquely identifiable up to a constant factor from the modulo samples [17]. Furthermore, from the perspective of compressed sensing, researchers have derived the minimum number of modulo measurements required to achieve the recovery guarantee for sparse vectors [18]. Additionally, the Cramér–Rao bounds (CRBs) are established, taking into account the known folding-count, for both quantized and unquantized scenarios, serving as benchmarks for performance evaluation [19].

Algorithms for reconstruction can mainly be divided into two categories, one is based on the time domain processing, and the other is based on the frequency domain processing. In the time domain, Bhandari et al. [20] proposed an unlimited sampling algorithm (USAlg) for bandlimited signals, which is based on the higher-order differences operator and provided a recovery guarantee when the oversampling rate exceeds $2\pi e$ times the Nyquist...
rate. USAAlg has been extended to address different signal models and problems, such as the mixture of sinusoids, finite-rate-of-innovation signals, multidimensional signals, DOA estimation, computed tomography, graph signals [21], [22], [23], [24], [25], [26], [27], [28], etc. However, due to higher-order differences, USAAlg is sensitive to noise, especially when the noise is added to the original signal. The authors in [29] and [30] proposed another method in the time domain named integer-forcing decoding to unwrap the random Gaussian distributed vectors, where the covariance matrix is known. The blind version is also proposed where an empirical method is used to estimate the covariance matrix [31]. In addition, the linear prediction (LP) approach is introduced for stationary stochastic Gaussian processes in both the temporal and spatiotemporal cases and approaches the limits given by rate-distortion theory [30]. The LP approach is also applied for bandlimited signals and the recovery guarantee is provided when the oversampling factor is greater than 1 [16]. The LP method requires that the initial few samples of the signal are always within the dynamic range, and the autocorrelation function is known. Furthermore, a blind version of the LP method is proposed, which incorporates both an automatic adjustment at the modulo level and a fully adaptable modulo unwrapping mechanism [32], [33]. For the recovery approach based on the Fourier domain, Bhandari et al. [34] proposed an algorithm for recovering the periodic bandlimited signals by solving a spectral estimation problem with Prony’s method when the number of folds is known, and this method can be used to handle the nonideal foldings in practical hardware. Another kind of method in the Fourier domain uses the projected gradient descent method to estimate the residual [35], [36]. Besides, to recover the clean original signal from noisy modulo samples, two-stage algorithms are proposed, which perform denoising on modulo samples before recovering [37], [38].

Hardware prototype for US is also designed, which is first introduced in [34] followed by the work in [39]. To deal with hysteresis and folding transients in practical hardware, a thresholding approach with a recovery guarantee is proposed which can also be applied to the case when the nonidealities are not present, and a low sampling rate version is also provided [40]. In addition, thresholding with general filters and average sampling are utilized to deal with nonidealities [41], [42]. Considering the practical hardware, three signal folding architectures are summarized in [43], and the reconstruction algorithms are proposed for bandpass signals.

In this article, we study the FMCW radar applied in range-estimation problem where one transmit and one receive antenna pair is used in the context of US, as shown in Fig. 1. This problem can be formulated as 1-D line spectral estimation (LSE) via US. In [7], the unlimited sampling framework (USF) is employed to address the problem, utilizing US ADCs at the hardware level and applying USAAlg at the algorithmic level. However, it is worth noting that the original signal in [7] operates at a low noise level, indicating a relatively weak noise intensity compared to the signal. As previously mentioned, USAAlg is particularly sensitive when noise is introduced to the original signal. In practice, the original signal is often affected by noise in many scenarios. Additionally, unlike situations in numerous papers where noise is added after modulo sampling, a noisy original signal not only influences the modulo values of samples but also affects the folding times. Therefore, studying robust reconstruction algorithms in the presence of noise before modulo sampling is of vital importance and is more challenging.

We study the LSE via US in the presence of noise before modulo sampling in this article. The main contributions are threefold: First, we establish an optimization problem to recover the folding instants and folding times. We show that the proposed objective function is meaningful even when the frequencies do not lie on the discrete Fourier transform (DFT) grid exactly, as the leakage onto a certain frequency band can be controlled. Second, an iterative algorithm named USLSE is proposed. Since the original optimization problem with quadratic objective function and integer constraints is NP-hard, we innovatively solve this problem by approximating the matrix in the quadratic form as a banded matrix. Such an approximation is reasonable as we show that the banded matrix consists of most of the energy of the original matrix, and the quadratic form with a banded matrix with integer constraints can be solved globally via dynamic programming (DP). In addition, provided that the estimates generated by the DP are close to the global solution, the residue calculated by subtracting the estimates from the global solution is sparse, and orthogonal matching pursuit (OMP) is adopted to further refine the estimates. We could run multiple iterations of DP followed by OMP to refine the estimates. Finally, substantial numerical simulations and real experiments are conducted to demonstrate the

![Fig. 1. Range estimation using FMCW radar in the framework of US.](image)
robustness of the proposed USLSE compared to the state-of-the-art methods. Numerical results also demonstrate that the proposed approach is effective for recovering bandlimited signals.

A. Notation

We use \( \mathbb{Z} \), \( \mathbb{R} \), and \( \mathbb{C} \) to denote sets of integers, real numbers, and complex numbers, respectively. Let \( \mathbb{Z}^* \) be the set of complex numbers with both the real part and the imaginary part being integers. For a complex vector \( x \in \mathbb{C}^N \), \( x_i \) or \( [x] \) denotes the \( i \)th element of \( x \) for any \( i = 1, \ldots, N \). For a complex matrix \( A \in \mathbb{C}^{N \times M} \), \( A_{ij} \) denotes the \( (i, j) \)th element of \( A \) for any \( i = 1, \ldots, N \) and \( j = 1, \ldots, M \). Let \( S_N \) and \( S_M \) be subsets of \( \{1, \ldots, N\} \) and \( \{1, \ldots, M\} \), respectively. \( |S_N| \) denotes the cardinality of \( S_N \). \( A_{S_N, S_M} \) denotes the submatrix by choosing the rows indexed by \( S_N \) and the columns indexed by \( S_M \) of \( A \). For simplicity, \( A_{S_N} \) is used to denote the submatrix by choosing the elements of \( x \) indexed by \( S_N \). Similarly, \( x_{S_N} \) denotes the subvector by choosing the elements of \( x \) indexed by \( S_N \). Let \( \Re\{A\} \) and \( \Im\{A\} \) denote the real and imaginary parts of \( A \), respectively. Let \( \text{diag}(x) \) return a diagonal matrix with the diagonal being \( x \). Let \( (\cdot)^\ast, (\cdot)^T, \) and \( (\cdot)^H \) be the conjugate, transpose, and Hermitian transpose operators, respectively. Let \( \|x\|_1, \|x\|_2 \), and \( \|x\|_\infty \) be the Manhattan norm, Euclidean norm, and maximum norm, respectively. Let \( |A| \) and \( |A| \) be the elementwise floor and round operators for both real and imaginary parts of \( A \), respectively.

II. PROBLEM SETUP

In FMCW radar, the transmitted chirp \( s(t) \) whose frequency increases linearly with time is

\[
s(t) = \sqrt{P} \cos \left( 2\pi (f_0 + \frac{B}{T_c} t) \times t + \phi \right) \tag{1}
\]

where \( f_0 \) is the carrier frequency, \( B \) is the bandwidth, and \( T_c \) is the duration of the chirp. When the transmitted signal encounters a target and reflects back to the radar receiver, the received signal, after dechirping, can be expressed as [44]

\[
g(t) = \sum_{k=1}^{K} a_k e^{j(2\pi \frac{f_0}{c} d_k + \phi_k)} + w(t) \tag{2}
\]

where \( K \) is the number of targets, \( d_k \) is the distance from the \( k \)th target to the radar, \( c \) is the speed of light, and \( w(t) \) is the noise. Let \( d_k \in (0, d_{\text{max}}) \), \( k = 1, \ldots, K \) where \( d_{\text{max}} \) is the maximum distance of targets. Let \( f_k \triangleq \frac{2d_k}{T_c} \) and \( c_k \triangleq \frac{\lambda c}{T_k} \), the received signal (2) can be simplified as

\[
g(t) = \sum_{k=1}^{K} c_k e^{j2\pi f_k t} + w(t) \tag{3}
\]

which is the continuous LSE problem with \( K \) frequencies where \( f_k \in (0, f_{\text{max}} = \frac{2d_{\text{max}}}{T_c}) \) and \( c_k \in \mathbb{C} \) denote the frequency and amplitude of the \( k \)th signal, respectively. Obviously, \( f_{\text{max}} \) is the Nyquist frequency, i.e., \( f_{\text{Nyq}} = f_{\text{max}} \). Let \( N \) denote the oversampling factor with respect to the Nyquist frequency \( f_{\text{Nyq}} \). Thus, the sampling frequency is \( f_s = \gamma f_{\text{Nyq}} = \gamma f_{\text{max}} \). The sampling interval is \( T_s = 1/f_s = 1/(\gamma f_{\text{max}}) \), and \( g[n] \triangleq g(nT_s) \) are the original samples, which can be written as

\[
g[n] = x[n] + w[n], \quad n = 0, \ldots, N - 1 \tag{4}
\]

where \( \omega_k = 2\pi f_k/f_s \) is the normalized angular frequency of \( k \)th signal, \( w[n] \) are the noise sampled from \( w(t) \), and \( N \) is the number of samples. In addition, for \( f_k \in \{0, f_{\text{max}}\} \) and \( f_s = \gamma f_{\text{max}} \), we have

\[
\omega_k = 2\pi f_k/f_s \in \left( 0, \frac{2\pi}{\gamma} \right), \quad k = 1, \ldots, K. \tag{6}
\]

Uniform modulo sampling of \( g(t) \) where a modulo operator is applied before sampling yields

\[
y[n] \triangleq \mathcal{M}_\lambda \{ \Re\{g(n)\} \} + j\mathcal{M}_\lambda \{ \Im\{g(n)\} \} \tag{7}
\]

where \( \mathcal{M}_\lambda(t) \) is the centered modulo mapping defined as

\[
\mathcal{M}_\lambda(t) \triangleq 2\lambda \left( \left\lfloor \frac{t}{2\lambda} + \frac{1}{2} \right\rfloor - \frac{1}{2} \right) \tag{8}
\]

where \( \lfloor t \rfloor \triangleq t - \lfloor t \rfloor \) denotes the fractional part of \( t \), \( \lambda \) denotes the threshold of ADC. Note that the original samples can be decomposed as the sum of modulo samples and a simple function [20], i.e.,

\[
g[n] = y[n] + 2\lambda \epsilon[n], \quad n = 0, 1, \ldots, N - 1 \tag{9}
\]

where \( \epsilon[n] \in \mathbb{Z}^* \) and \( \mathbb{Z}^* \triangleq \mathbb{Z} + j\mathbb{Z} \) is the set of complex numbers with both real and imaginary parts being integers. Furthermore, the matrix representation of (4) and (9) can be formulated as

\[
g = x + w = A(\omega)c + w \tag{10a}
\]

\[
g = y + 2\lambda \epsilon \tag{10b}
\]

where \( A(\omega) \triangleq \begin{bmatrix} a(\omega_1), a(\omega_2), \ldots, a(\omega_K) \end{bmatrix} \in \mathbb{C}^{N \times K} \) and \( a(\omega_k) \triangleq \begin{bmatrix} e^{j\omega_k}, \ldots, e^{j(N-1)\omega_k} \end{bmatrix}^T \) is the array manifold steering vector, \( g, x, w, y, \) and \( \epsilon \) are vector representations of the corresponding sequences. For example, \( g \triangleq [g[0], g[1], \ldots, g[N-1]]^T \). \( \omega \triangleq [\omega_1, \ldots, \omega_K]^T \) and \( c \triangleq [c_1, \ldots, c_K]^T \). It is worth noting that \( g_n = g[n-1], n = 1, \ldots, N \). In this article, we aim to recover \( \{c, \omega\} \) from modulo samples \( y \).

To estimate parameters \( \{c, \omega\} \), least squares estimation is equivalent to solving the following optimization problem:

\[
\min_{c, \omega, \epsilon} \|y + 2\lambda \epsilon - A(\omega)c\|_2^2 \tag{11a}
\]

\[
s.t. \epsilon \in \mathbb{Z}^{N \times N}. \tag{11b}
\]

\[\dagger\]If \( \{w[n]\}_{n=0}^{N-1} \) are independent and identically distributed (i.i.d.) and follow a Gaussian distribution, then the least-squares estimation is equivalent to maximum likelihood estimation by viewing \( \epsilon \) as deterministic unknown parameters.
However, (11) is NP-hard generally as $\epsilon$ are constrained on the lattice $Z^N$. In detail, (11) can be solved in the following steps. Without loss of generality, suppose that $|\epsilon|_\infty \leq \max(\|\Re(\epsilon)\|_\infty, \|\Im(\epsilon)\|_\infty) = V_{\max}$ where $V_{\max} \in Z$. We use enumeration to solve the problem. The number of possible $\epsilon$ is $(2V_{\max} + 1)^2$. For each $\epsilon$, (11) degenerates into an LSE, which can be approximately solved and we can obtain the objective value. Finally, we choose the estimates of the parameters corresponding to the best (minimum) objective value. Consequently, the computation complexity is $(2V_{\max} + 1)^2N$ times the complexity of the LSE algorithm. Note that the computation complexity is exponential in the number of measurements $N$, which makes it computationally intractable.

Another heuristic approach is to use a linear model to represent the nonlinear LSE model, i.e., the original $A(\omega)\epsilon$ can be represented as

$$A(\omega)\epsilon = A_{\text{dic}}\tilde{\epsilon}$$  

where $A_{\text{dic}} \in C^{N \times N_{\text{dic}}}$ is the dictionary matrix with the $(n_1, n_2)$th element being $e^{\frac{2\pi j n_1 n_2}{N_{\text{dic}}} }, n_1 = 1, \ldots, N, n_2 = 1, \ldots, N_{\text{dic}}$ and $N_{\text{dic}} \geq N$. When $N_{\text{dic}} > K$, $\tilde{\epsilon}$ is approximately sparse. Thus, (11) can be relaxed as

$$\min_{\epsilon, \tilde{\epsilon}} \|y + 2\lambda \epsilon - A_{\text{dic}}\tilde{\epsilon}\|_2^2 + \mu \|\tilde{\epsilon}\|_1$$  

s.t. $\epsilon \in Z^N$  

where $\mu$ is a regularization parameter to control the sparsity of the signal. Problem (13) is still NP-hard as there exist integer constraints that $\epsilon \in Z^N$. We solved the relaxed version of (13) by relaxing the nonconvex constraints (13b) as the linear inequality constraints, but the solution is poor. In particular, if we ignore the integer constraints first, the solution to (13) is trivial and far from the true values, i.e., $\tilde{\epsilon} = -y/(2\lambda)$ and $\tilde{\epsilon} = 0$. Besides, we kept the nonconvex constraints (13b) and used the proximal approach to solve (13) with a random initial point. However, the local solution is still poor. Therefore, it is challenging to solve (13) using polynomial time algorithms, and more structures should be utilized. In the ensuing section, based on oversampling, the signal part is eliminated, and an optimization problem involved only with respect to the first-order difference of the simple function is established.

III. LSE THROUGH FREQUENCY DOMAIN PROCESSING

Intuitively, oversampling benefits in two facets: First, oversampling makes the simple function $2\lambda \epsilon$ similar at the nearby instants. This causes the value of the first-order finite difference of the simple function to be near zero, which reduces the state space of the first-order finite difference. Second, oversampling widens the spectrum between the maximum signal frequency and the Nyquist frequency. Since the leakage due to the line spectral signal into the spectrum between the maximum signal frequency and the sampling frequency is small, which will be discussed ahead, the spectrum between the maximum signal frequency and the sampling frequency is only approximately related to the simple function. This leads us to expect that the recovery process becomes simpler. In the following, we exploit the earlier benefits to formulate the LSE.

Let $\epsilon$ denote the first-order finite difference of $\epsilon$, where the $n$th element is defined as

$$\epsilon_n = \epsilon_{n+1} - \epsilon_n, n = 1, \ldots, N - 1$$  

or in matrix form $\epsilon = J\epsilon$ where $J = I_{2N} - I_{2N-1}$. and $I \in \mathbb{R}^{N \times N}$ is the identity matrix. Note that $\epsilon$ is still on the lattice $Z^N$, thus $\Re(\epsilon)$ and $\Im(\epsilon)$ will be in the set $\{-V, \ldots, 0, \ldots, V\}$ where $V = \max(\|\Re(\epsilon)\|_\infty, \|\Im(\epsilon)\|_\infty)$ is a positive integer. As the oversampling factor $\gamma$ increases, the dynamic range of both the real and imaginary parts of $\epsilon$, i.e., $V$, will decrease, and Proposition 1 shows the result of the noiseless case. Applying the first-order finite difference to (10), we have

$$y + 2\lambda \epsilon = x + w = JA(\omega)\epsilon + w$$  

where $y \triangleq Jy, x \triangleq Jx$, and $w \triangleq Jw$.

Proposition 1 In the noiseless case, given the oversampling factor $\gamma$, both $\Re(\epsilon)$ and $\Im(\epsilon)$, $n = 1, 2, \ldots, N - 1$, will be in the set $\{-V, \ldots, 0, \ldots, V\}$ with a bounded value of $V$ such that

$$V \leq \left\lfloor \frac{K\pi c_{\max}}{\lambda \gamma} + 1 \right\rfloor$$  

where $c_{\max} = \max|c_k|$.

Proof The proof is postponed to Appendix A.

Performing DFT on (15) yields

$$\bar{y} + 2\lambda F\epsilon = \bar{x} + \bar{w} = FJ(\omega)\epsilon + \bar{w}$$  

where $\bar{y} \triangleq Fy, \bar{x} \triangleq FX, \bar{w} \triangleq FW$, and $F \in C^{(N-1) \times (N-1)}$ is the 'DFT' matrix with the $(n_1, n_2)$th element being $\frac{1}{\sqrt{N-1}} e^{-j\frac{2\pi (n_1-1)(n_2-1)}{N-1}}$. Given that $\omega_k \leq 2\pi/\gamma$ for $k = 1, \ldots, K$ as stated in (6), if $\omega_k$ lies on the grid $\{\frac{2\pi n}{N-1}\}_{n=0}^{N-2}$ for all $k$, the $(n+1)$th element of $\bar{x}$ is

$$\bar{x}[n] = F_{n+1}[J(\omega)\epsilon] = F_{n+1}[A(\omega)]_{n+1}(D - I_K)\epsilon = [0, 0, \ldots, 0](D - I_K)\epsilon = 0$$  

for $n > \lfloor \frac{N-1}{\gamma} \rfloor$, where $D = \text{diag}(e^{j\omega_0}, \ldots, e^{j\omega_K})$. When frequencies $\omega$ are not on the grid, despite the presence of spectral leakage, it can be expected that the magnitude of the $(n+1)$th element of the first-order difference line spectral signal in the frequency domain, i.e., $|\bar{x}[n]|$ or $|\bar{x}[n]|$, is small for some $n > \lfloor \frac{N-1}{\gamma} \rfloor$. In the following, a formal analysis is provided. We first derive the upper bound of the magnitude of the $(n+1)$th element of $\bar{x}$, which is shown in Proposition 2. For $(N-1)(\frac{\gamma}{\gamma} + \beta) \leq n \leq (N-1)(1 - \beta)$ where $\beta$ is a constant factor such that $\frac{1}{N-1} < \beta < \frac{\gamma - 1}{2\gamma}$.
according to (20), we have

\[
\begin{align*}
\|\tilde{x}[n]\| & \leq \frac{2K\pi c_{\max}s_{\max}\delta_{\max}}{\sqrt{N-1}} \min\{\sin\left(\frac{\pi n}{N-1} - \frac{\pi}{\gamma}\right), \sin\left(\frac{\pi n}{N-1}\right)\} \\
& \leq \frac{2K\pi c_{\max}s_{\max}\delta_{\max}}{\sqrt{N-1}} \min\{\sin(\frac{\pi}{\gamma} + \beta), \sin(\beta\pi)\} \\
& \leq \frac{2K\pi c_{\max}s_{\max}\delta_{\max}}{\sqrt{N-1}} \sin(\beta\pi)
\end{align*}
\]

(19)

where \( a \) is due to both \( \sin\left(\frac{\pi n}{N-1} - \frac{\pi}{\gamma}\right) \) and \( \sin\left(\frac{\pi n}{N-1}\right) \) are greater than 0 and are concave for \((N-1)(\frac{1}{\gamma} + \beta) \leq n \leq (N-1)(1 - \beta)\). Thus, \( \|\tilde{x}[n]\| \) can be arbitrarily small with increasing \( N \). This demonstrates that with enough samples, one could choose the samples \( \tilde{x}[n] \) with index \( \{(N-1)(\frac{1}{\gamma} + \beta) + 1, (N-1)(\frac{1}{\gamma} + \beta) + 2, \ldots, (N-1)(1-\beta)\} \) such that the leakage is negligible, thus an optimization problem only with respect to \( \xi \) can be established, which is introduced ahead.

**Proposition 2** Define \( a_k \in \mathbb{Z}, a_k \leq \left\lfloor \frac{N-1}{\gamma} \right\rfloor \) and \( \delta_k \in [-0.5, 0) \cup (0, 0.5), k = 1, \ldots, K \) such that \( \omega_k = \frac{2\pi}{N-1} (a_k + \delta_k) \). For \( n > \left\lfloor \frac{N-1}{\gamma} \right\rfloor \), the magnitude of the \((n+1)\)th element of \( \tilde{x} \) can be upper bounded by

\[
\|\tilde{x}[n]\| \leq \frac{2K\pi c_{\max}s_{\max}\delta_{\max}}{\sqrt{N-1}} \min\{\sin\left(\frac{\pi n}{N-1} - \frac{\pi}{\gamma}\right), \sin\left(\frac{\pi n}{N-1}\right)\}
\]

(20)

where \( c_{\max} \triangleq \max_{i=1}^{K}|c_i| \), \( s_{\max} \triangleq \max_{i=1}^{K}\sin(\frac{\pi n}{N-1}) \), and \( \delta_{\max} \triangleq \max_{i=1}^{K}|\delta_i| \).

**Proof** The proof is postponed to Appendix B. \( \square \)

Let \( S \) be a subset of \( \{1, 2, \ldots, N-1\} \) defined as

\[
S \triangleq \left\{\left\lfloor \frac{N-1}{\gamma} \right\rfloor + N_\beta + 2, \ldots, \left\lfloor \frac{N-1}{\gamma} \right\rfloor + N - N_\beta + 1\right\}
\]

(21)

where \( N_\beta \triangleq (N-1)\beta \) and \( \frac{1}{N-1} < \beta < \frac{1}{2} \) is a constant factor. By choosing the \( S \) elements of (17), a submodel

\[
\tilde{x}_S + 2\lambda F_S \xi = \tilde{x}_S + \tilde{w}_S
\]

(22)

can be obtained. Note that the number of measurements of model (22) is \( |S| = \left|\left\lfloor \frac{N-1}{\gamma} \right\rfloor + N_\beta + 1\right| \approx (N-1)(1 - \frac{1}{2} - 2\beta) \). On the one hand, a smaller \( \beta \) means that more data can be used to perform signal recovery, and the measurement matrix \( 2\lambda F_S \xi \) has a lower mutual coherence. On the other hand, \( \beta \) increases, the impact caused by spectrum leakage will be reduced, as stated in Proposition 2 and (19). Thus, there exists a tradeoff for selecting \( \beta \).

We validated this phenomenon through simulation for the algorithm we proposed later on in Section V-A. By treating the right-hand side of (22), i.e., \( \tilde{x}_S + \tilde{w}_S \), as noise, based on the least-squares method, the subproblem of estimating \( \xi \) can be formulated as

\[
\begin{align*}
\text{minimize} & \quad \|z_S + F_S \xi\|^2 \\
\text{s.t.} & \quad \xi \in \mathcal{V}, n = 1, \ldots, N - 1
\end{align*}
\]

(23a)

\[Q_{i,j} = \left\{ Q_{i,j} | i - j | \leq p \right\} \text{ otherwise}
\]

(26)

where \( z_S \triangleq \mathcal{V}/(2\lambda), \mathcal{V} = \{m + jn | V, m, n \in \mathbb{N}\} \) is the bounded lattice and \( V \) is a small integer, as shown in Proposition 1. It should be noted that there exists ambiguity in solving the problem (23). Specifically, when \( \xi^* \) is a solution to problem (23), it is noteworthy that \( \xi^* = \xi + c1 \), where \( c \in \mathbb{Z} \) and \( 1 \) is the all-ones vector, also serves as a solution to problem (23) if \( \xi \in \mathcal{V}^{N-1} \). Problem (23) is an integer quadratic programming problem which is also NP-hard generally. In the following, an efficient algorithm based on DP is proposed to approximately solve (23).

**IV. ALGORITHM**

This section provides a polynomial-time algorithm for solving (23) approximately. In general, this problem is difficult to solve. Our approach consists of two steps: First, the original problem (23) is solved approximately by exploiting the fine structure of (23). That is, we approximate the original problem (23) in a reasonable way and use the DP method to solve the problem globally [45]. Second, the approximate solution acts as a good candidate, and heuristic approaches are used to further refine the quality of the solution by solving the original problem (23). At the end of this section, we will provide the validation of the algorithm by showing the rationality of solving the approximate optimization problem via DP.

**A. DP Method**

By expanding the quadratic term, problem (23) can be rewritten as

\[
\begin{align*}
\text{minimize} & \quad \xi^H Q \xi + 2\Re\{b^H \xi\} \\
\text{s.t.} & \quad \xi \in \mathcal{V}, n = 1, \ldots, N - 1
\end{align*}
\]

(24a)

\[Q_{i,j} = F_{S,i}^H F_{S,j}.\]

(25)

When \( S = \{1, \ldots, N-1\} \), \( F_S = F \) is the DFT matrix. Thus, \( F_{S,i} \) and \( F_{S,j} \) are orthogonal when \( i \neq j \) and \( F_{S,i}^H F_{S,j} = 1 \) when \( i = j \) and \( Q \) is an identity matrix. For \( S \) defined in (21), which is a subset of \( \{1, \ldots, N-1\} \), although \( Q \) is not an identity matrix, elements with relatively large magnitudes still concentrate near the diagonal. In detail, when \( i = j \), the magnitude of \( Q_{i,j} \) will get closer to 1 as \( |S| \) increases. When \( i \neq j \), the magnitude of \( Q_{i,j} \) will be relatively small, especially when the wrap-around distance between \( i \) and \( j \), i.e., \( \min(|i - j|, N - 1 - |i - j|) \), is far away from 0. Thus, the matrix \( Q \) only has relatively large values on its diagonal and a few adjacent positions (wrap-around distance). Approximating \( Q \) as a \( p \)-th-order diagonal matrix \( \tilde{Q} \) is reasonable, where \( \tilde{Q} \) is a submatrix of \( Q \) and is defined as follows:

\[
\tilde{Q}_{i,j} = \left\{ Q_{i,j} | i - j | \leq p \right\} \text{ otherwise}
\]

(26)

Authorized licensed use limited to the terms of the applicable license agreement with IEEE. Restrictions apply.
where \( i = 1, \ldots, N - 1 \) and \( j = 1, \ldots, N - 1 \). Note that \( Q \) is still a Hermitian matrix. In Section IV-C, a detailed analysis of the rationality of approximating \( Q \) as \( \tilde{Q} \) will be provided. By replacing \( Q \) with \( \tilde{Q} \) in (24), an approximate optimization problem can be formulated as

\[
\begin{align*}
\text{minimize} & \quad \epsilon^H \tilde{Q} \epsilon + 2\Re\{b \epsilon_n\} \\
\text{subject to} & \quad \epsilon_n \in \mathcal{V}, \quad n = 1, \ldots, N - 1.
\end{align*}
\]  

\((27a)\)

\((27b)\)

We first consider the quadratic term in (27), i.e., \( \epsilon^H \tilde{Q} \epsilon \). For \( \tilde{Q}_{i,j} = 0 \) when \( |i-j| > p \), the cost function \( \epsilon^H \tilde{Q} \epsilon \) is a function of \( \epsilon_n \) and its nearby terms \( \{ \epsilon_{n'} \mid 0 < |q-n| \leq p \} \). In addition, the second term in (27) can be decomposed as independent sums, i.e., \( 2\Re\{b^* \epsilon_{n}\} = \sum_{j=1}^{N-1} 2\Re\{b^* \epsilon_{n}\} \). Thus, parameters \( \epsilon \) in problem (27) have a \( p \)th order Markov property, and the global optimum can be obtained by using a DP method.

Let \( h(\epsilon) = \epsilon^H \tilde{Q} \epsilon + 2\Re\{b \epsilon_n\} \), which is the objective function of problem (27), and we decompose it into \( P = N - p - 1 \) terms, i.e.,

\[
h(\epsilon) = \sum_{n=1}^{P-1} \epsilon^H \tilde{Q}^{(n)} \epsilon + 2\Re\{b^* \epsilon_n\} + \left( \epsilon^H \tilde{Q}^{(P)} \epsilon + 2\Re\{b^* \epsilon_{N-1} \} \right)
\]

\((28)\)

where \( \tilde{Q} = \sum_{n=1}^{N-1} \tilde{Q}^{(n)} \) and \( \tilde{Q}^{(n)} \) are all Hermitian matrices defined as follows. For \( n = 1, \ldots, P - 1 \), \( \tilde{Q}^{(n)} \in \mathbb{C}^{(N-1)\times(N-1)} \) is a submatrix of \( \tilde{Q} \) retaining the elements in columns \( n \) through \( N - 1 \) of \( \tilde{Q} \)’s \( n \)th row, as well as the elements in rows \( n \) through \( N - 1 \) of \( \tilde{Q} \)’s \( n \)th column defined as

\[
\tilde{Q}_{i,j}^{(n)} = \begin{cases} 
\tilde{Q}_{i,j} & (i = n \text{ and } j \geq i) \text{ or } (j = n \text{ and } i \geq j) \\
0, & \text{otherwise}.
\end{cases}
\]

For the last term, \( \tilde{Q}^{(P)} \in \mathbb{C}^{(N-1)\times(N-1)} \) is a submatrix of \( \tilde{Q} \) retaining both \( \{P, \ldots, N - 1\} \) rows and columns of \( \tilde{Q} \) defined as

\[
\tilde{Q}_{i,j}^{(P)} = \begin{cases} 
\tilde{Q}_{i,j} & (i \geq P) \text{ or } (j \geq P) \\
0, & \text{otherwise}.
\end{cases}
\]

\((29)\)

Because \( \tilde{Q}_{i,j} = 0 \) when \( |i-j| > p \), we have \( \tilde{Q}_{i,n} = 0 \) for \( j > n + p \) and \( \tilde{Q}_{i,n} = 0 \) for \( i > n + p \). Thus, the \( n \)th term, \( n = 1, \ldots, P - 1 \), of the decomposed loss (28) can be simplified as

\[
\epsilon^H \tilde{Q}^{(n)} \epsilon + 2\Re\{b^* \epsilon_n\} = \epsilon_n^H \tilde{Q}_{n,n} \epsilon_n + 2\Re\{\epsilon_n^H \tilde{Q}_{n,n+1,n+p} \epsilon_{n+1,n+p} + 2\Re\{b^* \epsilon_n\} \}
\]

\((30)\)

which is only related to variables \( \epsilon_{n+1,n+p} \) and denoted as \( h_n(\epsilon_{n+1,n+p} \) for the last term in (28). According to (29), we have

\[
\epsilon^H \tilde{Q}^{(P)} \epsilon + 2\Re\{b^* \epsilon_{N-1}\} = \epsilon^H \tilde{Q}_{P,N-1} \epsilon + 2\Re\{b^* \epsilon_{P,N-1}\} \]

\((31)\)

which is only related to the last \( p + 1 \) variables \( \epsilon_{P,N-1} \) and denoted as \( h_P(\epsilon_{P,N-1}) \). In summary, we can decompose the objective function (28) into \( P \) terms, i.e.,

\[
h(\epsilon) = \sum_{n=1}^{P} h_n(\epsilon_{n+1,n+p})
\]

\((32)\)

where \( h_n(\epsilon_{n+1,n+p}) \) are defined in (30) and (31).

Below the DP is utilized to solve the approximate optimization problem (27) by using the decomposed form of the objective function \( h(\epsilon) \) (32). We first define the optimization problem at stage \( n \) (for the first \( n \) variables) as

\[
\text{DP}_n(\epsilon_{n+1,n+p}) = \min_{\epsilon_n, \epsilon_{n+1,n+p}} \sum_{i=1}^{n} h_i(\epsilon_{i+1})
\]

\((33)\)

\(n = 1, \ldots, P - 1\). Observe that the optimization problem (27) is equivalent to

\[
\text{minimize}_{\epsilon_n, \epsilon_{n+1,n+p}} \sum_{i=1}^{n} h_i(\epsilon_{i+1}) + h_n(\epsilon_{n+1,n+p})
\]

\((34)\)

\(n = 1, \ldots, P - 1\), where \( \text{DP}_0(\epsilon_{1}) = 0 \) is the initialization term. For \( \epsilon_n \in \mathcal{V}, n = 1, \ldots, N - 1 \), the enumeration method is utilized here to solve the recursion (35), and the computational complexity is \( O(p^2|\mathcal{V}|^{p+1}) \). In addition, the optimal solution \( \hat{\epsilon}_n \) will be recorded by RCn(\( \epsilon_{n+1,n+p} \)) for \( n = 1, \ldots, P - 1 \). Finally, we will solve the final optimization problem (34) based on \( \text{DP}_{P-1}(\epsilon_{P,N-2}) \). Similarly, an enumeration method can be applied to obtain the optimal solution \( \hat{\epsilon}_{P,N-1} \), and the computational complexity of the final step is \( O(p^2|\mathcal{V}|^{p+1}) \). To obtain the corresponding optimal variables, a backward process

\[
\hat{\epsilon}_n = \text{RC}_n(\hat{\epsilon}_{n+1,n+p})
\]

\((36)\)

is recursively executed, for \( n = P - 1, \ldots, 1 \). The initial estimates \( \hat{\epsilon}_{P,N-1} \) is the solution to (34). The DP algorithm is summarized in Algorithm I.

B. USLSE Algorithm

The DP method has been used to solve the problem (24) approximately. Here, we propose an enhanced iterative algorithm to further refine the solution \( \hat{\epsilon} \). The algorithm proceeds as follows: By viewing \( \hat{\epsilon} \) as the initial value, \( \epsilon \) can be decomposed into two parts, i.e., \( \epsilon = \hat{\epsilon} + \Delta \epsilon \), and the original optimization problem (23) is

\[
\text{minimize}_{\Delta \epsilon} \| z_S + F_S \hat{\epsilon} + F_S \Delta \epsilon \|^2_2
\]

\((37a)\)

Authorized licensed use limited to the terms of the applicable license agreement with IEEE. Restrictions apply.
Algorithm 1: DP Algorithm.

1: Inputs: Observations $z_S$, measurement matrix $F_S$,
state space $V$, and order $P$.
2: Compute $b = F_S^H z_S$ and $Q(n)$, $n = 1, \ldots, P$.
3: Forward process:
4: Initialize $D P(\hat{\epsilon}_1, P) = 0$.
5: for $n = 1, \ldots, P - 1$ do
6: Compute $D P(\hat{\epsilon}_{n+1}, P)$ according to (35) and store it, and delete $D P_{n-1}(\hat{\epsilon}_{n+p-1})$.
7: Store the corresponding optimal solution $\hat{\epsilon}_n$ in $R C_n(\hat{\epsilon}_{n+1}, P)$.
8: end for
9: Obtain the estimates $\hat{\epsilon}_{P,N-1}$ by solving (34), and delete $D P_{P-1}(\hat{\epsilon}_{P,N-2})$.
10: Backward process:
11: for $n = P - 1, \ldots, 1$ do
12: Compute $\tilde{\epsilon}_n = R C_n(\hat{\epsilon}_{n+1}, P)$, and delete $R C_n(\hat{\epsilon}_{n+1}, P)$.
13: end for
14: Outputs: $\tilde{\epsilon}_n$.

s.t. $\Delta \epsilon_n \in Z^*$, $n = 1, \ldots, N - 1$. (37b)

Given that $\tilde{\epsilon}$ serves as a good approximation to the optimal solution, $\Delta \epsilon$ will be a sparse complex integer vector. Although problem (37) remains an integer quadratic programming problem, which is generally hard to solve, sparse-based methods such as OMP or basis pursuit can be applied here [46], [47]. Here, since $\Delta \epsilon_n \in Z^*$, we run the OMP algorithm, and for each iteration, the current estimate will be rounded to the lattice $Z^*$, and the OMP algorithm will terminate when the current estimate after rounding is 0. Finally, the estimates will be updated as $\tilde{\epsilon} := \tilde{\epsilon} + \Delta \tilde{\epsilon}$ provided that the refined estimate decreases the objective function value; otherwise, we still use the old estimate. We term the above procedure as “DP+OMP”.

Multiple iterations of “DP+OMP” can be executed to enhance performance. Given the estimate $\tilde{\epsilon}$, we expand (37) and obtain

$$ \min_{\Delta \tilde{\epsilon}} (\Delta \tilde{\epsilon})^H Q \Delta \tilde{\epsilon} + 2 \Re \{ (z_S + F_S \tilde{\epsilon})^H F_S \Delta \tilde{\epsilon} \} \quad (38a) $$

s.t. $\Delta \epsilon_n \in Z^*$, $n = 1, \ldots, N - 1$. (38b)

We still use the same banded matrix $\tilde{Q}$ to approximate $Q$ and run DP to solve the problem approximately. Similarly, we then use the OMP to refine the estimates. The algorithm that runs iter of iterations of DP and OMP is termed as “DP+OMP(iter).”

Note that DP+OMP(iter) provides the estimate $\tilde{\epsilon}$. The problem of additive constant ambiguity in $\tilde{\epsilon}$, as discussed in problem (23), is resolved by adjusting $\tilde{\epsilon}$ as follows: $\tilde{\epsilon} := \tilde{\epsilon} - \lfloor \frac{1}{N-1} \sum_{i=1}^{N-1} \hat{\epsilon}_i \rfloor$, where $\lfloor \cdot \rfloor$ denotes the rounding operator. To recover $\epsilon$ form $\tilde{\epsilon}$, the antidifference operator denoted as $\hat{\epsilon} = S(\tilde{\epsilon})$ will be used and is defined as

$$ \hat{\epsilon}_1 = 0 \quad (39a) $$

It should be noted that there is also an additive constant ambiguity of $\hat{\epsilon}$. We can obtain the recovered line spectral signal via $\hat{\epsilon} = y + 2 \Delta \hat{\epsilon}$. Finally, an LSE algorithm is applied to obtain $[\hat{\omega}, \hat{\epsilon}]$. Overall, the USLSE algorithm is summarized in Algorithm 2. USLSE is a two-stage algorithm where line spectral signal reconstruction is performed first, followed by the LSE. The whole process is shown in Fig. 2.

In the following, we will analyze the computational complexity of the USLSE algorithm. Note that USLSE is a two-stage algorithm: In the first stage, a signal unfolding algorithm is applied, and in the second stage, the LSE and detection algorithm is executed. In this article, we mainly focus on the first stage. Therefore, we will only analyze the computational complexity of the signal unfolding algorithm. In the unfolding stage of the USLSE, we first apply FFT to obtain data in the frequency domain whose computation complexity is $O(N \log N)$. Then, we iteratively execute DP followed by OMP to estimate the first-order finite difference of the simple function $\epsilon$. For the DP algorithm, the computational complexity is $O(\epsilon^2)$. For the OMP method, the computational complexity is $O((\gamma - 1)N^2)$ where $H = \| \Delta \epsilon \|_0$ is the number of nonzero elements in $\Delta \epsilon$. Given the maximum number of iterations $\text{Iter}_{\text{max}}$, the computational complexity is $O(\text{Iter}_{\text{max}}(\epsilon^2) | V |^{p+1} N + (\gamma - 1)N^2)$). Finally, the estimate of the line spectral signal $\hat{g}$ is obtained from $\hat{\epsilon}$ with computational complexity $O(N)$. In summary, the computational complexity of the first stage of USLSE is $O(\text{Iter}_{\text{max}}(\epsilon^2) | V |^{p+1} N + (\gamma - 1)N^2)$, which is dominated by DP and OMP algorithms. We can see that for the DP algorithm, its computational complexity increases exponentially with $p$ for a fixed $| V |$. In addition, although the computational complexity of the OMP algorithm with respect to $N$ is quadratic while the computational complexity of the DP algorithm is linear with respect to $N$, in practical applications, unless a very large $N$ is used, the computational time consumed by the DP algorithm is much greater than that of the OMP.

In addition, we compare the computational complexity of the unfolding stage of USLSE with other state-of-the-art signal unfolding algorithms, such as LP [30], USAlg [20], Prony [34], and QCQP [37]. The results are shown in Table I, where $R$ denotes the number of correlated samples in LP, $D$ is the difference order in USAlg. For LP and USAlg, the computational complexities are low as it is linear with respect to $N$. The Prony method consists of fast Fourier transform and root finding, and its computational complexity is $O(N \log N + P^3)$, where $P$ denotes the number of nonzero elements of the first-order difference of $\epsilon$, i.e., $P = \| \epsilon \|_0$. Therefore, when $\epsilon$ is sparse, the computational complexity is low. The worst-case computational complexity of QCQP is $O(N^3)$, which is a cubic polynomial with respect to $N$. Finally, for the USLSE algorithm, the computational complexity increases exponentially with $p$ and is a quadratic
Algorithm 2: USLSE Algorithm.

1: **Inputs:** Modulo samples \( y \), number of sinusoids \( K \), and oversampling factor \( \gamma \).
2: Compute \( \tilde{y} \).
3: Set subset \( S \), state space \( V \), and Markov model order \( p \).
4: Initialize \( \tilde{z} = 0 \).
5: for iter = 1 : \( \text{iter}_{\text{max}} \) do
6: Compute \( \tilde{A} = DP(z + F_S \tilde{z}, F_S, V, p) \) (Algorithm 1).
7: Update \( \tilde{z} := \tilde{z} + \Delta \tilde{z} \).
8: Compute \( \Delta \tilde{z} \) by solving (37) (OMP).
9: Update \( \tilde{z} := \tilde{z} + \Delta \tilde{z} \).
10: end for
11: Update \( \tilde{x} := \tilde{x} - [\sum_{i=1}^{N-1} \tilde{c}_i/(N - 1)] \).
12: Compute \( \tilde{e} = S(\tilde{x}) \) (up to an additive constant).
13: Compute \( \tilde{g} = \tilde{y} + 2\tilde{\varphi} \).
14: Estimate parameters \( \{\tilde{\varphi}, \tilde{c}\} \) by applying an LSE algorithm on \( \tilde{g} \).
15: **Outputs:** \( \{\tilde{\varphi}, \tilde{c}\} \).

C. Rationality of Approximating \( \mathbf{Q} \) as \( \tilde{\mathbf{Q}} \)

In this section, the rationality of approximating \( \mathbf{Q} \) as \( \tilde{\mathbf{Q}} \) is validated by evaluating the ratio between the squared Frobenius norm of \( \mathbf{Q}_p \) to that of \( \mathbf{Q} \), where \( \mathbf{Q}_p \) is the \( p \)-th order approximation of \( \mathbf{Q} \) defined in (26). Proposition 3 provides both the exact expression and a lower bound for the ratio \( \|\mathbf{Q}_p\|_F^2 / \|\mathbf{Q}\|_F^2 \). Note that for \( p = 0 \), a simple bound is \( \|\mathbf{Q}_p\|_F^2 \geq 1 - \eta \). In addition, when \( N \) and \( M \) are significantly larger than \( p \), the lower bound can be approximated as a function only with respect to \( \eta \) and \( p \), i.e., \( 1 - \eta + \frac{8}{\pi(1-\eta)} \sum_{k=1}^{p} \frac{\sigma_k^2}{k^2} \), based on the fact that \( \frac{M-1-k}{N-1-k} \approx \frac{M-1}{N-1} = \eta \) for \( k = 1, \ldots, p \). For a practical case, where the oversampling factor is \( \gamma = 10 \) and \( \eta = \frac{M-1}{N-1} \) is slightly larger than \( \frac{1}{10} \) set as \( 1/10 < \eta < 1/8 \), the results are then listed as follows: for \( p = 1 \), the lower bound is \( 1 - \eta + \frac{4\sigma_1^2}{\pi(1-\eta)} \), and the minimum value is approximately \( 0.890 \) with \( \eta = 1/8 \); for \( p = 2 \), the lower bound is \( 1 - \eta + \frac{16\sigma_1^2}{\pi(1-\eta)} \), and the minimum value is approximately \( 0.904 \) with \( \eta = 1/8 \); for \( p = 3 \), the lower bound is \( 1 - \eta + \frac{24\sigma_1^2}{\pi(1-\eta)} \), and the minimum value is approximately \( 0.918 \) with \( \eta = 1/8 \); for \( p = 4 \), the lower bound is \( 1 - \eta + \frac{32\sigma_1^2}{\pi(1-\eta)} \), and the minimum value is approximately \( 0.933 \) with \( \eta = 1/8 \). It can be seen that the ratio of the squared Frobenius norm of the approximated \( \tilde{\mathbf{Q}}_p \) to that of \( \mathbf{Q} \) is at least \( 0.890 \), corresponds to the case when \( p = 1 \). For \( p = 3 \) and \( p = 4 \), the ratio of the squared Frobenius norm of the approximated \( \tilde{\mathbf{Q}}_p \) to that of \( \mathbf{Q} \) is at least \( 0.918 \). Therefore, approximating \( \mathbf{Q} \) as \( \tilde{\mathbf{Q}}_p \) is reasonable in this case.

PROPOSITION 3. Let \( M \triangleq N - |S| \), the ratio of the energy of the banded matrix \( \tilde{\mathbf{Q}}_p \) with bandwidth \( p \) to that of \( \mathbf{Q} \) is

\[
\frac{\|\tilde{\mathbf{Q}}_p\|_F^2}{\|\mathbf{Q}\|_F^2} = 1 - \frac{M-1}{N-1} + 2 \frac{\sum_{k=1}^{p} \sin^2(\frac{k}{N-1} \pi)}{(N-1)^2(N-M)}. \quad (40)
\]

In addition, for \( p \leq \frac{N-1}{2} \), \( \frac{\|\tilde{\mathbf{Q}}_p\|_F^2}{\|\mathbf{Q}\|_F^2} \) can be lower bounded by

\[
\frac{\|\tilde{\mathbf{Q}}_p\|_F^2}{\|\mathbf{Q}\|_F^2} \geq 1 - \eta + \frac{8}{\pi^2} \sum_{k=1}^{p} \frac{u_k^2}{k^2(1 - \frac{M-1-k}{N-1})}. \quad (41)
\]
where \( \eta \equiv \frac{M-1}{N-1} \), and
\[
\eta_{kn} = \begin{cases} 
kn - \lfloor \eta \rfloor, & 0 \leq kn - \lfloor \eta \rfloor < \frac{1}{2} \\
1 + \lfloor \eta \rfloor - kn, & \frac{1}{2} \leq kn - \lfloor \eta \rfloor < 1 
\end{cases}.
\]

**Proof.** The result \((40)\) can be obtained by a straightforward calculation. Based on \((40)\), the lower bound \((41)\) can be derived as follows:
\[
\frac{\|\tilde{Q}\|_F^2}{\|Q\|_F^2} \geq 1 - \eta + \frac{2}{\pi^2} \sum_{k=1}^{p} \frac{N - 1 - k}{N - M} \sin^2 \left( \frac{k(M-1)\pi}{N-1} \right) \frac{1}{k^2}.
\]
where \(\gamma \) is due to \(\sin \left( \frac{k\pi}{N-1} \right) \leq \frac{k}{N-1} \). In addition, \(\geq\) is derived from the inequality \(\frac{\lambda x}{2} \leq \sin x \) \(\forall 0 \leq x \leq \frac{\pi}{4} \), and it is established through discussions involving different intervals of \(\frac{k(M-1)\pi}{N-1} \).

V. NUMERICAL SIMULATION

In this section, numerical experiments are conducted to verify the performance of the proposed USLSE. The signal-to-noise ratio (SNR) is defined as \(\text{SNR} = \frac{\|x\|_2^2}{\|\tilde{x} - x\|_2^2} \), where \(x\) is the noise-free signal and \(\tilde{x}\) is the noise. In this section, we set \(w\) as white Gaussian noise. The normalized mean-squared error (NMSE) is \(\text{NMSE}(\hat{x}) = \frac{\|\hat{x} - \tilde{x}\|_2^2}{\|x\|_2^2} \), where \(\hat{x}\) denotes the estimate of the algorithm. For the first stage of USLSE, the methods DP, DP+OMP, DP+OMP(ite) are used to solve the optimization problem \((23)\) where \(\text{iter} = 2, 3, 4\), and the Markov assumption with different orders \((p = 1, 2, 3, p = 4)\) are implemented. Given the estimates \(\hat{x}\), the estimates of the simple function \(\hat{c}\) up to an additive constant can be obtained by using the antiderivative operator \((39)\). The ambiguity is resolved by using \(\hat{c} := \hat{c} + \left( \sum_{i=1}^{\lfloor N/\gamma \rfloor} (\sum_{j=1}^{i} c_j) / N \right)\), where \(\lfloor \cdot \rfloor\) denotes the rounding operator. Finally, a state-of-the-art LSE algorithm will be applied to obtain the estimates \(\hat{x}\) and \((\hat{c}, \hat{\omega})\) based on \(\hat{g} = y + 2\hat{c}\). Here, Newtonalized OMP (NOMP) \([48]\) due to its fast implementation and high estimation accuracy is used to perform LSE. For different implementations of USLSE, the only difference is the method used to solve the optimization problem \((23)\). Thus, we will use the name of the method solving \((23)\) as the name of different implementations of USLSE. We define the event in which the algorithm successfully recovers the true signal \(x\) provided that \(\text{NMSE}(\hat{x}) < -15 \) dB and the probability of successful recovery is adopted as the performance metric. Additionally, we utilize the complementary cumulative distribution function (CCDF) of the mean-squared error (MSE) of frequencies \(\omega\) as another performance metric. The simulation runs on a desktop computer equipped with an 8-core Mac M1 CPU.

---

\(^{2}\) Different thresholds such as \(-10\) dB, \(-15\) dB, and \(-20\) dB are tried in the simulation, and the results are similar.

Five experiments are conducted to validate the theoretical analysis and verify the excellent performance of USLSE. For the first experiment, the effect of the hyperparameter \(\beta\) on USLSE is investigated. Then, the performance of USLSE compared to other state-of-the-art algorithms will be studied. Additionally, the performance of USLSE versus SNR is investigated in the third experiment. Furthermore, Section V-D studies the performance of USLSE in the scenario where both strong and weak targets coexist. Finally, the reconstruction of bandlimited signals through USLSE is conducted. Parameters of the first and third experiments (see Sections V-A and V-C) are set as follows: The oversampling factor is \(\gamma = 10\), and the number of measurements is \(N = 512\). The number of frequencies is \(K = 3\), and the frequencies are randomly generated from \((0, 2\pi / \gamma)\). The magnitudes of the weight coefficients are drawn i.i.d. from a Gaussian distribution \(\mathcal{N}(1, 0.1)\), and the phases are uniformly drawn i.i.d. from \((0, 2\pi)\). The dynamic range of the ADC is set as \(\lambda = 0.7\). We set \(V = 1\) for USLSE in all simulations. All results are averaged over 300 Monte Carlo (MC) trials.

A. Effect of Hyperparameter \(\beta\) on Model Performance

According to \((21)\), \(\frac{1}{N-1} \approx 0 < \beta < \frac{1}{2\gamma} \approx 0.5\) is a hyperparameter to set up the optimization problem \((23)\) and should be chosen in advance. We will investigate the performance of the algorithms (OMP, DP, DP+OMP, DP+OMP(iter)) versus \(\beta\), where OMP denotes the algorithm that uses the OMP solving problem \((23)\) followed by NOMP to perform the LSE. We set \(\text{SNR} = 22\) dB, and the probability of successful recovery and the averaged running time are shown in Fig. 3. To make the figures more readable, the result of OMP, which is invariant for different \(p\), is shown in all the figures for comparison.

In Fig. 3(a) where \(p = 1\), the probabilities of successful recovery of all algorithms decrease as \(\beta\) increases. Besides, DP+OMP(iter) with \(\text{iter} = 2, 3, 4\) have the same probability of successful recovery and perform best, followed by DP+OMP, OMP, and DP. For \(p\) ranging from 2 to 4 shown in Fig. 3(b), 3(c), and 3(d), respectively, DP+OMP(2) – (4) perform best and achieve similar probability of successful recovery. In detail, the probabilities of successful recovery of DP+OMP(2) – (4) increase first from 0.48 to 0.78 as \(\beta\) increases to \(\beta = 0.07\) and then decreases for \(p = 2\), from 0.58 to 0.86 as \(\beta\) increases to about \(\beta = 0.04\) and then decreases for \(p = 3\), from 0.7 to 0.86 as \(\beta\) increases to about \(\beta = 0.02\) and then decreases for \(p = 4\). Besides, DP+OMP(2) – (4) benefits from OMP in particular for \(p = 1\) and \(p = 2\). The reason is that with a lower value of \(p\), the gap between the optimization problem solved by DP \((27)\) and the problem solved by OMP \((23)\) may be larger due to the increased inaccuracy in the approximation of \(Q\). As shown in Fig. 3, for each \(p\), the running time increases in the order of OMP, DP, DP+OMP, DP+OMP(2), DP+OMP(3), DP+OMP(4). Furthermore, the running time of the DP-based algorithms increases as \(p\) increases. For all
algorithms, the running time of DP+OMP(4) with \( p = 4 \) is the longest, which is 4.319 s. Except for \( p = 1 \) in Fig. 3(a), the performances of all DP-based algorithms increase first and then decrease as \( \beta \) increases. The reason is that a higher \( \beta \) leads to a lower \( \gamma \) in (22) and a smaller number of measurements. Thus, there is a tradeoff when it comes to selecting the optimal \( \beta \) for DP-based algorithms. According to Fig. 3(b)–3(d), as \( p \) increases, the optimal \( \beta \) decreases, and the optimal probabilities of successful recovery of \( p = 3 \) and \( p = 4 \) are almost the same, which is greater than that of \( p = 2 \). In fact, although as \( p \) increases, \( Q \) approximates \( Q \) better, we cannot conclude that a better estimate will be obtained. We can observe that when \( \beta = 0.05 \), the performance of DP-based algorithms with \( p = 2 \) is better than those with \( p = 3 \), and the performance of DP-based algorithms with \( p = 3 \) is better than those with \( p = 4 \). By carefully investigating the numerical simulation, we found that as \( \beta \) increases, DP-based algorithms with larger \( p \) tend to overfit and result in poor performance.

In the following simulations, DP-based algorithms with \( p = 2, 3, 4 \) will be adopted, and the hyperparameter \( \beta \) will be set as \( \beta = 0.07, 0.04, \) and 0.02, respectively, corresponding to the optimal choices as shown in Fig. 3(b)–3(d).\(^3\) As the performance of the OMP decreases as \( \beta \) increases as shown in Fig. 3, the set \( \{[\frac{N-1}{\gamma}] + 2, \ldots, N - 1\} \).

\(^3\)We also conducted simulations for different SNRs, and the results all demonstrate that \( \beta = 0.07, 0.04, \) and 0.02 are good choices for DP-based algorithms with \( p = 2, 3, \) and 4, respectively.

### B. Comparison With Other Algorithms

In this section, the USLSE is compared with existing methods such as LP [30], USAlg [20], Prony [34], and QCQP [37] that recover the spectrum from modulo samples. It should be noted that \( g \) is a complex vector in our model while USAlg and QCQP are proposed to deal with the real signal. Thus, we will process the real and imaginary parts of \( g \) separately when using USAlg and QCQP. For the LP method, we assume that the initial few unfolded samples of the signal can be obtained and the autocorrelation function is known. Although the blind version of LP is proposed in [33], that method requires adaptive adjustment of the ADC’s dynamic range, which is not the case in this article. For USLSE, DP+OMP(2) with \( p = 3 \) is used to solve the problem (23). The recovery results for the real part of the line spectral signal using all algorithms are shown in Fig. 4, with similar results observed for the imaginary part.

Fig. 4(a) shows the results in a noisy and off-grid scenario with a relatively high oversampling factor \( \gamma \approx 150 \). LP, USAlg, QCQP, and USLSE successfully recover the original signal while Prony fails. In Fig. 4(b), where the signal is noiseless and the frequencies are precisely on the grid, all algorithms, except QCQP, effectively recover the unfolded signal. As in the cases of Fig. 4(c) and (d), the cardinality of nonzero elements in \( \xi \), i.e., \( \|\xi\|_0 \), is greater than \( N/2 \), the Prony method cannot be utilized in these two scenarios. Fig. 4(c) presents the results in a noisy and off-grid setting with \( \gamma \approx 12 \). LP and USLSE achieve successful recovery of the original signal while USAlg and QCQP do not. In Fig. 4(d) with an SNR of 30 dB and a relatively low
oversampling factor $\gamma \approx 6$, all algorithms, except LP, fail to recover the original line spectral signal. In summary, QCQP is effective in recovering signals from noisy modulo samples but requires a relatively high oversampling factor. Prony’s performance is sensitive to noise and relies on sparsity in $\epsilon$, leading to its failure in Fig. 4(a), (c), and (d). In cases with a large oversampling factor, as in Fig. 4(a), or with a noiseless original signal, as in Fig. 4(b), USAlg successfully recovers the unfolded samples. However, in Fig. 4(c) and (d), USAlg fails due to the application of a higher-order differences operator to the noisy original signal. Except for Fig. 4(d), where the oversampling factor is relatively small, USLSE effectively recovers the noisy/noiseless signal when frequencies are either on the grid or not on the grid. LP successfully recovers the original signals in all scenarios with a known autocorrelation function. Although USLSE outperforms USAlg, Prony, and QCQP in the settings of Fig. 4, it should be noted that USAlg can recover signals with unlimited amplitude when the oversampling factor is larger than $2\pi \varepsilon$ in the noiseless case, Prony can be used to deal with nonideal foldings in practical hardware, and QCQP can recover the clean signal from noisy modulo samples.

The running time of all algorithms is also evaluated and the results are shown in Table II. We observe that the execution times for each algorithm remain fairly consistent across various figures in Fig. 4. Additionally, LP, USAlg, and QCQP exhibit the shortest execution times, followed by Prony and USLSE. Furthermore, the average running times of LP, USAlg, and QCQP are all less than 0.01 s, which are significantly shorter than that of Prony ($= 0.233$ s) and USLSE ($= 0.584$ s).

C. Performance Versus SNR

In this section, the performances of USLSE under Markov approximation with $p = 2, 3, 4$ versus SNR are investigated, and USAlg will be used as a comparison. As QCQP method requires a large oversampling factor, and the Prony method is sensitive to the noise, the probabilities of successful recovery of QCQP and Prony are close to 0 in this scenario. In addition, the LP method requires that the initial few unfolded samples of the signal can be obtained and the autocorrelation function is known, which is not the case discussed in this article. Therefore, the results of QCQP, Prony, and LP are not presented in this section. In addition, problem (23) can be viewed as a mixed integer quadratic programming (MIQP); thus, the professional optimization software Gurobi is also used here to solve it for comparison [49]. The maximum running time of Gurobi is set to be the same as the running time of DP+OMP(4) with $p = 4$, which is the longest running time among all other algorithms.

The probabilities of successful recovery versus SNR are shown in Fig. 5. To make the figures more readable, the results of OMP, Gurobi, and USAlg, which are invariant for different $p$, are shown in all the figures for comparison.
First, note that for all figures in Fig. 5 DP+OMP(2) – (4) are almost identical. In Fig. 5(a), when SNR ≤ 20 dB, the probabilities of successful recovery of all algorithms are lower than 0.35. When 22 dB ≤ SNR ≤ 34 dB, DP+OMP(2) – (4) perform best followed by DP+OMP and DP, and all DP-based algorithms outperform USAlg, Gurobi, and OMP. When SNR ≥ 36 dB, USAlg performs best followed by DP+OMP and DP, and all DP-based algorithms provide a better approximation of Q, respectively.

Fig. 6 shows the CCDF of the MSE of frequencies for different algorithms. In Fig. 6(a) where SNR = 22 dB, when CCDF = 0.5 we see that DP-based algorithms perform much better than other algorithms. Furthermore, DP-based algorithms with p = 3 and p = 4 show the same performance and perform better than DP-based algorithms with p = 2. In Fig. 6(b) where SNR = 26 dB, when CCDF = 0.5, we see that DP-based algorithms show the same performance and perform much better than other algorithms. Finally, in Fig. 6(c) where SNR = 30 dB, when CCDF = 0.5, we observe that DP-based algorithms and USAlg show the same performance and perform much better than the other algorithms.

D. Coexistence of Strong and Weak Targets

In this section, we conducted a simulation to demonstrate the performance of USLSE in the scenario where both strong and weak targets coexist. We set K = 2. Let c1 and ω1 be the amplitude and frequency of the strong target, and c2 and ω2 be the amplitude and frequency of the weak target, respectively. To avoid the possibility that the two targets may not be resolved due to the close frequency spacing, we randomly generate the two frequencies such that their wrap-around distance is greater than 5 × 2π/N, where the wrap-around distance of ω1 and ω2 is defined as dist(ω1, ω2) ≜ minω∈Z |ω1 − ω2 + 2πa|. In addition, the weak target is detected when the wrap-around distance of the estimate ˆω and ω2 is less than 2π/N × 0.5, i.e., dist(ˆω, ω2) < 2π/N × 0.5. The detection probability PD is used as a criterion for performance evaluation. In addition, we set N = 512, γ = 10, and λ = 0.5. Furthermore, the magnitude of the weight of the weak target is set to |c2| = 0.1, and the noise is drawn from the complex Gaussian distribution CN(0, 0.005). Let R ≜ (|c1| + |c2|)/λ denote the ratio of the maximum amplitude of the line spectral signal to the ADC threshold, and SWR ≜ 20 log(|c1|/|c2|) denote the ratio of the magnitude of the weight of the strong target to the weak target. It should be noted that the number of maximum folding times is [(R + 1)/2]. In this simulation, alongside the outcomes obtained from various algorithms employing modulo ADC, we also show the results of the traditional ADC with dynamic range being [−λ, λ], i.e., when the amplitude of the signal is within

---

For further details and visualizations, please refer to the original paper.
Fig. 6. CCDF of the MSE of frequencies for $N = 512$, $\gamma = 10$, and $\lambda = 0.7$. Panels (a)–(c) correspond to SNR = 22 dB, SNR = 26 dB, and SNR = 30 dB, respectively.

Fig. 7. Detection probabilities of the weak target versus $R$ or SW$_r$ for $N = 512$, $\gamma = 10$, and $\lambda = 0.5$. Panels (a)–(c) correspond to the second-order ($p = 2$), third-order ($p = 3$), and fourth-order ($p = 4$) Markov approximations of $Q$, respectively. (a) $p = 2$. (b) $p = 3$. (c) $p = 4$.

$[-\lambda, \lambda]$, no distortion incurs, when its amplitude $A$ exceeds $\lambda$, it returns $\text{sign}(A)\lambda$. The NOMP algorithm is employed to process the measurements output by the traditional ADC.

Fig. 7 illustrates the detection probabilities $P_D$ of all algorithms versus $R$ or SW$_r$. To make the figures more readable, the results of traditional ADC, USAlg, Gurobi, and OMP, which are invariant for different $p$ are shown in all figures for comparison. For all cases where $p = 2$, $p = 3$, and $p = 4$, we see that the detection probabilities $P_D$ tend to decrease for all algorithms as $R$ and SW$_r$ increase. For traditional ADC, it exhibits satisfactory performance when SW$_r \leq 25$ dB ($R \leq 3.8$, i.e., the maximum folding times is less than or equal to 2) but almost fails to detect the weak signal effectively when SW$_r \geq 30$ dB ($R \geq 6.5$, i.e., the maximum folding times is greater than or equal to 3). In addition, when SW$_r \leq 34$ dB ($R \leq 10.2$), DP-based algorithms show better performance than other algorithms in the case of $p = 3$ or $p = 4$. Furthermore, when SW$_r > 38$ dB ($R \geq 16.0$) where the SNR is relatively high, USAlg performs best in all figures. We also observe that DP+OMP(3) with $p = 3$ and $p = 4$ perform better than DP+OMP(3) with $p = 2$ when SW$_r \leq 34$ dB ($R \leq 10.2$).

E. Bandlimited Signals Reconstruction

Instead of applying the USLSE on the line spectral signal, our algorithm can also be utilized to recover the bandlimited signal. In this section, we will investigate the performances of USAlg, Gurobi, OMP, and all DP-based algorithms under the Markov approximation with $p = 2, 3, 4$ versus SNR on bandlimited signals. The signal is generated as follows. First, we generate a bandlimited real signal $g_{\Omega}(t)$ with bandwidth $\Omega$ [34]. Then, we transform it into the frequency domain $\tilde{g}_{\Omega}(\omega)$. By setting the amplitude of $\tilde{g}_{\Omega}(\omega)$ to 0 for $-\Omega \leq \omega \leq 0$, $\tilde{g}_{\Omega}(\omega)$ is obtained. Finally, we obtain the complex bandlimited signal in the time domain $g_{\Omega}(t)$ by applying inverse Fourier transformation on $\tilde{g}_{\Omega}(\omega)$. We set $N = 400$ and $\gamma = 10$, and the results are shown in Fig. 8. To make the figures more readable, the results of OMP, Gurobi, and USAlg, which are invariant for different $p$, are shown in all the figures for comparison.

In Fig. 8, OMP can hardly recover the original signal. For all cases, i.e., Fig. 8(a)–(c), the probabilities of successful recovery of USAlg, Gurobi, and all DP-based algorithms increase as the SNR increases. When SNR $\leq 22$ dB, DP+OMP(2)–(4) perform best followed by DP+OMP, DP, Gurobi, and USAlg for $p = 3$ and $p = 4$. When SNR $\geq 24$ dB, DP+OMP(2)–(4) perform best followed by DP + OMP, DP, USAlg, and Gurobi for $p = 2, 3, 4$ and 4. Moreover, when SNR $\geq 30$ dB, all algorithms, except for Gurobi and OMP, achieve perfect successful recovery. Furthermore, the probabilities of the successful recovery of DP-based algorithms increase in our settings as $p$ increases when SNR $\geq 24$ dB. It should be noted that, compared to $p = 3$ and $p = 4$, DP-based algorithms with $p = 2$ benefit more from OMP and iterative refinement.

VI. EMPIRICAL EXPERIMENTS

Empirical experiments are conducted to demonstrate the performance of USLSE by using an AWR1642 radar. The sampling frequency is $f_s = 10$ MHz. The chirp rate is $\kappa = 30 \times 10^{12}$ Hz/s. The maximum radial distance is $r_{\text{max}} = c f_s/(2\kappa)$ = 50m where $c$ is the speed of light. The number of measurements is $N = 256$. Two experiments are conducted. We run the NOMP-CFAR, which can maintain the desired
Fig. 8. Probabilities of successful recovery for bandlimited signals versus SNR with $N = 400$ and $\gamma = 10$. Panels (a)–(c) correspond to the second-order ($p = 2$), third-order ($p = 3$), and fourth-order ($p = 4$) Markov approximations of $Q$, respectively. (a) $p = 2$. (b) $p = 3$. (c) $p = 4$.

Fig. 9. Field experiment setup. (a) There are two targets, referred to as People 1 and People 2. (b) There are three targets, referred to as Corner, People 1, and People 2. Two people can be regarded as weak signals compared to the corner.

Fig. 10. Recovery results of experiment 1 for USLSE, USAlg, and QCQP algorithms. For USLSE, DP+OMP(2) with $p = 4$ is applied here. (a) Real part. (b) Imaginary part.

false alarm rate without knowledge of the noise variance to process the original measurements and output the range estimates as ground truth [50]. The modulo samples are generated in software from the original measurements.

A. Experiment 1

The field experiment 1 is shown in Fig. 9(a), where two people named People 1 and People 2 with radial distances being about 1.8 m and 2.8 m are put in front of the radar. The maximum amplitude of the real part and the imaginary part of the real data is about 1200, and the threshold is set as $\lambda = 250$, which means the maximum folding time is 2. We set $\gamma = 10$ and $\beta = 0.01$. Reconstruction of signals of USLSE, USAlg, and QCQP algorithms are shown in Fig. 10. In this scenario, $\varepsilon$ does not meet the sparsity requirement of the Prony method, leading to its ineffectiveness. For USLSE, we use the fourth-order Markov assumption of $Q$, and the DP+OMP(2) method is utilized to solve the problem (27). For USAlg, the first-order difference is applied as $D = 1$ minimizes both $\|D_\mathcal{R}\{g\}\|_\infty$ and $\|D_\mathcal{I}\{g\}\|_\infty$, where $D_\mathcal{R}\{g\}$ ($D_\mathcal{I}\{g\}$) is the $D$th-order difference of $\mathcal{R}\{g\}$ ($\mathcal{I}\{g\}$) defined in [20]. USAlg and QCQP failed to recover both the real part and the imaginary part while USLSE successfully recovers both the real and imaginary parts. Besides, we use NOMP-CFAR on the original
measurements $g$ to obtain the ranges of the targets, which are used as the so-called ground truth. In addition, we will also apply NOMP-CFAR on the modulo samples $y$ and the recovered samples estimated by USLSE, USAAlg, and QCQP. The range estimation results ranging from 0 m to 4 m are shown in Fig. 11. For the modulo samples $y$ and the signal estimated by USAAlg, no targets are detected by NOMP-CFAR. For QCQP, two weak targets are detected around People 1 and it generates many false alarms. For USLSE, which perfectly recovers the signal, the estimates are the same as the ground truth. As shown in Fig. 11, the range estimate of People 1 is about 1.8 m. The ranges of two detected points correspond to People 2 are 2.8 and 2.9 m. There also exists a leakage component whose range estimate is 0.01 m.

B. Experiment 2

The field experiment consisting of a corner in addition to People 1 and People 2 is shown in Fig. 9(b). The radial distances of Corner, People 1, and People 2 are about 1.8 m, 2.5 m, and 3.3 m, respectively. Compared to the corner, two people can be regarded as weak signals. The maximum amplitude of both the real part and the imaginary part is about 5500, and the threshold is set as $\lambda = 550$, which means that the maximum folding time is 5. We set $\gamma = 10$ and $\beta = 0.08$. Reconstruction of signals of USLSE, USAAlg, and QCQP algorithms are shown in Fig. 12. In this situation, $\epsilon$ also fails to satisfy the sparsity criterion essential for the Prony method, rendering it ineffective. For USLSE, we use the second-order Markov assumption of $Q$ and the DP+OMP(2) method is utilized to solve the problem (27). Similar to experiment 1, for USAAlg, the second-order difference is the optimal choice and is applied here. USAAlg and QCQP failed to recover both the real part and the imaginary part while USLSE almost successfully recovered the signal except for the small head part of the real part of the signal. The LSE results with ranges ranging from 0 to 4.5 m are shown in Fig. 13. For the modulo samples $y$, no targets are detected by NOMP-CFAR. For QCQP, only a minor leakage is output, with no other targets identified. It can be seen that USAAlg only detects the strongest corner and generates many false alarms. While USLSE detects Corner, People 1, and People 2, and the leakage component. Besides, USLSE does not generate false alarms.

VII. CONCLUSION

In this article, we studied LSE via US. The USLSE algorithm is proposed, which exploits the oversampling and the bandlimited property of the line spectrum to estimate the folding instants and folding times and then recover the line spectral signal and perform LSE. For USLSE, we novelly combine DP and OMP to iteratively refine the estimates. The robustness of USLSE compared to the existing approach is demonstrated through numerical experiments. Additionally, we also evaluate the performance of USLSE by processing the data acquired by the mmWave radar and show that USLSE is able to jointly estimate two people and one corner with the oversampling factor of 10 and the folding time is 5.

Several relevant future concerns are worth noting. First, exploring the equivalent conditions for uniquely identifying line spectral signals from modulo samples in future research is worthwhile, as it can serve as a performance benchmark for subsequent algorithms. Second, since the algorithm proposed in this article relies on oversampling, there exists a tradeoff between high sampling rates and increased dynamic range. To enable broader applications of our algorithm, a detailed analysis in this regard is worth investigating. Finally, our proposed algorithm relies on several approximations. Investigating the impact of these approximations on the algorithm’s noise resilience and dynamic range enhancement is worth further exploration from an analytical point of view.

APPENDIX A

PROOF OF PROPOSITION 1 (THE UPPER BOUND OF $V$)

According to the definition of $g[n]$, we have

$$g[n] = g[n+1] - g[n] = \sum_{k=1}^{K} |c_k| e^{i\omega_k} \left( e^{i\omega_k n} - 1 \right)$$

for $n = 0, \ldots, N - 2$, and the upper bound of $|g[n]|$ can be obtained, i.e.,

$$|g[n]| \leq \sum_{k=1}^{K} |c_k| e^{i\omega_k} \left| e^{i\omega_k n} - 1 \right| \leq K c_{\text{max}} \omega_{\text{max}}$$

where $c_{\text{max}} = \max_k |c_k|$ and $\omega_{\text{max}} = \max_k \omega_k$. In addition, we have

$$|\Re\{g[n]\}| \leq K c_{\text{max}} \omega_{\text{max}} \|\Im\{\epsilon\}\|_1 \leq K c_{\text{max}} \omega_{\text{max}}.$$  

Thus, the upper bound of $V$ can be obtained

$$V = \max_{\epsilon} \{\|\Re\{\epsilon\}\|, \|\Im\{\epsilon\}\|\}$$
Let $\tilde{x}[n]$ denote the $(n+1)$th element of $\tilde{x}$. $[\tilde{x}[n]]_0$ can be upper bounded by

$$
[\tilde{x}[n]]_0 = \sum_{m=0}^{N-1} F_{n+1,j_{m+1}} x[m] \\
= \frac{1}{\sqrt{N+1}} \sum_{m=0}^{N-1} e^{-j\frac{2\pi m}{N}} (x[m+1] - x[m]) \\
\leq \frac{1}{\sqrt{N+1}} \sum_{k=1}^{K} c_k e^{-j\frac{2\pi k}{N}} (e^{j\omega_k m} - e^{-j\omega_k m}) \\
= \frac{2}{\sqrt{N+1}} \sum_{k=1}^{K} |c_k| \sin\left(\frac{\omega_k}{2}\right) \sum_{m=0}^{N-2} e^{j\omega_k (m+1)} \\
= \frac{1}{\sqrt{N+1}} \sum_{m=0}^{N-1} e^{j\omega_k m}
$$

where $n = 0, \ldots, N - 2$. Let $\omega_k = \frac{2\pi}{N}(a_k + \delta_k)$, where $a_k \in \{0, 1, 2, \ldots, \frac{N-1}{2}\}$, and $\delta_k \in [-0.5, 0) \cup (0, 0.5)$, $k = 1, \ldots, K$. For $n = 0, \ldots, N - 2$, we have

$$
[\tilde{x}[n]]_0 = \sum_{m=0}^{N-2} e^{j\omega_k m} \\
= \frac{1 - e^{j(N-1)(\omega_k - \frac{2\pi}{N})}}{1 - e^{j(\omega_k - \frac{2\pi}{N})}} = \frac{\sin(\frac{\pi \delta_k}{N})}{\sin(\frac{\pi}{N} - \frac{\pi}{2})}.
$$

Inserting (46) into (45), for $n > \left\lfloor \frac{N-1}{2} \right\rfloor$, one obtains

$$
[\tilde{x}[n]]_0 \leq \frac{2}{\sqrt{N+1}} \sum_{k=1}^{K} |c_k| \sin\left(\frac{\omega_k}{2}\right) \sin\left(\frac{\pi \delta_k}{N} - \frac{\pi}{2}\right) \\
\leq \frac{a}{\sqrt{N+1}} \min\{\frac{2\pi}{N} - \delta_k\} \leq \frac{2K \pi c_{\max} s_{\max} \delta_{\max}}{\sqrt{N-1}} \\
\leq \frac{a}{\sqrt{N-1}} \min\{\frac{2\pi}{N} - \delta_k\} \leq \frac{2K \pi c_{\max} s_{\max} \delta_{\max}}{\sqrt{N-1}}
$$

where $a$ is due to that $\sin(\frac{\pi}{N} - \frac{\pi}{2})$ is greater than 0 and is concave for $\omega_k \in (0, \frac{2\pi}{N})$, $c_{\max} \triangleq \max_{i=1}^{K} |c_i|$, $s_{\max} \triangleq \max_{i=1}^{K} |\delta_i|$. and $\delta_{\max} \triangleq \max_{i=1}^{K} |\delta_i|$. and $\delta_{\max}$

REFERENCES

[1] A. G. Stove, “Linear FMCW radar techniques,” Proc. Inst. Electr. Eng.—F (Radar Signal Process.), vol. 139, no. 5, pp. 343–350, 1992.

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

[3] F. G. Gerolmo and M. Brandt-Pearce, “Detection and estimation of LFMCW radar signals,” IEEE Trans. Aerosp. Electron. Syst., vol. 48, no. 1, pp. 405–418, Jan. 2012.

[4] B. Jin, J. Zhu, Q. Wu, Y. Zhang, and Z. Xu, “One-bit LFMCW radar: Spectrum analysis and target detection,” IEEE Trans. Aerosp. Electron. Syst., vol. 56, no. 4, pp. 2732–2750, Aug. 2020.

[5] R. Zhang et al., “Integrated sensing and communication with massive MIMO: A unified tensor approach for channel and target parameter estimation,” IEEE Trans. Wireless Commun., 2024, to be published, doi: 10.1109/TWC.2024.3351856.

[6] R. Zhang, C. Li, J. Li, and G. Wang, “Range estimation and range-Doppler imaging using signed measurements in LFMCW radar,” IEEE Trans. Aerosp. Electron. Syst., vol. 55, no. 6, pp. 3531–3550, Dec. 2019.

[7] T. Feuillen, M. Alae-Kerahroodi, A. Bhandari, B. S. M. R, and B. Ottersten, “Unlimited sampling for FMCW radars: A proof of concept,” in Proc. IEEE Radar Conf., 2022, pp. 1–5.

[8] F. Engel, P. Heidenreich, A. M. Zoubir, F. K. Jondral, and M. Wintermantel, “Advances in automotive radar: A framework on computationally efficient high-resolution frequency estimation,” IEEE Signal Process. Mag., vol. 34, no. 2, pp. 36–46, Mar. 2017.
[9] B. Vandersmissen et al., “Indoor person identification using a low-power FMCW radar,” IEEE Trans. Geosci. Remote Sens., vol. 56, no. 7, pp. 3941–3952, Jul. 2018.

[10] F. J. Block, “Performance of wideband digital receivers in jamming,” in Proc. IEEE Mil. Commun. Conf., 2006, pp. 1–7.

[11] D. Park, J. Rhee, and Y. Joo, “A wide dynamic-range CMOS image sensor using self-reset technique,” IEEE Electron Device Lett., vol. 28, no. 10, pp. 890–892, Oct. 2007.

[12] J. Yuan, H. Y. Chan, S. W. Fung, and B. Liu, “An activity-triggered 95.3 dB DR, 6 dB THD CMOS imaging sensor with digital calibration,” IEEE J. Solid-State Circuits, vol. 44, no. 10, pp. 2834–2843, Oct. 2009.

[13] K. Sasagawa et al., “An implantable CMOS image sensor with self-reset pixels for functional brain imaging,” IEEE Trans. Electron Devices, vol. 63, no. 1, pp. 215–222, Jan. 2016.

[14] A. Krishna, S. Rudresh, V. Shaw, H. R. Sabella, C. S. Scelamantula, and C. S. Thakur, “Unlimited dynamic range analog-to-digital conversion,” 2019, arXiv:1911.09371.

[15] A. Bhandari and F. Krahmer, “On identifiability in unlimited sampling,” in Proc. Int. Conf. Sampling Theory Appl., 2019, pp. 1–4.

[16] 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.

[17] S. Mulleti and Y. C. Eldar, “Modulo sampling of FRI signals,” IEEE Access, vol. 12, pp. 60369–60384, Apr. 2024.

[18] D. Prasanna, C. Sriram, and C. R. Murthy, “On the identifiability of sparse vectors from modulo compressed sensing measurements,” IEEE Signal Process. Lett., vol. 28, pp. 131–134, Dec. 2021.

[19] Y. Cheng, J. Karlsson, and J. Li, “CRB analysis for Mod-ADC with known folding-count,” in Proc. IEEE Veh. Technol. Conf., 2023, pp. 1–5.

[20] A. Bhandari, F. Krahmer, and R. Raskar, “On unlimited sampling and reconstruction,” IEEE Trans. Signal Process., vol. 69, pp. 3827–3839, Dec. 2021.

[21] A. Bhandari, F. Krahmer, and R. Raskar, “Unlimited sampling of sparse sinusoidal mixtures,” in Proc. Int. Conf. Img. Inf. Theory, 2018, pp. 336–340.

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

[23] V. Boui, F. Krahmer, and A. Bhandari, “Multidimensional unlimited sampling: A geometrical perspective,” in Proc. Eur. Signal Process. Conf., 2021, pp. 2314–2318.

[24] S. Fernández-Menduíña, F. Krahmer, G. Leus, and A. Bhandari, “DoA estimation via unlimited sensing,” in Proc. Eur. Signal Process. Conf., 2021, pp. 1866–1870.

[25] A. Bhandari, M. Beckmann, and F. Krahmer, “The modulo Radon transform and its inversion,” in Proc. Eur. Signal Process. Conf., 2021, pp. 770–774.

[26] S. Fernández-Menduíña, F. Krahmer, G. Leus, and A. Bhandari, “Computational array signal processing via modulo non-linearities,” IEEE Trans. Signal Process., vol. 70, pp. 2168–2179, Mar. 2022.

[27] M. Beckmann, F. Krahmer, and A. Bhandari, “HDOM tomography via modulo Radon transform,” in Proc. IEEE Int. Conf. Image Process., 2020, pp. 3025–3029.

[28] A. Bhandari and F. Krahmer, “HDOM imaging from quantization noise,” in Proc. IEEE Int. Conf. Image Process., 2020, pp. 101–105.

[29] O. Ordentlich and U. Erez, “Integer-forcing source coding,” IEEE Trans. Inf. Theory, vol. 63, no. 2, pp. 1253–1269, Feb. 2017.

[30] O. Ordentlich, G. Tabak, P. K. Hanumolu, A. C. Singer, and G. W. Wornell, “A modulo-based architecture for analog-to-digital conversion,” IEEE J. Sel. Topics Signal Process., vol. 12, no. 5, pp. 825–840, Oct. 2018.

[31] E. Romanov and O. Ordentlich, “Blind unwrapping of modulo reduced Gaussian vectors: Recovering MSBs from LSBs,” IEEE Trans. Inf. Theory, vol. 67, no. 3, pp. 1897–1919, Mar. 2021.
Jiang Zhu (Senior Member, IEEE) received the B.E. degree in electronic science and technology from Harbin Engineering University, Harbin, China, in 2011, and the Ph.D. degree in information and communication engineering from Tsinghua University, Beijing, China, in 2016.

He was a Visiting Student with Lehigh University, Bethlehem, PA, USA, in 2015. He joined Ocean College, Zhejiang University, Hangzhou, China, in 2016, and is currently an Associate Professor. His current research interests include statistical signal processing, compressed sensing, and their applications.

Fengzhong Qu (Senior Member, IEEE) received the B.E. degree in information engineering and M.E. degree in communication engineering from Zhejiang University, Hangzhou, China, in 2002 and 2005, respectively, and the Ph.D. degree from the Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL, USA, in 2009.

From 2009 to 2010, he was an Adjunct Research Scholar with the Department of Electrical and Computer Engineering, University of Florida. Since 2011, he has been with Ocean College, Zhejiang University, Zhoushan, China, where he is currently a Full-Time Professor and the Vice-Dean. His current research interests include underwater acoustic communications and networking, and seafloor observatories.

Dr. Qu has been actively serving the technical community by organizing many IEEE international conferences and serving on the editorial boards of a number of journals, including IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS, IEEE TRANSACTIONS ON INTELLIGENT VEHICLES, IET Communications, and China Communications.

De Wen Soh received the B.S. degree in mathematics from Stanford University, Stanford, CA, USA, in 2009, and the Ph.D. degree in electrical engineering from Yale University, New Haven, CT, USA, in 2016, where he worked on high-dimensional graphical model learning under the supervision of Sekhar Tatikonda.

He is currently an Assistant Professor with the Singapore University of Technology and Design, Singapore. His research interests include graphical model estimation, graph signal processing, network analytics, transport modeling, high-dimensional statistical theory, and artificial intelligence.