A 1.5-bit Quantization Scheme and Its Application to Direction Estimation

Xicheng Lu, Wei Liu, *Senior Member, IEEE*, and Akram Alomainy, *Senior Member, IEEE*

***Abstract*—In massive multiple-input multiple-output (MIMO) systems, the balance between cost and performance has made low-bit, especially 1-bit, analog-to-digital converters (ADCs) an indispensable part of the solution. To address the severe quantization errors inherent in such quantization schemes, in this paper, a novel 1.5-bit ADC quantization scheme is proposed. Its information entropy varies with the user-defined threshold but remains bounded by $\log_2 3$, requiring only one additional comparator beyond a 1-bit quantizer. This three-level design effectively reduces the large quantization loss incurred by 1-bit ADCs. Leveraging the Price theorem and Mehler's formula, we derive the 1.5-bit correlation estimator and analyze the approximation error using a first-order Taylor expansion. Our findings reveal that, at low signal-to-noise ratios (SNRs), the eigenvalues of the 1.5-bit covariance matrix are nearly identical to those of the unquantized covariance matrix. This allows direct parameter estimation without the need to reconstruct the unquantized covariance. Moreover, we show that the approximation error for 1.5-bit measurements is much smaller than that of 1-bit quantization in high SNR conditions. Based on the derived correlation estimator, an algorithm is proposed for recovering the unquantized covariance matrix using a gradient descent method. Simulation results obtained by applying our proposed algorithm to DOA estimation show that, the 1.5-bit scheme is robust to the choice of the threshold value, and significantly outperforms 1-bit quantization without much increase in cost.**

***Index Terms*—1-bit quantization, DOA estimation, 1.5-bit quantization, low-bit quantization.**

I. Introduction

Direction of arrival (DOA) estimation is a key research topic in array signal processing, with widespread applications in radar, sonar, and wireless communications [1]–[4]. In particular, large-scale antenna array systems such as massive MIMO have become a focus of research in the community [5], which provide an opportunity for achieving a much improved DOA estimation performance; however, if implemented in a fully digital manner, significant increase in computational complexity is inevitable. To mitigate concerns about hardware cost, power consumption, and real-time processing capabilities for a large-scale antenna array, low-bit analogue to digital converters (ADCs) have gained popularity in such applications owing to their low-cost and complexity.

In low-bit quantizer design, the zero-threshold 1-bit quantizer is widely known for its simple hardware structure and minimal cost, requiring only a zero-threshold comparator to preserve signs of the sampled data. Early research on 1-bit systems focused on topics such as Bussgang's theorem [6], the arcsine law [7], [8], and frequency estimation using 1-bit samples [9], [10], demonstrating that the 1-bit correlation results can be used to reconstruct the full-precision correlation result under the zero-mean Gaussian assumption, so that satisfactory DOA estimation performance can still be achieved using 1-bit samples. However, the power information is lost in the zero-threshold quantizer, while in many applications such as localization in wireless sensor networks, the power information is crucial to the success of some developed algorithms. As the result, non-zero threshold 1-bit quantizer capable of recovering both power and correlation information of the unquantized data was then developed [11]. Recently, several mixed-ADC architectures have been proposed by combining one-bit and high-resolution ADCs to mitigate the severe quantization loss inherent in purely one-bit systems [12], [13].

For DOA estimation under 1-bit quantization, the arcsine law plays a central role. In [14], a time-varying threshold was introduced for 1-bit quantization, and the arcsine law was numerically modified to reduce quantization errors. Since practical signals are often non-stationary, the modified arcsine law was further extended in [15] with 1-bit time-varying thresholds to cover non-stationary signals. Meanwhile, the performance of the 1-bit non-zero threshold quantizer proposed in [11] was analyzed in [16] for Gaussian stationary signals. In addition, it was shown in [17] that the arcsine law remains valid for complex elliptically symmetric (CES) distributions, allowing heavy-tailed distributions (e.g., complex Cauchy, complex T-distribution) to follow the arcsine law.

Along the development of 1-bit quantizers, under sub-Gaussian distributions, a two-bit quantizer was introduced in [18], with two distinct quantizations performed on each sample using independent dither signals, enabling full covariance information recovery. Moreover, an adaptive threshold scheme was proposed to optimize the involved dithering parameters [19].Besides that, recent theoretical advances [20] show that dithering-based quantization with outlier truncation achieves near-minimax optimal recovery for quantized compressive sensing and covariance estimation, even under heavy-tailed or non-Gaussian models. It is demonstrated in [21] that leveraging Toeplitz structure enables bit-efficient covariance estimation. For the 2-bit case, typically its covariance estimation relies on algorithms that iteratively reconstruct the unquantized covariance matrix under sign constraints [22]. Earlier work [23] analyzed 2-bit correlation estimators based on numerical double integration, but did not provide a closed-

This work is supported by The Hong Kong Polytechnic University Start-Up Fund under Project P0053642 (corresponding author: Wei Liu).

X. Lu and A. Alomainy are with the School of Electronic Engineering and Computer Science, Queen Mary University of London, London, UK (e-mail:xicheng.lu@qmul.ac.uk, a.alomainy@qmul.ac.uk); W. Liu is with the Department of Electrical and Electronic Engineering, The Hong Kong Polytechnic University, Hong Kong, (email: wei2.liu@polyu.edu.hk)

form estimator analogous to the arcsine law for 1-bit quantization. Therefore, multi-bit quantization methods, while offering more information, often entail higher computational complexity and lack simple closed-form solutions for covariance recovery.

Compressive sensing (CS) has also been adapted for 1-bit measurements using several strategies. Early zero-threshold methods [24]–[26] use simple sign detectors, but they suffer from amplitude ambiguity. Non-zero thresholds are employed to embed extra information for norm recovery and noise robustness, although they are sensitive to threshold selection [27]. A Bayesian framework via generalized approximate message passing (GAMP) leverages separable priors to extend applicability beyond sparse signals in [28]. Time-varying thresholds [29] and dithering-based methods [30] offer dynamic adaptation and compatibility with structured matrices, albeit at increased computational cost.

In this paper, to address the problem of excessive quantization error associated with conventional 1-bit ADCs, without significantly increasing the cost in terms of storage requirement and sampling speed, a 1.5-bit quantization framework is developed. Specifically, the proposed scheme employs two symmetric nonzero thresholds, partitioning the signal amplitude into three levels, thereby capturing more detailed information than the traditional 1-bit approach. This design significantly reduces quantization error while maintaining low hardware complexity. Notably, in contrast to the typical 2-bit quantizer, the 1.5-bit scheme utilizes only two symmetric nonzero thresholds, thereby requiring one less comparator. When the quantization thresholds are optimally chosen, the 1.5-bit quantizer provides substantial performance improvement over the conventional 1-bit approach, while achieving estimation accuracy that is very close to that of the 2-bit quantizer, but with reduced implementation complexity. Moreover, unlike our previous conference publication [31] based on the empirical observation that the correlation estimator exhibits an arcsine-like relationship as in 1-bit quantization, this work rigorously derives a closed-form expression for the 1.5-bit correlation estimator, establishing a solid theoretical foundation for the proposed method.

Our main contributions are summarized as follows:

- **Closed-form correlation estimator for the 1.5-bit and multi-bit quantizers:** Building on the 1.5-bit quantizer proposed in our conference publication [31], we leverage Price's theorem [32], Mehler's formula [33], and Hermite polynomials [34] to derive a closed-form polynomial-form correlation estimator. This theoretical framework is further extended to the general multi-bit quantization case, allowing for a unified closed-form correlation estimator for arbitrary symmetric multi-bit quantizers. We also point out that, in contrast to [23], which numerically evaluates the 2-bit correlation estimator via double integration, our results provide explicit analytical formulas.
- **Feasibility and practical advantages of direct use of multi-bit sample covariance:** It is rigorously shown that the covariance matrix obtained from 1.5-bit measurements can be employed directly for DOA estimation under low-SNR scenarios, and for high-SNR cases its performance remains almost indistinguishable from that based on the reconstructed unquantized covariance, however, for the classic 1-bit scheme, the performance of this method degrades noticeably [35]. Importantly, this property is also generalized to the multi-bit setting, where the normalized unquantized covariance can be reliably estimated using the multi-bit sample covariance, thus bypassing the need for computationally expensive polynomial inversion in practical applications.
- **Gradient-based covariance reconstruction, order selection, robust thresholding, and comparative analysis:** Building on our derived correlation estimator, a gradient-descent approach is developed for reconstructing the unquantized normalized covariance matrix. In parallel, we determine through simulations a suitable polynomial order for truncated expansions, balancing computational overhead and estimation accuracy. Furthermore, we empirically validate the robustness of threshold selection in the 1.5-bit scheme and provide a practical reference threshold. We also present a reference threshold for the 2-bit case, and conduct extensive comparisons with 1-bit and 2-bit methods. The results demonstrate that direct DOA estimation using the 1.5-bit and 2-bit sample covariance matrices achieves comparable accuracy to that of the reconstructed unquantized version, while the 1.5-bit quantizer, requiring only one additional comparator than the 1-bit case, delivers substantial performance gains. In contrast, the performance improvement from 1.5-bit to 2-bit quantization is marginal, despite its much higher hardware complexity.

The paper is organized as follows. The 1-bit quantization scheme is briefly reviewed in Section II. The 1.5-bit quantization design and its closed-form correlation estimator are presented in Section III. Its application to DOA estimation is detailed in Section IV. Simulations are performed in Section V and Section VI concludes the paper.

II. Low-Bit Quantization Schemes

Consider a general b-bit quantizer that partitions the real axis $\mathbb{R}$ into $A = 2^b$ adjacent intervals:

$$I_k = [t_{k-1}, t_k), \quad k = 1, 2, \ldots, A, \tag{1}$$

where $t_0 = -\infty$, $t_A = +\infty$, and $t_k \in \mathbb{R}$ for $k = 1, \ldots, A-1$. When an input p falls into the interval I_k, the quantizer outputs a corresponding discrete level z_k:

$$z_k = Q_b(p) \quad \text{if } p \in I_k, \tag{2}$$

where $Q_b(\cdot)$ denotes the quantization operation.

As a special case, consider the 1-bit quantizer ($b = 1$) with a threshold α: $I_1 = (-\infty, \alpha)$ and $I_2 = [\alpha, +\infty)$. For the classical sign quantizer, set $\alpha = 0$:

$$Q_1(p) = \begin{cases} z_1 = -1, & p < 0, \\ z_2 = +1, & p \geq 0. \end{cases} \tag{3}$$

Consider two zero-mean jointly Gaussian random variables x_1 and x_2 with variance σ_x^2, and define their correlation as $r_x = \mathbb{E}[x_1 x_2]$.

Applying the 1-bit quantizer yields $y_i = \mathrm{sgn}(x_i)$ for $i = 1, 2$. The correlation r_y between y_1 and y_2 is then related to r_x via the well-known arcsine law [7]: $r_y = \frac{2}{\pi} \arcsin\left(\frac{r_x}{\sigma_x^2}\right)$, where σ_x^2 is the signal variance.

For the 2-bit case ($b = 2$), the real axis is partitioned into four intervals by three thresholds, typically chosen as $-\Lambda$, 0, and $+\Lambda$ (with $\Lambda = \kappa\sigma_x$, $\kappa > 0$): $I_1 = (-\infty, -\Lambda)$, $I_2 = [-\Lambda, 0)$, $I_3 = [0, \Lambda)$, and $I_4 = [\Lambda, +\infty)$.

The quantization rule is then

$$Q_2(x, \Lambda) = \begin{cases} z_1 = -3, & x < -\Lambda, \\ z_2 = -1, & -\Lambda \le x < 0, \\ z_3 = +1, & 0 \le x < \Lambda, \\ z_4 = +3, & x \ge \Lambda. \end{cases} \tag{4}$$

The 2-bit rule and the following proposed 1.5-bit quantizer share the same outer decision thresholds $\pm\Lambda$. In the 2-bit scheme, an additional threshold at the origin yields four output levels $-3, -1, +1, +3$. The 1.5-bit design removes this inner threshold: all samples between $\pm\Lambda$ are mapped to the central output level 0, producing the three-level alphabet $-1, 0, +1$. We next detail this 1.5-bit scheme.

A. Proposed 1.5-bit Quantization Scheme

The 1.5-bit quantizer employs the same pair of symmetric thresholds $\pm\Lambda$ as the 2-bit case, but omits the central zero threshold:

$$Q_{1.5}(x, \Lambda) = \begin{cases} +1, & x > \Lambda, \\ 0, & -\Lambda \le x \le \Lambda, \\ -1, & x < -\Lambda. \end{cases} \tag{5}$$

Note. Unless explicitly stated otherwise, we assume throughout this paper that the variance σ_x^2 of each unquantized signal is known.

Here "1.5-bit" is just a nominal term, and clearly the actual information entropy of the quantizer output is affected by the threshold; hence the maximum information entropy of a 1.5-bit quantizer is $log_2 3$. Essentially, this refers to a three-level quantization scheme that lies between the traditional 1-bit (two-level) and 2-bit (four-level) schemes, and represents the most hardware-efficient extension beyond 1-bit, while its practical performance in DOA estimation remains largely unexplored. Our work here will try to clarify how much performance can be gained by moving from 1-bit to 1.5-bit quantization, and how close this comes to the 2-bit (4-level) case.

III. 1.5-bit Quantization and Correlation Estimation

A. Derivation of 1.5-bit Quantization Correlation Estimator

Consider two arbitrary zero-mean Gaussian signals, $x_1(t)$ and $x_2(t)$, with standard deviations σ_{x_1} and σ_{x_2}, respectively, and cross-correlation $r_x = \mathbb{E}[x_1(t)x_2(t)]$. Their normalized correlation coefficient is defined as $\rho_x = \frac{\mathbb{E}[x_1(t)x_2(t)]}{\sigma_{x_1}\sigma_{x_2}}$.

After the quantization operation, the quantized signals are given by $y_1(t) = \mathcal{Q}(x_1(t), \Lambda_1)$ and $y_2(t) = \mathcal{Q}(x_2(t), \Lambda_2)$, with quantization thresholds defined as $\Lambda_1 = \kappa\sigma_{x_1}$ and $\Lambda_2 = \kappa\sigma_{x_2}$.

The objective is to derive a mathematical relationship between the cross-correlation of the quantized signals $\mathbb{E}[y_1(t)y_2(t)]$, and the original normalized correlation coefficient ρ_x. This derivation is crucial because it enables us to recover the original covariance matrix from quantized measurements, which is required for applications such as DOA estimation. To facilitate this derivation, we make use of the following Price's theorem [32].

Theorem 1. *To jointly Gaussian random variables x_1 and x_2 with joint probability density function $p(x_1, x_2)$ and for a function $g(x_1, x_2)$ that vanishes at infinity, the nth derivative of $\mathbb{E}[g(x_1, x_2)]$ with respect to the correlation coefficient ρ is given by*

$$\frac{\partial^n}{\partial\rho^n}\mathbb{E}[g(x_1, x_2)] = \mathbb{E}\left[\frac{\partial^{2n} g(x_1, x_2)}{\partial x_1^n \partial x_2^n}\right]. \tag{6}$$

To simplify the notation in the derivation process, we drop the explicit time dependence and denote $x_1(t)$ and $x_2(t)$ simply by x_1 and x_2, and similarly $y_1(t)$ and $y_2(t)$ by y_1 and y_2, and the normalized correlation coefficient ρ_x of x_1 and x_2 as ρ. In this context, we set $g(x_1, x_2) = y_1 y_2 = \mathcal{Q}(x_1, \Lambda_1)\,\mathcal{Q}(x_2, \Lambda_2)$, Then, the Price's theorem can be applied to derive the 1.5-bit correlation estimator $\mathbb{E}[y_1 y_2]$. Using Theorem 1, we have:

$$\begin{aligned} \frac{\partial\mathbb{E}[y_1 y_2]}{\partial\rho} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} &\left(\delta\left(x_1 - \frac{\Lambda_1}{\sigma_{x_1}}\right) + \delta\left(x_1 + \frac{\Lambda_1}{\sigma_{x_1}}\right)\right) \\ &\cdot \left(\delta\left(x_2 - \frac{\Lambda_2}{\sigma_{x_2}}\right) + \delta\left(x_2 + \frac{\Lambda_2}{\sigma_{x_2}}\right)\right) \\ &\cdot p(x_1, x_2)\, dx_1\, dx_2, \end{aligned} \tag{7}$$

where $\delta(\cdot)$ is the Dirac delta function. It can be rewritten as:

$$\begin{aligned} \frac{\partial\mathbb{E}[y_1 y_2]}{\partial\rho} &= 2p\left(\frac{\Lambda_1}{\sigma_{x_1}}, \frac{\Lambda_2}{\sigma_{x_2}}\right) + 2p\left(\frac{\Lambda_1}{\sigma_{x_1}}, -\frac{\Lambda_2}{\sigma_{x_2}}\right) \\ &= 2p(\kappa, \kappa) + 2p(\kappa, -\kappa), \end{aligned} \tag{8}$$

where

$$\begin{aligned} p(\kappa, \kappa) = &\frac{1}{2\pi\sigma_{x_1}\sigma_{x_2}\sqrt{1-\rho^2}} \\ &\exp\left(-\frac{1}{2(1-\rho^2)}\left[\frac{\kappa^2}{\sigma_{x_1}^2} + \frac{\kappa^2}{\sigma_{x_2}^2} - \frac{2\rho\kappa^2}{\sigma_{x_1}\sigma_{x_2}}\right]\right), \end{aligned} \tag{9}$$

$$\begin{aligned} p(\kappa, -\kappa) = &\frac{1}{2\pi\sigma_{x_1}\sigma_{x_2}\sqrt{1-\rho^2}} \\ &\exp\left(-\frac{1}{2(1-\rho^2)}\left[\frac{\kappa^2}{\sigma_{x_1}^2} + \frac{\kappa^2}{\sigma_{x_2}^2} + \frac{2\rho\kappa^2}{\sigma_{x_1}\sigma_{x_2}}\right]\right). \end{aligned} \tag{10}$$

The right-hand side of (8) can be simplified according to Mehler's formula [33], which is:

$$\frac{1}{\sqrt{1-\rho^2}} e^{-\frac{\rho^2\left(u^2+v^2\right)-2\rho uv}{1-\rho^2}} = \sum_{n=0}^{\infty} \frac{\rho^n}{2^n n!} H_n(u)\, H_n(v) \tag{11}$$

where $H_n(u)$ is a Hermite polynomial that can be generated by Rodrigues' formula [36]

$$H_n(u) = (-1)^n e^{u^2} \frac{d^n}{du^n} e^{-u^2}. \tag{12}$$

After integrating ρ on both sides of (8), we have:

$$\mathbb{E}[y_1y_2] = \frac{1}{\pi}e^{-\kappa^2}\left[\sum_{n=0}^{\infty}\frac{\rho^{n+1}}{2^n(n+1)!}H_n^2\left(\frac{\kappa}{\sqrt{2}}\right) + \sum_{n=0}^{\infty}\frac{\rho^{n+1}}{2^n(n+1)!}H_n\left(\frac{\kappa}{\sqrt{2}}\right)H_n\left(\frac{\kappa}{-\sqrt{2}}\right)\right]. \tag{13}$$

Lemma 1. For any integer $n \geq 0$, the Hermite polynomial $H_n(x)$ satisfies the parity property $H_n(-x) = (-1)^n H_n(x)$. In particular, $H_n(x)$ is an even function if n is even and an odd function if n is odd.

The proof of Lemma 1 is provided in Appendix A.

Based on Lemma 1, (13) can be simplified to:

$$\mathbb{E}[y_1y_2] = \frac{2}{\pi}e^{-\kappa^2}\sum_{n=0}^{\infty}\frac{\rho^{2n+1}}{2^{2n}(2n+1)!}H_{2n}^2\left(\frac{\kappa}{\sqrt{2}}\right). \tag{14}$$

To transform (14) to a more general form, we need to represent $H_{2n}^2\left(\frac{\kappa}{\sqrt{2}}\right)$ in a generic expression. The power series of e^{-x^2} is $e^{-x^2} = \sum_{k=0}^{\infty}\frac{(-1)^k x^{2k}}{k!}$. Taking the $2n$th derivative of the function x^{2k} with respect to x, we have $\frac{d^{2n}}{dx^{2n}}x^{2k} = \frac{(2k)!}{(2k-2n)!}x^{2k-2n}$. Thus, the Hermite polynomial H_{2n} for $x = \frac{\kappa}{\sqrt{2}}$ is:

$$H_{2n}\left(\frac{\kappa}{\sqrt{2}}\right) = (-1)^{2n}e^{(\frac{\kappa}{\sqrt{2}})^2}\sum_{k=n}^{\infty}\frac{(-1)^k(2k)!}{k!(2k-2n)!}\left(\frac{\kappa}{\sqrt{2}}\right)^{2k-2n} \tag{15}$$

Finally, the cross-correlation $\mathbb{E}[y_1y_2]$ with ρ is obtained:

$$\begin{aligned}\mathbb{E}[y_1y_2] &= \frac{2}{\pi}e^{-\kappa^2}\sum_{n=0}^{\infty}\frac{\rho^{2n+1}}{2^{2n}(2n+1)!}H_{2n}^2\left(\frac{\kappa}{\sqrt{2}}\right)\\ &= \frac{2}{\pi}e^{-\kappa^2}\sum_{n=0}^{\infty}\frac{\rho^{2n+1}}{2^{2n}(2n+1)!}\\ &\quad\sum_{k=n}^{\infty}\sum_{j=n}^{\infty}\frac{(-1)^{k+j}(2k)!(2j)!}{k!j!(2k-2n)!(2j-2n)!}\left(\frac{\kappa}{\sqrt{2}}\right)^{2k+2j-4n}\end{aligned} \tag{16}$$

When $\kappa = 0$, i.e., the 1-bit zero-threshold case, based on Lemma 1, we see that when n is odd, $H_n(0) = 0$, and when n is even, $H_{2n}(0) = (-1)^n\frac{(2n)!}{2^n(n)!}$; hence:

$$\begin{aligned}\mathbb{E}[y_1y_2] &= \frac{2}{\pi}\left[\sum_{k=0}^{\infty}\frac{\rho^{2k+1}}{2^{2k}(2k+1)!}H_{2k}(0)^2\right]\\ &= \frac{2}{\pi}\sum_{k=0}^{\infty}\frac{(2k)!}{4^k(k!)^2(2k+1)}\rho^{2k+1},\end{aligned} \tag{17}$$

where $\sum_{k=0}^{\infty}\frac{(2k)!}{4^k(k!)^2(2k+1)}\rho^{2k+1}$ is the Taylor series of $arcsine(\rho)$. Thus, we obtain the well-known arcsine law $\mathbb{E}[y_1y_2] = \frac{2}{\pi}\arcsin(\rho)$ for 1-bit quantization with a zero threshold. Although this result was established more than half a century ago by Van Vleck [7] using geometric integration arguments, the above derivation starts from the Price-Hermite expansion of 1.5-bit. Therefore, this perspective connects the classical 1-bit result with 1.5-bit quantizer via the zero-threshold limit. And further based on the work of [35], the series expression of the arcsine law played a key role in the subsequent comparison of the approximation error at 1.5-bit in Lemma 5.

To ensure accuracy for the nth order polynomial solutions, we next prove convergence of (16) under an arbitrary κ.

Lemma 2. Define (16) as: $f(\rho) = \sum_{n=0}^{\infty} a_{2n+1}\,\rho^{2n+1}$, where the coefficient is :

$$\begin{aligned}a_{2n+1} &= \frac{1}{\pi}\,e^{-\kappa^2}\,\frac{1}{2^{2n}\,(2n+1)!}\\ &\sum_{k=n}^{\infty}\sum_{j=n}^{\infty}\frac{(-1)^{k+j}\,(2k)!\,(2j)!}{k!\,j!\,(2k-2n)!\,(2j-2n)!}\left(\tfrac{\kappa}{\sqrt{2}}\right)^{2k+2j-4n}.\end{aligned} \tag{18}$$

Then, for $|\rho| \leq 1$, the series $\sum_{n=0}^{\infty} a_{2n+1}\,\rho^{2n+1}$ converges absolutely.

The proof is provided in Appendix B, and it will be used in the subsequent feasibility proof for recovering the unquantized correlation with a correlation estimator using a gradient descent method.

Remark 1. *Since the full closed-form expression of* (16) *is too complex, for simplicity in the following analysis, unless otherwise stated, we use the form of* $\mathbb{E}[y_1y_2] = \frac{2}{\pi}e^{-\kappa^2}\sum_{n=0}^{\infty}\frac{\rho^{2n+1}}{2^{2n}(2n+1)!}H_{2n}^2\left(\frac{\kappa}{\sqrt{2}}\right)$.

B. Derivation of the Correlation Estimator for General b-bit Quantization

Although Price-Hermite-based frameworks for correlation estimator of 1-bit non-zero threshold quantization have been investigated in recent works [11], [16], to the best of our knowledge, a unified explicit expression for the general multi-bit case has not yet been presented in literature, and the recent multi-bit based DOA covariance recovery problem is mainly solved by constructing optimization and sign constraint problems [22]. In this subsection, we provide a derivation of the correlation estimator for arbitrary b-bit quantization.

After the b-bit quantization process of x_1 and x_2, the outputs are $y_1 = Q_b(x_1)$, and $y_2 = Q_b(x_2)$, where $Q_b(\cdot)$ denotes a generic b-bit quantizer (see Section II), specified by output levels $\{z_k\}_{k=1}^{A}$ and thresholds $\{t_k\}_{k=1}^{A-1}$. For the first signal x_1, define normalized thresholds $u_k = t_k/(\sqrt{2}\sigma_{x_1})$, $(k = 1,\ldots,A-1)$; for the second signal x_2, similarly define levels $\{z_m\}_{m=1}^{A}$, thresholds $\{t_m\}_{m=1}^{A-1}$, and normalized thresholds $v_m = t_m/(\sqrt{2}\sigma_{x_2})$, $(m = 1,\ldots,A-1)$.

Our goal is to characterize the relationship between the quantized correlation $\mathbb{E}[y_1y_2]$ and the underlying coefficient ρ for arbitrary b-bit quantization.

Applying Theorem 1 and exploiting the piecewise-constant nature of $Q_b(\cdot)$, we have

$$\frac{\partial\mathbb{E}[y_1y_2]}{\partial\rho} = \mathbb{E}\left[\frac{\partial^2}{\partial x_1\partial x_2}\left(Q_b(x_1)Q_b(x_2)\right)\right]. \tag{19}$$

This yields a sum of Dirac delta functions at the quantizer thresholds, weighted by output level differences $\Delta z_k = z_{k+1} - z_k$, giving

$$\frac{\partial\mathbb{E}[y_1y_2]}{\partial\rho} = \sum_{k=1}^{A-1}\sum_{m=1}^{A-1}\Delta z_k\Delta z_m\; p(u_k, v_m). \tag{20}$$

By expanding $p(u_k, v_m)$ using Mehler's formula [33], we obtain:

$$\frac{\partial \mathbb{E}[y_1 y_2]}{\partial \rho} = \frac{1}{2\pi} \sum_{n=0}^{\infty} \frac{\rho^n}{2^n n!} A_n, \tag{21}$$

where

$$A_n = \sum_{k=1}^{A-1} \sum_{m=1}^{A-1} \Delta z_k \Delta z_m \exp(-u_k^2 - v_m^2) H_n(u_k) H_n(v_m). \tag{22}$$

Integrating both sides with respect to ρ, and noting that for symmetric quantizers $\mathbb{E}[y_1 y_2]\big|_{\rho=0} = 0$, we finally obtain the general b-bit quantization mapping:

$$\mathbb{E}[y_1 y_2] = \frac{1}{2\pi} \sum_{n=0}^{\infty} \frac{A_n}{2^n n!(n+1)} \rho^{n+1}. \tag{23}$$

IV. 1.5-bit quantization scheme for DOA estimation

A. General Array Signal Model

Consider K narrowband far-field sources, denoted as s_k for $k = 1, 2, \ldots, K$, impinging on an M-sensor linear array from directions $\theta_k \in [-90°, 90°]$. The sensor locations are defined by the set $\mathbb{D}$, where each sensor is positioned at $\frac{r_i \lambda}{2}$ with $r_i \in \mathbb{Z}$, $i = 1, 2, \ldots, M$, and λ being the wavelength of the signals.

The received signal at the sensors can be expressed as:

$$\mathbf{x}(t) = \mathbf{A}(\theta)\mathbf{s}(t) + \mathbf{n}(t), \tag{24}$$

where $\mathbf{x}(t) \in \mathbb{C}^{M\times 1}$ denotes the observed signal vector at time t, $\mathbf{s}(t) \in \mathbb{C}^{K\times 1}$ represents the source signal vector, and $\mathbf{n}(t) \in \mathbb{C}^{M\times 1}$ the additive noise vector. The steering matrix is $\mathbf{A}(\theta) = [\mathbf{a}(\theta_1), \mathbf{a}(\theta_2), \ldots, \mathbf{a}(\theta_K)]$, with $\mathbf{a}(\theta_i) = [e^{-jr_1\pi\sin(\theta_i)}, e^{-jr_2\pi\sin(\theta_i)}, \ldots, e^{-jr_M\pi\sin(\theta_i)}]^T$.

The covariance matrix of the received signals is:

$$\mathbf{R}_{\mathbf{x}} = \mathbb{E}\{\mathbf{x}(t)\mathbf{x}^H(t)\} = \sum_{i=1}^{K} \sigma_i^2 \mathbf{a}(\theta_i)\mathbf{a}^H(\theta_i) + \sigma_n^2 \mathbf{I}_M, \tag{25}$$

where σ_i^2 denotes the power of the ith source signal $s_i(t)$, $i = 1, 2, \ldots, K$, and σ_n^2 denotes the power of the additive noise $n(t)$ at each sensor.

The normalized correlation coefficient ρ_{jk} between signals received at sensors j and k can be expressed as:

$$\rho_{jk} = \Re(\rho_{jk}) + j\Im(\rho_{jk}) = \frac{[\mathbf{R}_{\mathbf{x}}]_{jk}}{\sqrt{[\mathbf{R}_{\mathbf{x}}]_{jj}[\mathbf{R}_{\mathbf{x}}]_{kk}}}. \tag{26}$$

When $j = k$, the normalized correlation coefficient $\rho_{jk} = 1$.

According to [35], if the power of sources is equal, i.e., $\sigma_1^2 = \sigma_2^2 = \cdots = \sigma_k^2 = \sigma_s^2$ and define $\xi = \frac{\sigma_s^2}{\sigma_n^2}$, the correlation ρ_{jk} for signals at the jth and kth sensors is:

$$\rho_{jk} = \frac{\sum_{i=1}^{K} \boldsymbol{a}_j(\theta_i)\boldsymbol{a}_k^*(\theta_i)}{K + \frac{1}{\xi}}. \tag{27}$$

Observing (27), it is easy to see that ρ_{jk} increases as the SNR increases (i.e., the source power increases given a fixed noise power) or as the number of sources K decreases.

B. DOA Estimation Based on Normalized Covariance Reconstruction Using the Gradient Descent Method

When applying the 1.5-bit quantizer to received array signal $\mathbf{x}(t)$, the qth element output of $\mathbf{y}(t)$ is:

$$y_q(t) = \frac{1}{\sqrt{2}}(\mathcal{Q}(\Re\{x_q(t)\}, \Lambda) + j\mathcal{Q}(\Im\{x_q(t)\}, \Lambda)), \tag{28}$$

where $\Re\{x_q(t)\}$ and $\Im\{x_q(t)\}$ denote the real and imaginary parts of the qth element $x_q(t)$of $\mathbf{x}(t)$, respectively. The factor $\frac{1}{\sqrt{2}}$ is used to normalize the power. The $\mathcal{Q}$ function outputs $+1$, 0, or -1 depending on the input value $x_q(t)$ and the threshold value Λ.

As the received $\mathbf{y}(t)$ is complex, we decompose it into two real-valued vectors $\tilde{\mathbf{y}}(t) = [\mathbf{v}(t)^T, \mathbf{b}(t)^T]^T$, where $\mathbf{v}(t) = \Re\{\mathbf{y}(t)\}$ and $\mathbf{b}(t) = \Im\{\mathbf{y}(t)\}$ are both $M \times 1$ real-valued vectors. The covariance matrix of $\tilde{\mathbf{y}}(t)$ can be expressed as:

$$\tilde{\mathbf{R}}_{\mathbf{y}} = \mathbb{E}[\tilde{y}(t)\tilde{y}(t)^T] = \begin{bmatrix} \mathbf{R}_{\mathbf{vv}} & \mathbf{R}_{\mathbf{vb}} \\ \mathbf{R}_{\mathbf{vb}}^T & \mathbf{R}_{\mathbf{vv}} \end{bmatrix}, \tag{29}$$

where each block is an $M \times M$ matrix defined as $\mathbf{R}_{\mathbf{vv}} = \mathbb{E}[\mathbf{v}(t)\,\mathbf{v}(t)^T]$, $\mathbf{R}_{\mathbf{bb}} = \mathbb{E}[\mathbf{b}(t)\,\mathbf{b}(t)^T]$, and $\mathbf{R}_{\mathbf{vb}} = \mathbb{E}[\mathbf{v}(t)\,\mathbf{b}(t)^T]$. Similarly, the original complex signal $\mathbf{x}(t)$ can be expressed as $\tilde{\mathbf{x}}(t) = [\mathbf{u}(t)^T, \mathbf{i}(t)^T]^T$, where $\mathbf{u}(t) = \Re\{\mathbf{x}(t)\}$ and $\mathbf{i}(t) = \Im\{\mathbf{x}(t)\}$. The unquantized covariance matrix $\tilde{\mathbf{R}}_{\mathbf{x}}$ with separate real and imaginary parts is then given by:

$$\tilde{\mathbf{R}}_{\mathbf{x}} = \mathbb{E}[\tilde{\mathbf{x}}\tilde{\mathbf{x}}^T] = \begin{bmatrix} \mathbf{R}_{\mathbf{uu}} & \mathbf{R}_{\mathbf{ui}} \\ \mathbf{R}_{\mathbf{ui}}^T & \mathbf{R}_{\mathbf{ii}} \end{bmatrix}. \tag{30}$$

The covariance matrix we hope to reconstruct for DOA estimation is:

$$\mathbf{R}_{\mathbf{x}} = \mathbb{E}\big[\mathbf{x}(t)\,\mathbf{x}(t)^H\big] = \mathbf{R}_{\mathbf{uu}} + \mathbf{R}_{\mathbf{ii}} - j\big(\mathbf{R}_{\mathbf{ui}} - \mathbf{R}_{\mathbf{ui}}^T\big). \tag{31}$$

The normalized covariance matrix $\bar{\mathbf{R}}_{\mathbf{x}}$ is

$$\bar{\mathbf{R}}_{\mathbf{x}} = {\mathbf{D}_{\mathbf{x}}}^{-1}\,\mathbf{R}_{\mathbf{x}}\,{\mathbf{D}_{\mathbf{x}}}^{-1} \tag{32}$$

where $\mathbf{D}_{\mathbf{x}} = \mathrm{diag}([\sigma_{x_1}, \ldots, \sigma_{x_M}])$ is assumed to be known.

Correspondingly, we define the normalized block covariance matrix as

$$\tilde{\bar{\mathbf{R}}}_{\mathbf{x}} = \begin{bmatrix} \bar{\mathbf{R}}_{\mathbf{uu}} & \bar{\mathbf{R}}_{\mathbf{ui}} \\ \bar{\mathbf{R}}_{\mathbf{ui}}^T & \bar{\mathbf{R}}_{\mathbf{ii}} \end{bmatrix}, \tag{33}$$

where, $\bar{\mathbf{R}}_{\mathbf{uu}} = {\mathbf{D}_{\mathbf{x}}}^{-1}\mathbf{R}_{\mathbf{uu}}{\mathbf{D}_{\mathbf{x}}}^{-1}$, and similar for the other blocks, with all diagonal entries of $\tilde{\bar{\mathbf{R}}}_{\mathbf{x}}$ equal to 1. The normalized complex covariance matrix is then recovered by

$$\bar{\mathbf{R}}_{\mathbf{x}} = \bar{\mathbf{R}}_{\mathbf{uu}} + \bar{\mathbf{R}}_{\mathbf{ii}} - j\big(\bar{\mathbf{R}}_{\mathbf{ui}} - \bar{\mathbf{R}}_{\mathbf{ui}}^T\big). \tag{34}$$

Our method directly outputs $\bar{\mathbf{R}}_{\mathbf{x}}$, which suffices for subspace-based DOA estimation algorithms such as MUSIC, since the spatial spectrum depends only on normalized correlation.

Recovering the normalized covariance matrix $\bar{\mathbf{R}}_{\mathbf{x}}$ from $\tilde{\mathbf{R}}_{\mathbf{y}}$ involves a reconstruction function based on (14). Instead of a simple expression like the arcsine law as in the 1-bit case, the equation represents an infinite series.

To approximate the infinite series on the right-hand side of (14) by a finite sum, define

$f_N(\rho) = \frac{2}{\pi}e^{-\kappa^2}\sum_{n=0}^{N}\frac{\rho^{2n+1}}{2^{2n}(2n+1)!}H_{2n}^2\left(\frac{\kappa}{\sqrt{2}}\right)$ as the truncated series, where the sum is taken over the first $N+1$ terms, and assume the normalized threshold κ is already known. Since $f_N(\rho_{jk})$ is a polynomial of order $N+1$, solving $f_N(\rho_{jk}) = [\tilde{\mathbf{R}}_{\mathbf{y}}]_{jk}$ yields $N+1$ roots in the complex plane. To ensure that ρ_{jk} falls within the physically valid domain (real part within the range of [-1,1]) and numerically well behaved (small imaginary part), additional constraints need to be imposed on ρ_{jk}. So we construct the following minimization problem:

$$\begin{aligned} \min_{\rho_{jk}} \quad & L(\rho_{jk}) = \left| f_N(\rho_{jk}) - [\tilde{\mathbf{R}}_{\mathbf{y}}]_{jk} \right|^2 \\ \text{s.t.} \quad & -1 \le \Re(\rho_{jk}) \le 1, \\ & 0 < \Im(\rho_{jk}) < \varepsilon. \end{aligned} \tag{35}$$

where ε is a tiny value of tolerance.

Lemma 3. Let $\mathcal{D} = \left\{\rho \in \mathbb{C} : -1 \le \Re(\rho) \le 1, 0 < \Im(\rho) < \varepsilon\right\}$ be the domain of interest. Assume that the truncated series $f_N(\rho)$ converges absolutely on $\mathcal{D}$, so $f_N(\rho)$ is analytic on $\mathcal{D}$. Define $L(\rho) = \left| f_N(\rho) - \hat{R} \right|^2$. Then $L(\rho)$ is continuously differentiable on $\mathcal{D}$, and its gradient satisfies a local Lipschitz condition. Specifically, for every $\rho_0 \in \mathcal{D}$ (where ρ_0 is any fixed point in the domain $\mathcal{D}$), there exists a neighborhood $U \subset \mathcal{D}$ of ρ_0 and a constant $K > 0$ such that for all $\rho_1, \rho_2 \in U$ (i.e., any two points in this neighborhood), $\|\nabla L(\rho_1) - \nabla L(\rho_2)\| \le K \|\rho_1 - \rho_2\|$. Since $L(\rho)$ is continuously differentiable (C^1) on $\mathcal{D}$, its gradient $\nabla L(\rho)$ remains well-behaved in any sufficiently small neighborhood around ρ_0, ensuring that the local Lipschitz condition holds.

The proof of Lemma 3 is provided in Appendix C.

Based on Lemma 3, (35) can be solved using the gradient descent method. The gradient of $L(p)$ is :

$$\nabla L(\rho_{jk}) = \frac{\partial L}{\partial \Re(\rho_{jk})} + i \cdot \frac{\partial L}{\partial \Im(\rho_{jk})}, \tag{36}$$

where

$$\frac{\partial L}{\partial \Re(\rho_{jk})} = 2\left(f_N(\rho_{jk}) - [\tilde{\mathbf{R}}_{\mathbf{y}}]_{jk}\right)\frac{\partial f_N}{\partial \Re(\rho_{jk})}, \tag{37}$$

and

$$\frac{\partial L}{\partial \Im(\rho_{jk})} = 2\left(f_N(\rho_{jk}) - [\tilde{\mathbf{R}}_{\mathbf{y}}]_{jk}\right)\frac{\partial f_N}{\partial \Im(\rho_{jk})}. \tag{38}$$

The partial derivative of $f_N(\rho_{jk}))$ with respect to the real and imaginary parts of ρ_{jk} is:

$$\begin{aligned} \frac{\partial f_N}{\partial \Re(\rho_{jk})} = \frac{2}{\pi}e^{-\frac{\kappa^2}{\sigma_{x_j}\sigma_{x_k}}}\sum_{n=0}^{\infty}\frac{n+1}{2^n(n+1)!}\rho_{jk}^n \\ H_n\left(\frac{\kappa}{\sqrt{2}\sigma_{x_j}}\right)H_n\left(\frac{\kappa}{\sqrt{2}\sigma_{x_k}}\right), \end{aligned} \tag{39}$$

$$\begin{aligned} \frac{\partial f_N}{\partial \Im(\rho_{jk})} = \frac{2}{\pi}e^{-\frac{\kappa^2}{\sigma_{x_j}\sigma_{x_k}}}\sum_{n=0}^{\infty}\frac{(n+1)i}{2^n(n+1)!}\rho_{jk}^n \\ H_n\left(\frac{\kappa}{\sqrt{2}\sigma_{x_j}}\right)H_n\left(\frac{\kappa}{\sqrt{2}\sigma_{x_k}}\right). \end{aligned} \tag{40}$$

After substituting (40) into (38) and (39) into (37), and then substituting (37) and (38) into (36), we have the full gradient expression:

$$\begin{aligned} \nabla L(\rho_{jk}) = & \frac{4}{\pi}e^{-\frac{\kappa^2}{\sigma_{x_j}\sigma_{x_k}}}\left(f_N(\rho_{jk}) - [\tilde{\mathbf{R}}_{\mathbf{y}}]_{jk}\right) \\ & \sum_{n=0}^{\infty}\frac{n+1}{2^n(n+1)!}\rho_{jk}^n H_n\Big(\frac{\kappa}{\sqrt{2}\sigma_{x_j}}\Big)H_n\Big(\frac{\kappa}{\sqrt{2}\sigma_{x_k}}\Big) \\ & + i\cdot\frac{4}{\pi}e^{-\frac{\kappa^2}{\sigma_{x_j}\sigma_{x_k}}}\left(f_N(\rho_{jk}) - [\tilde{\mathbf{R}}_{\mathbf{y}}]_{jk}\right) \\ & \sum_{n=0}^{\infty}\frac{(n+1)i}{2^n(n+1)!}\rho_{jk}^n H_n\Big(\frac{\kappa}{\sqrt{2}\sigma_{x_j}}\Big)H_n\Big(\frac{\kappa}{\sqrt{2}\sigma_{x_k}}\Big). \end{aligned} \tag{41}$$

Hence, the update rule of the gradient descent method is: $\rho_{jk}^{(m+1)} = \rho_{jk}^{(m)} - \alpha_m \nabla L\Big(\rho_{jk}^{(m)}\Big)$, where $\rho_{jk}^{(m)}$ is the estimate at the mth iteration, and $\alpha_m > 0$ is the step size.

Following the constraints of (35), after updating, we should project $\rho_{jk}^{(m+1)}$ to $\mathcal{D}$:

$$\begin{aligned} \mathrm{Proj}_{\mathcal{D}}(\rho_{jk}) = & \Big(\max(-1, \min(\Re(\rho_{jk}), 1))\Big) \\ & + i\Big(\min(\max(\Im(\rho_{jk}), 0), \varepsilon)\Big). \end{aligned} \tag{42}$$

When the following condition is satisfied $|L(\rho_{jk}^{(m+1)}) - L(\rho_{jk}^{(m)})| < \epsilon$, the iteration stops and we obtain the final result of ρ_{jk}.

A summary of the steps is provided in Table Algorithm 1. Given a finite number of snapshots P, the real-valued 1.5-bit covariance matrix $\tilde{\mathbf{R}}_{\mathbf{y}}$ (or its complex counterpart $\mathbf{R}_{\mathbf{y}}$) must be estimated empirically as

$$\tilde{\mathbf{R}}_{\mathbf{y}} = \frac{1}{P}\sum_{t=1}^{P}\tilde{\mathbf{y}}(t)\,\tilde{\mathbf{y}}(t)^T. \tag{43}$$

This sample covariance $\tilde{\mathbf{R}}_{\mathbf{y}}$ is then processed by Algorithm 1 to yield an estimate $\tilde{\bar{\mathbf{R}}}_{\mathbf{x}}$ of the real-valued normalized covariance matrix in the augmented real domain. By recombining the real and imaginary parts according to (34), we recover the estimated normalized complex covariance matrix $\bar{\mathbf{R}}_{\mathbf{x}}$.

When the signal variances $\{\sigma_{x_j}^2\}$ are known, the full covariance matrix can be reconstructed as $\mathbf{R}_{\mathbf{x}} = \mathbf{D}_{\mathbf{x}}\bar{\mathbf{R}}_{\mathbf{x}}\mathbf{D}_{\mathbf{x}}$, enabling direct application of subspace-based DOA algorithms such as MUSIC.

Nevertheless, for subspace-based methods, only the subspace structure of $\mathbf{R}_{\mathbf{x}}$ is relevant. The eigenvectors of the normalized and full covariance matrices differ only by a non-singular scaling, which preserves the spatial signatures required by MUSIC and related estimators. Therefore, even if the signal variances $\{\sigma_{x_j}^2\}$ are unknown, accurate DOA estimation remains feasible based solely on the estimated normalized correlation matrix $\bar{\mathbf{R}}_{\mathbf{x}}$.

Remark 2. The framework of reconstructing correlation coefficient matrix in Algorithm 1 can be applied to any b-bit quantization schemes, with only one modification to its gradient (see in line 8 of Algorithm 1). Specifically, replace

Algorithm 1 Normalized covariance matrix reconstruction from 1.5-bit array signals using the gradient descent method

Input: 1.5-bit array signal $\mathbf{y}(t)$, parameters κ, tolerance ϵ, maximum number of iterations $M_{\max}$, step size α_m, polynomial order N.

Output: the reconstructed original normalized complex covariance matrix $\bar{\mathbf{R}}_{\mathbf{x}}$.

$\tilde{\mathbf{y}}(t) = [\mathbf{v}(t)^T, \mathbf{b}(t)^T]^T \in \mathbb{R}^{2M}$.
Estimate sample covariance matrix $\tilde{\mathbf{R}}_{\mathbf{y}} \in \mathbb{R}^{2M\times 2M}$ based on P snapshots.
for each pair (j,k), $1 \le j,k \le 2M$ **do**
  Note: The element $[\tilde{\bar{\mathbf{R}}}_{\mathbf{x}}]_{jk}$ corresponds to a value ρ_{jk}, and $\rho_{jk} = 1$ when $j = k$.
  Initialize $\rho_{jk}^{(0)} \in \mathcal{D}$.
  repeat
    Compute gradient $\nabla L(\rho_{jk}^{(m)})$ using equation (41).
    Update:

$$\rho_{jk}^{(m+1)} \leftarrow \rho_{jk}^{(m)} - \alpha_m \nabla L\big(\rho_{jk}^{(m)}\big).$$

    Project:

$$\rho_{jk}^{(m+1)} \leftarrow \mathrm{Proj}_{\mathcal{D}}\left(\rho_{jk}^{(m+1)}\right),$$

  until $|L(\rho_{jk}^{(m)}) - L(\rho_{jk}^{(m-1)})| < \epsilon$ or $m \ge M_{\max}$
  Set final estimate: $\rho_{jk} \leftarrow \rho_{jk}^{(m)}$.
  Update the entry of: $[\tilde{\bar{\mathbf{R}}}_{\mathbf{x}}]_{jk} \leftarrow \rho_{jk}$.
end for
Reconstruct the original $\bar{\mathbf{R}}_{\mathbf{x}} \in \mathbb{C}^{M\times M}$:

$$\bar{\mathbf{R}}_{\mathbf{x}} = \bar{\mathbf{R}}_{\mathbf{uu}} + \bar{\mathbf{R}}_{\mathbf{ii}} - j\big(\bar{\mathbf{R}}_{\mathbf{ui}} - \bar{\mathbf{R}}_{\mathbf{ui}}^T\big).$$

the gradient of 1.5-bit scheme $L\big(\rho_{jk}^{(m)}\big)$ by the gradient of multi-bit scheme $L_b\big(\rho_{jk}^{(m)}\big)$

$$\nabla L_b(\rho_{jk}) = \frac{2}{\pi}\left(g_N(\rho_{jk}) - [\tilde{\mathbf{R}}_{\mathbf{y}}]_{jk}\right)\sum_{n=0}^{N}\frac{A_n}{2^n n!}\,\rho_{jk}^n, \quad (44)$$

where $g_N(\rho)$ is the correlation estimator for the b-bit quantizer (23), and A_n is defined in (22).

C. DOA Estimation Using the 1.5-bit Covariance Matrix

As can be seen from Section IV-B, unlike the 1-bit scheme, where we can use the simple arcsine law to recover the normalized covariance matrix $\bar{\mathbf{R}}_{\mathbf{x}}$, in the 1.5-bit case, an iterative optimization process has to be employed to recover $\bar{\mathbf{R}}_{\mathbf{x}}$ from $\tilde{\mathbf{R}}_{\mathbf{y}}$, as shown in Algorithm 1 and it has a very high computational complexity. In this part, we would like to show that in most of the cases, we can use $\mathbf{R}_{\mathbf{y}}$ directly for DOA estimation without going through the additional stage of recovering $\bar{\mathbf{R}}_{\mathbf{x}}$ from it. We refer to the approach that directly uses $\mathbf{R}_{\mathbf{y}}$ as an approximation to the unquantized covariance matrix for subspace estimation as Algorithm 2.

To derive the result, we first give the following lemma.

Lemma 4. Let all the diagonal entries of $\mathbf{R}_{\mathbf{x}}$ be p, i.e., $\mathrm{diag}(\mathbf{R}_{\mathbf{x}}) = p\mathbf{I}$, where $\mathbf{I}$ denotes the $M \times M$ identity matrix. Under the condition of a sufficiently small $|\rho| \in [0,1]$, $\mathbf{R}_{\mathbf{y}}$ can be approximated by $\mathbf{R}_{\mathbf{y}} \approx c\,\mathbf{R}_{\mathbf{x}} + \varepsilon\,\mathbf{I}$, where c and ε depend on κ and p. Consequently, $\mathbf{R}_{\mathbf{y}}$ and $\mathbf{R}_{\mathbf{x}}$ share the same set of eigenvectors, thus leading to identical signal and noise subspaces. In other words, $\mathbf{R}_{\mathbf{y}}$ can be used directly for subspace decomposition without reconstructing $\mathbf{R}_{\mathbf{x}}$.

Proof. If $j \neq k$, the truncated series can be expanded as:

$$\begin{aligned} f_N(\rho) &\approx f_N(0) + f_N'(0)\rho + \frac{f_N''(0)}{2}\rho^2 + \frac{f_N'''(0)}{6}\rho^3 \cdots \\ &\approx 0 + \frac{2}{\pi}e^{-\kappa^2}\rho + 0\cdot\rho^2 + \frac{(4\kappa^2-2)}{6\pi}e^{-\kappa^2}\rho^3 \cdots \\ &\approx \frac{2}{\pi}e^{-\kappa^2}\rho + \mathcal{O}(\rho^3). \end{aligned} \quad (45)$$

Assume $|\rho|$ is small enough. For non-diagonal entries (i.e., $(j \neq k)$ of $\mathbf{R}_{\mathbf{y}}$, one may approximate $[\mathbf{R}_{\mathbf{y}}]_{jk} \approx \frac{2}{p\pi}\,e^{-\kappa^2}\,[\mathbf{R}_{\mathbf{x}}]_{jk}$, where on the diagonal, it has $[\mathbf{R}_{\mathbf{y}}]_{jj} = 1$ and $[\mathbf{R}_{\mathbf{x}}]_{jj} = p$. Thus, collecting diagonal and off-diagonal elements, one obtains

$$\mathbf{R}_{\mathbf{y}} - \mathrm{diag}(\mathbf{R}_{\mathbf{y}}) \approx \left(\tfrac{2}{p\pi}\,e^{-\kappa^2}\right)\,\big(\mathbf{R}_{\mathbf{x}} - \mathrm{diag}(\mathbf{R}_{\mathbf{x}})\big), \quad (46)$$

Let $c = \frac{2}{p\pi}\,e^{-\kappa^2}$. Then, we have:

$$\mathbf{R}_{\mathbf{y}} \approx \mathbf{I} + c\,\big(\mathbf{R}_{\mathbf{x}} - p\mathbf{I}\big) = c\,\mathbf{R}_{\mathbf{x}} + \varepsilon\mathbf{I}, \quad (47)$$

where $\varepsilon = 1 - cp$. Since $\mathbf{R}_{\mathbf{x}}$ can be diagonalized as $\mathbf{R}_{\mathbf{x}} = \mathbf{V}\Lambda\mathbf{V}^H$, for any eigenvector $\mathbf{v}$ of $\mathbf{R}_{\mathbf{x}}$ associated with eigenvalue γ_i, we have

$$\mathbf{R}_{\mathbf{y}}\mathbf{v} \approx \big(c\,\mathbf{R}_{\mathbf{x}} + \varepsilon\,\mathbf{I}\big)\mathbf{v} = c\,\gamma_i\,\mathbf{v} + \varepsilon\,\mathbf{v} = \big(c\,\gamma_i + \varepsilon\big)\,\mathbf{v}. \quad (48)$$

Hence $\mathbf{v}$ remains an eigenvector of $\mathbf{R}_{\mathbf{y}}$, but the corresponding eigenvalue is shifted to $c\,\kappa_i + \varepsilon$. This eigenvalue shift preserves the relative order of eigenvalues for small ρ, which in turn preserves the partition between signal subspace and noise subspace. □

Therefore, based on Lemma 4, $\mathbf{R}_{\mathbf{y}}$ and $\mathbf{R}_{\mathbf{x}}$ share a similar set of eigenvectors under the small-ρ assumption, and once we obtain $\mathbf{R}_{\mathbf{y}}$, the subspace-based methods such as MUSIC can be applied directly by employing the 1.5-bit measurements based covariance matrix $\mathbf{R}_{\mathbf{y}}$ to estimate DOAs.

As we know, the maximum first-order approximation error of 1-bit is $\frac{2}{\pi}|arcsin(1) - 1| \approx 0.36$, and the error of 1.5-bit approximation is $\Delta f_N(\rho) = f_N(\rho) - \frac{2}{\pi}e^{-\kappa^2}\rho$; when $\rho = 1$, the maximum error would decrease with the increase of κ; however, it is not clear what happens to this error when $\rho \neq 1$. Inspired by the polynomial series of 1-bit case at the end of Section III-A, we conjecture that for the 1.5-bit case, we could further relax the condition that ρ be sufficiently small in Lemma 4, and thus we propose the following Lemma 5.

Lemma 5. Define the first-order approximation errors

$$\Delta f_{N1}(\rho) = \frac{2}{\pi}\big[\arcsin(\rho) - \rho\big], \quad (49a)$$

$$\Delta f_N(\rho) = f_N(\rho) - \frac{2}{\pi}e^{-\kappa^2}\rho. \quad (49b)$$

Then, for every $\kappa > 0$ and every $\rho \in [-1,1]$, $\big|\Delta f_N(\rho)\big| \le \Delta f_{N1}\big(|\rho|\big)$.

The proof of Lemma 5 is provided in Appendix D.

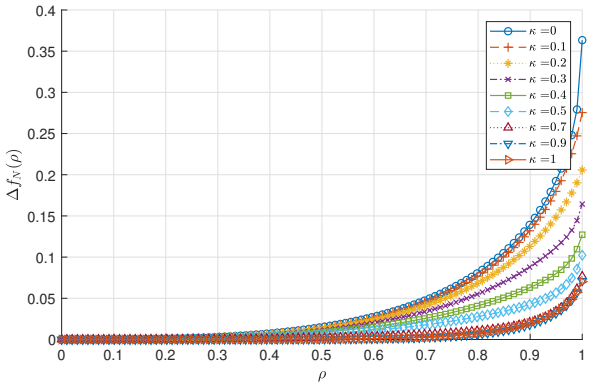


Fig. 1. $\Delta f_N(\rho)$ vs. ρ in different κ.

Fig. 1 illustrates how the first-order approximation changes with ρ for several threshold values κ. All 1.5-bit curves lie well below the 1-bit reference suggesting that the "small-ρ" condition is much less restrictive for the 1.5-bit quantizer. Lemma 5 formalizes this observation: for every $\kappa > 0$ and every $\rho \in [-1, 1]$, $\left|\Delta f_N(\rho)\right| \leq \Delta f_1(\rho)$, so the first-order linearization error of the 1.5-bit mapping is provably smaller than that of the classic 1-bit arcsine law. Hence the range of correlation coefficients for which the approximation $\mathbf{R_y} \approx c\mathbf{R_x} + \varepsilon\mathbf{I}$ is accurate is substantially wider than that required by the 1-bit approximation used in [35].

Corollary 1. With the same assumption of p, ρ, and $\mathbf{R_x}$ as in Lemma 4, consider $\mathbf{x}(t)$ passing through a b-bit quantizer $Q_b(\cdot)$ and output $\mathbf{y}_b(t)$, where its quantized covariance matrix is $\mathbf{R}_{\mathbf{y}_b}$. Assume the corresponding correlation estimator is given by (23). Then, the sample covariance matrix is

$$\mathbf{R}_{\mathbf{y}_b} \;=\; c_b\,\mathbf{R_x} \;+\; \varepsilon_b\,\mathbf{I} \;+\; O\big(\rho_{\max}^3\big), \tag{50}$$

where $c_b = \frac{A_0}{2\pi p}$, and $\varepsilon_b = \sigma_{y_b}^2 - c_b p$, with $A_0 = \sum_{k=1}^{A-1}\sum_{m=1}^{A-1} \Delta z_k\,\Delta z_m\, e^{-(u_k^2+v_m^2)}$ and $\sigma_{y_b}^2 = \mathbb{E}\big[\,Q_b(x)^2\big]$.

Proof. For a symmetric b-bit quantizer, the closed-form series (23) becomes $[\mathbf{R}_{\mathbf{y}_b}]_{jk} = \frac{1}{2\pi}\Big(A_0\,\rho_{jk} + O(\rho_{jk}^3)\Big)$.

By Lemma 1, if n is odd ($n = 2\ell + 1$), $H_n(x)$ is an odd function. The paired thresholds satisfy $u_{A-k} = -u_k$ and the level differences obey $\Delta z_{A-k} = -\Delta z_k$. Hence each complementary pair in the double sum that defines A_n cancels:

$$\begin{aligned} &\Delta z_k\,H_n(u_k) + \Delta z_{A-k}\,H_n(u_{A-k}) \\ &= \Delta z_k H_n(u_k) - \Delta z_k H_n(u_k) = 0. \end{aligned} \tag{51}$$

Therefore, all odd-index coefficients vanish, i.e. $A_1 = A_3 = A_5 = \cdots = 0$. Then,

$$[\mathbf{R}_{\mathbf{y}_b}]_{jk} = \begin{cases} \sigma_{y_b}^2, & j = k, \\ \dfrac{A_0}{2\pi p}\,[\mathbf{R_x}]_{jk} + O(\rho_{\max}^3), & j \neq k. \end{cases} \tag{52}$$

Collecting the diagonal and off-diagonal parts gives (50) and multiplying a Hermitian matrix by a positive scalar and adding a scaled identity do not change its eigenvectors, so $\mathbf{R}_{\mathbf{y}_b}$ and $\mathbf{R_x}$ have identical eigenvector set up to the neglected $O(\rho_{\max}^3)$ term. □

D. Computational Complexity Analysis

In this subsection, we analyze the computational complexity of our proposed two algorithms in Sections IV-B and IV-C. Since both algorithms end up using MUSIC or the same subspace-based method, we only need to focus on the computational complexity of the algorithm to obtain the covariance matrix that is ultimately applied to the subspace-based DOA estimation algorithm. For Algorithm 1 from Section IV-B, it requires an iterative optimization process to recover $\mathbf{R_x}$ from the quantized, real-valued covariance $\tilde{\mathbf{R}}_{\mathbf{y}}$, and each element $[\tilde{\mathbf{R}}_{\mathbf{x}}]_{jk}$ is recovered by solving (35) via the gradient descent method with the update

$$\rho_{jk}^{(m+1)} = \mathrm{Proj}_{\mathcal{D}}\left(\rho_{jk}^{(m)} - \alpha_m \nabla L\big(\rho_{jk}^{(m)}\big)\right), \tag{53}$$

where the gradient $\nabla L(\rho_{jk})$ is given by equation (41) and $f_N(\rho)$ is the truncated series

$$f_N(\rho) = \frac{2}{\pi} e^{-\kappa^2} \sum_{n=0}^{N} \frac{\rho^{2n+1}}{2^{2n}(2n+1)!} H_{2n}^2\Big(\frac{\kappa}{\sqrt{2}}\Big). \tag{54}$$

For the function (54), its gradient comprises $(N+1)$ terms, with each involving a power ρ^{2n+1}, a factorial expression, and a Hermite polynomial $H_{2n}\big(\frac{\kappa}{\sqrt{2}}\big)$. The Hermite polynomials are computed via the standard three-term recursion $H_0(x) = 1, \quad H_1(x) = 2x, \quad H_{m+1}(x) = 2x\,H_m(x) - 2m\,H_{m-1}(x)$, and they can be generated for all $n = 0, \ldots, N$ in $\mathcal{O}(N)$ total operations if precomputed. However, if these polynomials are computed on-the-fly using the above recursion, the cost to compute $H_{2n}\big(\frac{\kappa}{\sqrt{2}}\big)$ is $\mathcal{O}(n)$ for each term. Consequently, evaluating one term in the series requires $\mathcal{O}(n)$ operations, and for a total of $n = 0, \ldots, N$ yields a complexity of $\sum_{n=0}^{N} \mathcal{O}(n) = \mathcal{O}(N^2)$. Once all the required Hermite polynomials have been computed, evaluating each series term would require only basic arithmetic operations; but without pre-computation, the total cost becomes $\mathcal{O}(N^2)$. A similar argument applies to the gradient, which is formed by differentiating the same series term by term. Thus, both $f_N(\rho)$ and its gradient can be computed in $\mathcal{O}(N^2)$ time per iteration, and the iterative process converges in $M_{\max}$ iterations per covariance element; the per-element complexity is $\mathcal{O}(M_{\max}\,N^2)$. Since there are $\mathcal{O}(M^2)$ elements in the $2M \times 2M$ matrix $\tilde{\mathbf{R}}_{\mathbf{x}}$, the total complexity for the iterative reconstruction stage is $\mathcal{O}\big(M^2 \cdot M_{\max} \cdot N^2\big)$. In addition, the sample covariance matrix $\tilde{\mathbf{R}}_{\mathbf{y}}$ is estimated from P snapshots at a cost of $\mathcal{O}(P\,M^2)$. Finally, the reconstruction of the original covariance matrix $\mathbf{R}_x$ is a matrix reorganization operation with a cost of $\mathcal{O}(M^2)$. Thus, the overall complexity of Algorithm 1 is $\mathcal{O}\big(P\,M^2 + M^2 \cdot M_{\max}\,N^2\big)$.

In contrast, Algorithm 2 in Section IV.C bypasses the covariance recovery stage entirely by directly using the 1.5-bit covariance matrix $\mathbf{R_y}$ computed from the 1.5-bit measurements for DOA estimation. The direct estimation of $\mathbf{R_y}$ from P snapshots requires $\mathcal{O}(P\,M^2)$, which is significantly lower than that of Algorithm 1, especially when $M_{\max}$ or N is large. This analysis demonstrates that Algorithm 1 incurs a high computational cost due to the iterative optimization over all $\mathcal{O}(M^2)$ covariance elements. In many practical scenarios, Algorithm 2 provides a favorable trade-off by directly using

$\mathbf{R_y}$ for DOA estimation, substantially reducing computational complexity while still achieving acceptable estimation performance, as demonstrated later by our simulation results.

V. Simulations and Results

A. Simulation Setup

In this section, we present comprehensive numerical results to evaluate the performance of our proposed 1.5-bit quantization framework and associated DOA estimation algorithms, together with both 1-bit and 2-bit benchmarks.

For each quantization level (1-bit, 1.5-bit, and 2-bit), we implement both "Recon" and "Direct" MUSIC algorithms. Specifically, the "Recon" approach is based on Algorithm 1, which was originally formulated for 1.5-bit quantization. To adapt this framework for multi-bit schemes, such as 2-bit, it suffices to replace the gradient computation in line 8 of Algorithm 1, i.e., $\nabla L(\rho_{jk})$, with the corresponding multi-bit version $\nabla L_b(\rho_{jk})$ derived from the general b-bit correlation estimator (see Remark 2). Thus, "Recon 2-bit MUSIC" is implemented by this straightforward substitution. The "Direct" approach, in contrast, applies MUSIC directly to the sample covariance matrix of the quantized data (e.g., 1-bit, 1.5-bit, or 2-bit) without reconstruction, based on Lemma 4 and Corollary 1. As additional baselines for comparison, we include the unquantized MUSIC, as well as the 1-bit scheme and its classic algorithms "Recon 1-bit MUSIC" and "Direct 1-bit MUSIC" following [35], [37].

All experiments are performed using 500 independent Monte Carlo trials. The root mean square error (RMSE) is adopted as the main performance metric, defined as

$$\mathrm{RMSE} = \sqrt{\frac{1}{JK}\sum_{j=1}^{J}\sum_{k=1}^{K}\left(\hat{\theta}_{j,k} - \theta_k\right)^2}, \tag{55}$$

where K is the number of sources, J is the number of Monte Carlo runs, $\hat{\theta}_{j,k}$ is the estimated DOA of the kth source in the jth run, and θ_k is its true DOA. Throughout the experiments, the DOA estimation grid interval is fixed at 0.1°.

Unless otherwise stated, the step size is set to $\alpha = 0.1$ and the maximum number of iterations to $M_{\max} = 200$ for all gradient-based reconstruction algorithms.

B. Polynomial Order Selection

Algorithm 1 reconstructs unquantized covariance matrix via truncated polynomials. Using excessively high polynomial orders may lead to an unnecessarily high computational complexity without clear improvement in performance. To determine an appropriate order for our experiments, we choose the 70th-order polynomial as an approximation of "infinite" order. Define $O_{err}(n) = \left| f_{70}(\rho) - f_n(\rho) \right|$, which measures the difference between the nth-order approximation and the 70th-order reference. Fig. 2 depicts $O_{err}(n)$ from the 1st to the 30th order over a wide range of ρ, setting $\kappa = 0.3$. We observe that the approximation error grows with ρ for a given polynomial order, and that lower ρ values converge faster. Balancing accuracy and computational burden, we fix the polynomial order to $n = 10$ for Algorithm 1 in subsequent experiments.

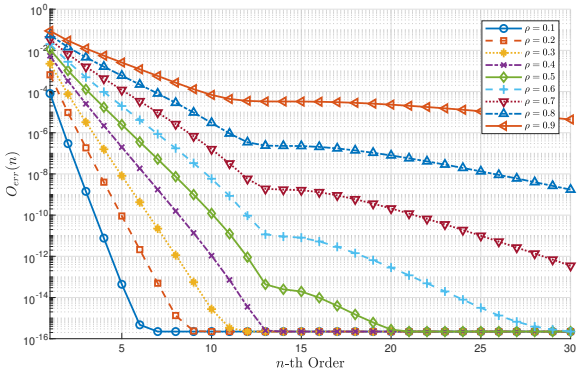


Fig. 2. $O_{err}(n)$ vs. nth Order, where $\kappa = 0.3$ and ρ varies from 0 to 1.

C. Numerical Verification of Lemma 4 and Lemma 5

To quantify the effectiveness of Lemma 4, we evaluate the relative approximation error between the exact quantized covariance matrix and its linear approximation:

$$\Delta\mathbf{R_y} = \frac{\|\mathbf{R_y}^{\mathrm{app}} - \mathbf{R_y}\|_F}{\|\mathbf{R_y}\|_F}, \tag{56}$$

where $\|\cdot\|_F$ denotes the Frobenius norm, $\mathbf{R_y}$ is constructed such that its diagonal entries are set to 1, while its off-diagonal entries are given by the closed-form estimator (14), and $\mathbf{R_y}^{\mathrm{app}}$ is given by (47).

We assume that $\mathbf{R_x}$ and p are known. Consider a ULA with $M = 10$ sensors and two equal-power sources impinging from -10° and 3.5°. For each SNR value, $P = 500$ snapshots are generated. For 1.5-bit quantization, the normalized threshold κ varies from 0.1 to 1.0.

The results in Fig. 3 demonstrate that, although Lemma 4 is formally established under the assumption of sufficiently small correlation (small $|\rho|$), the linear approximation for 1.5-bit quantization remains highly accurate across a much wider range of signal correlations and SNR values than its 1-bit counterpart. In particular, while the approximation error for 1-bit quantization increases rapidly as $|\rho|$ grows, the error for 1.5-bit quantization stays uniformly low even for moderately large $|\rho|$, which further validates Lemma 5.

Furthermore, as the normalized threshold κ increases, the valid range for the small-ρ linearization expands, relaxing the restrictive condition required in Lemma 4. However, it is important to note that increasing the threshold does not always improve performance, as excessively large thresholds may lead to signal loss or reduced information content. This observation motivates the next set of experiments, where we further investigate the impact of threshold selection on estimation accuracy.

D. Threshold Selection for the 1.5-bit Scheme

Since no closed-form rule currently exists for selecting the optimal threshold in the 1.5-bit case, we conduct an empirical sweep to identify a suitable value of κ.

All threshold selection experiments employ a ULA with $M = 10$ sensors and $N = 200$ snapshots. For $K = 2$, source DOAs are fixed at -10° and 3.5°; for $K = 6$, sources are uniformly spaced in $[-60^\circ, 60^\circ]$. We evaluate four

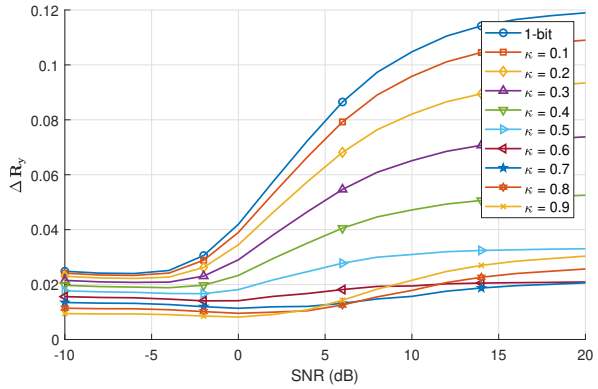
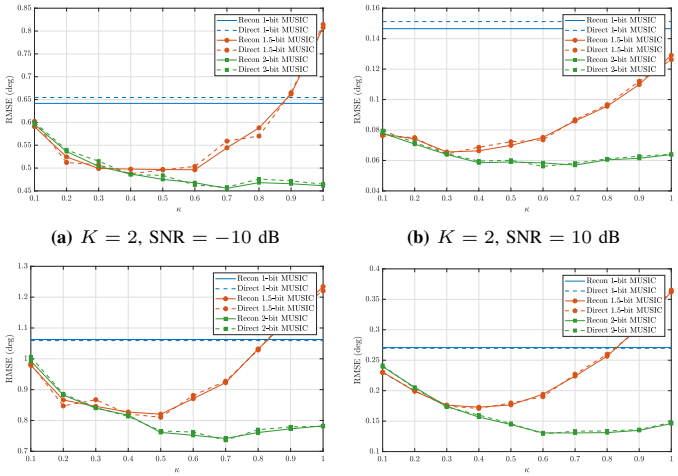


Fig. 3. $\Delta \mathbf{R}_{\mathbf{y}}$ vs. SNR in different κ.



(a) $K = 2$, SNR $= -10$ dB **(b)** $K = 2$, SNR $= 10$ dB

(c) $K = 6$, SNR $= -10$ dB **(d)** $K = 6$, SNR $= 10$ dB

Fig. 4. RMSE vs. κ: (a) $K = 2$ and SNR $= -10$ dB; (b) $K = 2$ and SNR $= 10$ dB; (c) $K = 6$ and SNR $= -10$ dB; (d) $K = 6$ and SNR $= 10$ dB.

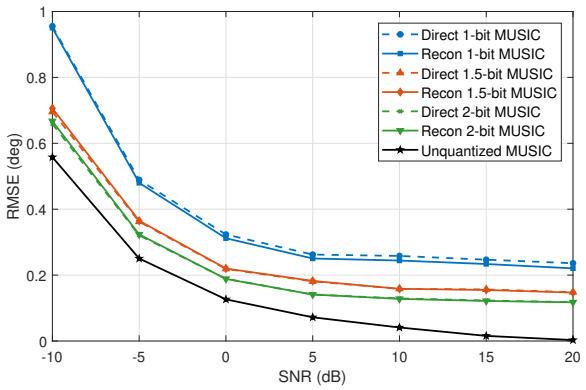


Fig. 5. RMSE versus SNR with P=200 snapshots.

representative scenarios, covering both low and high SNRs, and both sparse and dense sources. The results are depicted in Fig. 4a to 4d, corresponding to $(K = 2,\ \text{SNR} = -10\text{dB})$, $(K = 2,\ \text{SNR} = 10\ \text{dB})$, $(K = 6,\ \text{SNR} = -10\ \text{dB})$, and $(K = 6,\ \text{SNR} = 10\ \text{dB})$, respectively.

As shown in Fig. 4, the RMSE of the 1.5-bit scheme consistently attains its minimum near $\kappa = 0.4$, while the optimal value for 2-bit is around $\kappa = 0.7$. For instance, in Fig. 4b ($K = 2$, SNR $= 10$ dB), "Direct 1.5-bit MUSIC" achieves an RMSE of $0.0642°$, which is a 57.5% reduction compared to "Direct 1-bit MUSIC" ($0.1512°$); increasing to 2-bit only yields a modest 12.6% improvement ($0.0561°$). Similarly, in Fig. 4d ($K = 6$, SNR $= 10$dB), the improvement from 1-bit to 1.5-bit ($0.2696°$ vs. $0.1709°$, 36.6%) is larger than the subsequent 1.5-bit to 2-bit gain ($0.1709°$ vs. $0.1291°$, 24.5%). In the low SNR cases (Fig. 4a and 4c), the absolute RMSE values increase, but the qualitative trends remain the same.

It is also noteworthy that the gap between 1-bit and 1.5-bit quantizations is most striking when the number of sources is small (see Fig. 4a and 4b), and gradually narrows as more sources are added (Fig. 4c and 4d). This phenomenon can be attributed to the fact that, with only a few sources, the array output amplitudes are more concentrated around moderate values. In this case, introducing a third (zero) level in the 1-bit quantizer efficiently filters out weak signals and prevents severe information loss due to sign-only clipping, which resulting in substantial reduction in quantization error compared to 1-bit quantization. By contrast, adding an extra level to 1.5-bit to form 2-bit quantization has little effect, since very few samples reach these outer bins under such conditions. As the number of sources increases, the instantaneous amplitude distribution at the array output broadens, causing a larger proportion of samples to exceed the inner thresholds and populate the outermost quantization levels. In this case, 2-bit quantization begins to capture additional signal detail, leading to further improvement.

Based on these findings and experiments, we fix $\kappa = 0.4$ for all 1.5-bit experiments and $\kappa = 0.7$ for all 2-bit experiments in the remainder of this work.

E. Performance Comparisons

1) RMSE vs. SNR: Fig. 5 presents the RMSE versus SNR for a 12-sensor ULA with 6 sources uniformly spaced in $[-60°, 60°]$ and a fixed number of snapshots $P = 200$.

First, at low SNR (e.g., -10 or -5 dB), the performance of "Direct 1-bit MUSIC" closely matches "Recon 1-bit MUSIC," "Direct 1.5-bit MUSIC" matches "Recon 1.5-bit MUSIC," and "Direct 2-bit MUSIC" matches "Recon 2-bit MUSIC." This observation is in line with our theoretical analysis of Lemma 4 and Corollary 1: under the small-ρ condition, the sample covariance matrix of the quantized data provides an accurate estimation of their normalized unquantized covariance matrix.

As SNR increases, the differences among these schemes become increasingly pronounced. Specifically, the gap between "Direct 1-bit MUSIC" and "Recon 1-bit MUSIC" widens rapidly, This is because the approximation error increases with the correlation coefficient $|\rho|$. By contrast, "Direct 1.5-bit MUSIC" and "Recon 1.5-bit MUSIC" exhibit much smaller errors from their linear approximations. For 1.5-bit scheme, Lemma 5 establishes that, for any $\kappa > 0$, the first-order approximation error of the 1.5-bit quantizer is always no larger than that of 1-bit, which is also confirmed empirically in Fig. 3. As a result, "Direct 1.5-bit MUSIC" and "Recon 1.5-bit MUSIC" remain nearly indistinguishable over the entire SNR range. An interesting phenomenon is that even at high

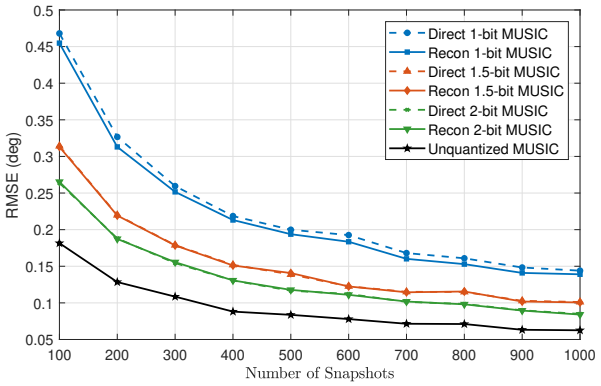


Fig. 6. RMSE versus snapshots at SNR=0dB.

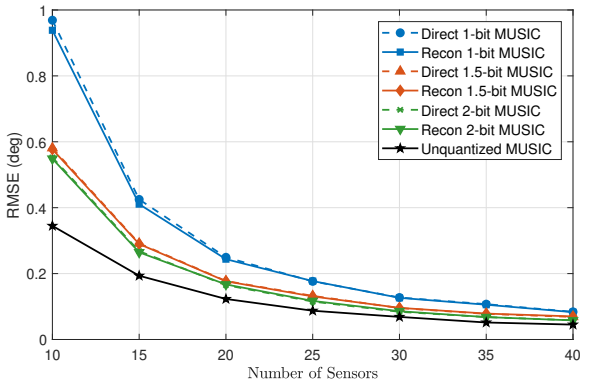


Fig. 7. RMSE versus number of sensors at SNR=0dB and P=50.

SNR, the performance gap between "Direct 2-bit MUSIC" and "Recon 2-bit MUSIC" remains negligible.

Notably, the performance gain in moving from 1-bit to 1.5-bit quantization is substantial, whereas the improvement from 1.5-bit to 2-bit is relatively minor. This indicates that increasing the quantization level from two to three yields much greater benefit for DOA estimation than further increasing it to four.

2) RMSE vs. Number of Snapshots: Fig. 6 shows the impact of the number of snapshots P on DOA estimation accuracy at fixed $\mathrm{SNR} = 0\,\mathrm{dB}$, using the same array and source settings as Section V-E1.

It is observed that, the curves for "Direct 1.5-bit MUSIC" and "Recon 1.5-bit MUSIC" are nearly identical over the entire snapshot range, as are those for the "2-bit" pair. This observation confirms that, for moderate and a sufficient number of snapshots, the linear approximation underlying the sample covariance approach is accurate, which corroborates our previous analysis and experiments.

More importantly, 1.5-bit quantization achieves the same RMSE as 1-bit quantization with substantially fewer snapshots. For example, at $P = 100$, the performance of "Direct 1.5-bit MUSIC" is comparable to that of "Direct 1-bit MUSIC" at $P = 200$. Likewise, with $P = 400$ snapshots, "Direct 1.5-bit MUSIC" matches the performance of "Direct 1-bit MUSIC" with $P = 800$ snapshots. Meanwhile, "Direct 2-bit MUSIC" at $P = 100$ outperforms "Direct 1.5-bit MUSIC" at $P = 200$. Similarly, the performance of "Direct 2-bit MUSIC" with $P = 300$ snapshots is comparable to "Direct 1.5-bit MUSIC" with $P = 400$, whereas "Direct 1-bit MUSIC" needs $P = 800$ snapshots to achieve similar accuracy.

This quantization gain enables significant reductions in both the required sampling rate and observation time for practical sensor array systems.

3) RMSE vs. Number of Sensors: Fig. 7 shows the RMSE as a function of the array size M for $P = 50$ snapshots , $K = 5$ sources uniformly spaced in $[-60°, 60°]$, and $\mathrm{SNR} = 0\,\mathrm{dB}$.

The result reveals that the difference between each "Direct 1-bit MUSIC" and "Recon 1-bit MUSIC" narrows as M increases. This may be due to more sensor pairs having lower spatial correlation, reducing the impact of first-order approximation errors in the arcsine or Hermite expansions. As a result, unlike previous experiments, the curves for "Recon 1-bit MUSIC" and "Direct 1-bit MUSIC" essentially coincide for $M \geq 20$.

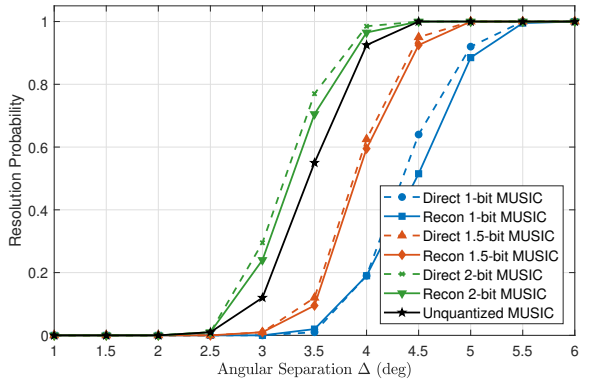


Fig. 8. Resolution probability versus angular separation at SNR=0 dB and P=500.

Moreover, the transition from 1-bit to 1.5-bit quantization yields the most significant improvement in RMSE, while the additional gain from 2-bit is comparatively modest. For instance, at $M = 30$, the RMSEs for 1-bit, 1.5-bit and 2-bit schemes are $0.1336°$, $0.0947°$ and $0.0876°$, respectively.

4) Resolution Probability: Fig. 8 presents the resolution probability versus angular separation Δ between two sources, for $M = 10$ sensors, $\mathrm{SNR} = 0\,\mathrm{dB}$, and $P = 500$ snapshots. The true DOAs are set as $[10.5°,\ 10.5° - \Delta]$, and Δ is swept from $1°$ to $10°$. For each separation Δ, a Monte Carlo trial is considered successful if, after sorting, both estimated DOAs fall within $\Delta/2$ of their corresponding true source locations. The resolution probability is then computed as the fraction of successful trials.

The results indicate that both the 1.5-bit and 2-bit schemes achieve resolution probabilities exceeding 0.95 for $\Delta \geq 4.5°$, outperforming the 1-bit scheme with only an around 0.6 resolution probability. Particularly, the direct use of the quantized covariance matrix can sometimes yield even higher resolution probabilities than reconstruction-based approaches. The reason may be that, for a finite number of snapshots, estimation noise can outweigh the errors introduced by first-order approximation in covariance recovery.

F. Robustness to Prior Variance mismatch

Recall that the threshold Λ is set as $\Lambda = \kappa\sigma_x$, where κ is the normalizing factor and σ_x is the signal standard deviation. However, in practice, σ_x is often estimated and subject to errors, leading to mismatches of thresholds that are not optimally selected.

We examine the robustness of the proposed method when the receiver uses an estimated signal variance. In detail, the standard deviation σ_x of each sensor is replaced by $\hat{\sigma}_x^{(q)} = (1+\varepsilon_q)\sigma_x$, where $\varepsilon_q \sim \mathcal{N}(0, \sigma_\varepsilon^2)$ is an independent calibration error for the q-th sensor. The designed κ therefore becomes $\kappa_q = \Lambda_{\text{actual}}^{(q)}/\hat{\sigma}_x^{(q)} = \kappa/(1+\varepsilon_q)$. A 10-element ULA with $K = 5$ sources uniformly spaced at $[-60°, 60°]$, SNR $= 0$ dB, and $P = 200$ snapshots is simulated while σ_ε varies from 0 to 0.2.

Fig. 9 plots the resulting RMSE and highlights two key observations regarding threshold-mismatch robustness in low-bit MUSIC-based DOA estimation.

First, "Direct" MUSIC is consistently more robust to variance errors than "Recon" MUSIC. This pattern holds for both 1.5-bit and 2-bit quantization. Understandably, variance mismatches distort the sample covariance matrix, and the analytic inversion required by "Recon" methods further amplifies such distortions through its nonlinear mapping. For "Direct" MUSIC, the subspace structure is affected only by the scaling of the sensor variances, so its performance degrades slowly as the variance estimation error σ_ε increases. In contrast, "Recon" MUSIC is highly sensitive, and the inversion amplifies sample noise and calibration errors, leading to a sharper performance drop.

On the other hand, 1.5-bit quantization is slightly more tolerant to variance errors than 2-bit quantization, regardless of "Recon" or "Direct". In the 1.5-bit scheme, the output variance of each sensor depends only weakly on its effective threshold, since only the inner levels $\pm\kappa_q$ are involved. This ensures that all sensors retain nearly equal variance and the overall subspace is still preserved, allowing both "Direct" and "Recon" MUSIC to be robust even under moderate calibration errors (e.g., $\sigma_\varepsilon \approx 0.05$). By contrast, 2-bit quantization introduces outer levels that make the output variance highly sensitive to any variance mismatch. As a result, even small errors (e.g., $\sigma_\varepsilon \approx 0.02$) quickly lead to large variance differences across sensors, breaking the energy balance required for accurate subspace estimation, and causing the performance of both "Direct 2-bit MUSIC" and especially "Recon 2-bit MUSIC" to deteriorate more rapidly. As σ_ε increases beyond approximately 0.05, the effective thresholds in the 1.5-bit scheme deviate substantially from their optimal values of 0.4. Meanwhile, as shown in Fig. 4, the 2-bit scheme appears to tolerate a wider range of threshold choices. Consequently, the performance of "Recon 1.5-bit MUSIC" begins to degrade more rapidly, eventually approaching the degradation rate observed for "Recon 2-bit MUSIC". The 1-bit scheme remains completely unaffected by variance mismatch because its threshold is always zero, but it is subject to a large amount of quantization noise, and its performance is still inferior to that of 1.5-bit and 2-bit when the variance mismatch is not large.

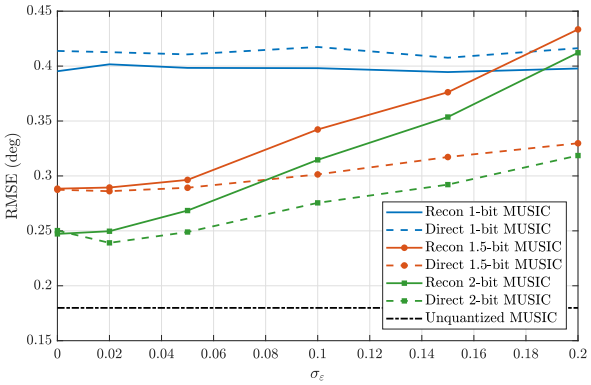


Fig. 9. RMSE versus threshold calibration error standard deviation σ_ε for $M = 10$, $K = 5$, $P = 200$, and SNR$= 0$ dB.

In summary, "Direct 1.5-bit MUSIC" achieves a favorable balance between accuracy and robustness: it substantially outperforms "Direct 1-bit MUSIC", maintains resilience to practical threshold mismatch, and is less sensitive than higher-bit approaches. Since "Recon" methods may amplify errors under mismatch or limited data, the "Direct 1.5-bit MUSIC" approach is recommended in scenarios where precise power calibration is difficult. It provides a practical compromise between estimation accuracy, robustness, and system complexity.

G. Computational Complexity Comparison

In this simulation, we consider a scenario with $M = 10$ sensors and $K = 6$ sources, each located at uniformly spaced angles in $[-60°, 60°]$, operating at an SNR of -10dB. Each trial uses 200 snapshots. We compare the proposed "Recon 1.5-bit MUSIC" with "Direct 1.5-bit MUSIC". For "Recon 1.5-bit MUSIC", we vary the polynomial order $N \in \{2, 5, 10\}$ and the maximum number of gradient-based iterations $M_{\max} \in \{50, 100, 200\}$, yielding 9 configurations. Table I presents the average running time per trial for all configurations. The experiments are conducted on a desktop PC with an AMD Ryzen 7 5800H CPU with Radeon Graphics, 16GB of 3200 MHz RAM, and a 64-bit Windows 11 operating system.

From Table I, one can see that "Direct 1.5-bit MUSIC" requires significantly less computational time than "Recon 1.5-bit MUSIC", since it does not perform the additional covariance reconstruction operation. Moreover, increasing either the polynomial order N or the maximum iteration count $M_{\max}$ causes the computational cost of "Recon 1.5-bit MUSIC" to increase by several orders of magnitude. In contrast, the "Direct 1.5-bit MUSIC" approach consistently achieves comparable estimation results as demonstrated earlier, but at significantly lower computational cost, indicating a more favorable trade-off between complexity and accuracy for practical applications.

VI. Conclusions

In this paper, a 1.5-bit quantization scheme has been proposed for DOA estimation and a closed-form expression

Table I: Average CPU time (s) for 1.5-bit methods

Algorithm	Avg CPU Time (s)
Direct 1.5-bit MUSIC	0.0016
Recon 1.5-bit MUSIC ($N = 2$, $M_{max} = 50$)	0.1006
Recon 1.5-bit MUSIC ($N = 2$, $M_{max} = 100$)	0.1980
Recon 1.5-bit MUSIC ($N = 2$, $M_{max} = 200$)	0.3960
Recon 1.5-bit MUSIC ($N = 5$, $M_{max} = 50$)	0.2359
Recon 1.5-bit MUSIC ($N = 5$, $M_{max} = 100$)	0.4894
Recon 1.5-bit MUSIC ($N = 5$, $M_{max} = 200$)	0.9583
Recon 1.5-bit MUSIC ($N = 10$, $M_{max} = 50$)	0.4909
Recon 1.5-bit MUSIC ($N = 10$, $M_{max} = 100$)	0.9857
Recon 1.5-bit MUSIC ($N = 10$, $M_{max} = 200$)	1.8878

derived to recover the unquantized correlation coefficients. Our theoretical analysis shows that 1.5-bit measurement data can, in most cases, be employed directly without reconstruction to achieve nearly identical DOA estimation accuracy compared to the reconstructed covariance matrix. We further investigated the empirical selection of the 1.5-bit threshold through simulations, providing practical guidance for choosing a suitable value. Numerical results show that the 1.5-bit scheme significantly improves performance over 1-bit, closely approaches 2-bit accuracy, and offers a superior trade-off between estimation quality and hardware complexity.

Appendix A
Proof of Lemma 1

We first substitute $u = -x$ into (12):

$$H_n(-x) = (-1)^n e^{x^2} \frac{d^n}{d(-x)^n} e^{-x^2}. \tag{57}$$

Rewrite (57) as

$$\begin{aligned} H_n(-x) &= (-1)^n e^{x^2} \cdot (-1)^n \frac{d^n}{dx^n} e^{-x^2} = e^{x^2} \frac{d^n}{dx^n} e^{-x^2} \\ &= (-1)^{-n} H_n(x) = (-1)^n H_n(x). \end{aligned} \tag{58}$$

which completes the proof.

Appendix B
Proof of Lemma 2

We begin by considering the limit

$$\lim_{n\to\infty} \left| \frac{a_{2n+3}\, \rho^{2n+3}}{a_{2n+1}\, \rho^{2n+1}} \right| = \lim_{n\to\infty} \left| \frac{a_{2n+3}}{a_{2n+1}} \right| \cdot |\rho|^2. \tag{59}$$

If this limit is lower than 1, then the series $\sum a_{2n+1}\rho^{2n+1}$ converges absolutely.

Then, we have

$$\begin{aligned} \frac{a_{2n+3}}{a_{2n+1}} &= \frac{\frac{1}{2^{2n+2}(2n+3)!} \sum_{k=n+1}^{\infty} \sum_{j=n+1}^{\infty} B_{k,j}}{\frac{1}{2^{2n}(2n+1)!} \sum_{k=n}^{\infty} \sum_{j=n}^{\infty} B_{k,j}} \\ &= \underbrace{\frac{\frac{1}{2^{2n+2}(2n+3)!}}{\frac{1}{2^{2n}(2n+1)!}}}_{\text{Constant factor}} \underbrace{\frac{\sum_{k=n+1}^{\infty} \sum_{j=n+1}^{\infty} B_{k,j}}{\sum_{k=n}^{\infty} \sum_{j=n}^{\infty} B_{k,j}}}_{R_n}. \end{aligned} \tag{60}$$

where $B_{k,j} = \frac{(-1)^{k+j}\,(2k)!\,(2j)!}{k!\,j!\,(2k-2n)!\,(2j-2n)!} \left(\frac{\kappa}{\sqrt{2}}\right)^{2(k+j-2n)}$. The constant factor is given by

$$\begin{aligned} \frac{\frac{1}{2^{2n+2}(2n+3)!}}{\frac{1}{2^{2n}(2n+1)!}} &= \frac{2^{2n}(2n+1)!}{2^{2n+2}(2n+3)!} \\ &= \frac{(2n+1)!}{4(2n+3)!} \\ &= \frac{1}{4} \cdot \frac{1}{(2n+3)(2n+2)}. \end{aligned} \tag{61}$$

As $n \to \infty$, this constant factor decays on the order of $\frac{1}{n^2}$.

For easy analysis, define R_n as:

$$R_n = \frac{\sum_{k=n+1}^{\infty} \sum_{j=n+1}^{\infty} B_{k,j}}{\sum_{k=n}^{\infty} \sum_{j=n}^{\infty} B_{k,j}}. \tag{62}$$

so that (59) can be rewritten as

$$\lim_{n\to\infty} \frac{1}{4} \cdot \frac{1}{(2n+3)\,(2n+2)} \,|R_n|\,|\rho|^2. \tag{63}$$

For any fixed n, observe that the set $\{(k,j) : k \geq n+1, j \geq n+1\}$ is a subset of $\{(k,j) : k \geq n, j \geq n\}$. Hence, we have

$$\left| \frac{\sum_{k=n+1}^{\infty} \sum_{j=n+1}^{\infty} B_{k,j}}{\sum_{k=n}^{\infty} \sum_{j=n}^{\infty} B_{k,j}} \right| \leq 1, \quad \text{i.e.} \quad |R_n| \leq 1. \tag{64}$$

As a result,

$$\lim_{n\to\infty} \frac{1}{4} \cdot \frac{1}{(2n+3)\,(2n+2)} \,|R_n|\,|\rho|^2 = 0 < 1. \tag{65}$$

Thus, we conclude that

$$\lim_{n\to\infty} \left| \frac{a_{2n+3}\, \rho^{2n+3}}{a_{2n+1}\, \rho^{2n+1}} \right| = 0 < 1, \tag{66}$$

which proves that the series $\sum_{n=0}^{\infty} a_{2n+1}\, \rho^{2n+1}$ converges absolutely.

Appendix C
Proof of Lemma 3

Since $f_N : \mathcal{D} \to \mathbb{C}$ is analytic on $\mathcal{D}$ based on Lemma. 2, writing $\rho = x + iy$ with $x, y \in \mathbb{R}$, there exist functions $u, v \in \mathcal{C}^{\infty}(\mathcal{D}, \mathbb{R})$, which means u and v are both infinitely differentiable. We may decompose $f_N(\rho) = u(x,y) + iv(x,y)$. Let $\hat{R} = a + ib$ with $a, b \in \mathbb{R}$. Then

$$L(x,y) = \Big[u(x,y) - a\Big]^2 + \Big[v(x,y) - b\Big]^2. \tag{67}$$

Since $u, v \in \mathcal{C}^{\infty}(\mathcal{D}, \mathbb{R})$, it follows that $L \in \mathcal{C}^{\infty}(\mathcal{D}, \mathbb{R})$. Hence, $L(\rho)$ is continuously differentiable on $\mathcal{D}$.

For any point $\rho_0 \in \mathcal{D}$, there exists a convex neighborhood $\mathcal{U} \subset \mathcal{D}$ containing ρ_0 such that $\overline{\mathcal{U}} \subset \mathcal{D}$ is compact. Because $\nabla^2 L$ is continuous, it is bounded on the compact set $\overline{\mathcal{U}}$, that is, there exists a constant $K_{\mathcal{U}} > 0$ such that

$$\|\nabla^2 L(\rho)\| \leq K_{\mathcal{U}} \quad \forall \rho \in \overline{\mathcal{U}}, \tag{68}$$

where $|\cdot|$ denotes Euclidean norm.

For any $\rho_1, \rho_2 \in \mathcal{U}$, let ξ be a point on the line segment between ρ_1 and ρ_2. The Mean Value Theorem for vector-valued functions implies that

$$\nabla L(\rho_1) - \nabla L(\rho_2) = \nabla^2 L(\xi)(\rho_1 - \rho_2). \tag{69}$$

Taking norms and using the bound on $\nabla^2 L$, we obtain

$$\|\nabla L(\rho_1) - \nabla L(\rho_2)\| \le \|\nabla^2 L(\xi)\| \|\rho_1 - \rho_2\| \le K_{\mathcal{U}} \|\rho_1 - \rho_2\|. \tag{70}$$

This inequality establishes that ∇L is Lipschitz continuous on $\mathcal{U}$; in other words, it satisfies a local Lipschitz condition.

Therefore, $L(\rho)$ is continuously differentiable on $\mathcal{D}$ and its gradient $\nabla L(\rho)$ satisfies a local Lipschitz condition.

Appendix D
Proof of Lemma 5

The power series of arcsine mentioned in Section III-A is

$$\arcsin(\rho) = \sum_{m=0}^{\infty} \frac{(2m)!}{4^m (m!)^2 (2m+1)} \, \rho^{2m+1} \tag{71}$$

which implies

$$\Delta f_{N1}(\rho) = \frac{2}{\pi} \sum_{m=1}^{\infty} a_m \, \rho^{2m+1}, \tag{72}$$

where $a_m = \frac{(2m)!}{4^m (m!)^2 (2m+1)} > 0$. The factor starts from $m = 1$ because the linear term is canceled, leaving the first–order error.

For the 1.5-bit case, its first-order error $f_N(\rho)$ is

$$\Delta f_N(\rho) = \frac{2}{\pi} e^{-\kappa^2} \sum_{n=1}^{\infty} b_n(\kappa) \, \rho^{2n+1}, \tag{73}$$

where $b_n(\kappa) = \frac{H_{2n}^2(\kappa/\sqrt{2})}{2^{2n}(2n+1)!} > 0$.

The generating function for the Hermite definition (12) is

$$e^{\,2xt - t^2} = \sum_{n=0}^{\infty} \frac{H_n(x)}{n!} \, t^n. \tag{74}$$

For any radius $r > 0$, the Cauchy's estimate applied to (74) yields

$$|H_{2n}(x)| \le (2n)! \, r^{-2n} \exp\bigl(2|x|r + r^2\bigr). \tag{75}$$

Choosing $r = \sqrt{n}\,(n \ge 1)$ gives

$$|H_{2n}(x)| \le (2n)! \exp\bigl(2|x|\sqrt{n} + n - n \ln n\bigr). \tag{76}$$

Because $2|x|\sqrt{n} \le x^2 + n$ and $n - n \ln n \le 0$ for $n \ge 3$, we have

$$|H_{2n}(x)| \le (2n)! \, e^{x^2}, \quad n \ge 3. \tag{77}$$

For the remaining $n = 0, 1, 2$, it is easy to prove that they satisfy the inequality. Hence, for all $n \ge 0$ and $x \in \mathbb{R}$,

$$|H_{2n}(x)| \le (2n)! \, e^{x^2}. \tag{78}$$

Letting $x = \kappa/\sqrt{2}$ in (78),

$$e^{-\kappa^2} \, H_{2n}^2\Bigl(\tfrac{\kappa}{\sqrt{2}}\Bigr) \le (2n)!^2, \tag{79}$$

we have

$$e^{-\kappa^2} \, b_n(\kappa) \le a_n, \qquad \forall n \ge 1, \ \forall \kappa > 0. \tag{80}$$

Because the coefficients in (72)–(73) are non-negative, (80) implies that for any $\rho \in [-1, 1]$

$$\begin{aligned} \bigl|\Delta f_N(\rho)\bigr| &= \frac{2}{\pi} \sum_{n=1}^{\infty} e^{-\kappa^2} b_n(\kappa) \, |\rho|^{2n+1} \\ &\le \frac{2}{\pi} \sum_{n=1}^{\infty} a_n \, |\rho|^{2n+1} = \Delta f_{N1}\bigl(\rho\bigr). \end{aligned} \tag{81}$$

The inequality is strict whenever $\kappa > 0$, which completes the proof.

References

[1] M. I. Skolnik, "Introduction to radar systems," *New York*, 1980.

[2] H. Krim and M. Viberg, "Two decades of array signal processing research: the parametric approach," *IEEE signal processing magazine*, vol. 13, no. 4, pp. 67–94, 1996.

[3] L. C. Godara, "Application of antenna arrays to mobile communications. ii. beam-forming and direction-of-arrival considerations," *Proceedings of the IEEE*, vol. 85, no. 8, pp. 1195–1245, 1997.

[4] H. L. Van Trees, *Optimum array processing: Part IV of detection, estimation, and modulation theory*. John Wiley & Sons, 2002.

[5] F. Shu, Y. Qin, T. Liu, L. Gui, Y. Zhang, J. Li, and Z. Han, "Low-complexity and high-resolution DOA estimation for hybrid analog and digital massive MIMO receive array," *IEEE Transactions on Communications*, vol. 66, no. 6, pp. 2487–2501, 2018.

[6] J. J. Bussgang, "Cross-correlation functions of amplitude-distorted gaussian signals," Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA, USA, Tech. Rep. 216, 1952.

[7] J. Van Vleck and D. Middleton, "The spectrum of clipped noise," *Proceedings of the IEEE*, vol. 54, no. 1, pp. 2–19, 1966.

[8] G. Jacovitti and A. Neri, "Estimation of the autocorrelation function of complex gaussian stationary processes by amplitude clipped signals," *IEEE transactions on information theory*, vol. 40, no. 1, pp. 239–245, 1994.

[9] A. Host-Madsen and P. Handel, "Effects of sampling and quantization on single-tone frequency estimation," *IEEE Transactions on Signal Processing*, vol. 48, no. 3, pp. 650–662, 2000.

[10] A. Host-Madsen and K. Andersen, "Lower bounds for estimation of frequency and phase of doppler signals," *Measurement Science and Technology*, vol. 6, no. 6, p. 637, 1995.

[11] C.-L. Liu and Z.-M. Lin, "One-bit autocorrelation estimation with non-zero thresholds," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 4520–4524, 2021.

[12] X. Zhang, Y. Cheng, H. C. So, and J. Li, "A novel mixed-ADC architecture for DOA estimation," *IEEE Signal Processing Letters*, vol. 31, pp. 611–615, 2024.

[13] X. Zhang, Y. Cheng, X. Shang, and J. Liu, "CRB analysis for mixed-adc based DOA estimation," *IEEE Transactions on Signal Processing*, vol. 72, pp. 3043–3058, 2024.

[14] A. Eamaz, F. Yeganegi, and M. Soltanalian, "Covariance recovery for one-bit sampled stationary signals with time-varying sampling thresholds," *Signal Processing*, vol. 206, p. 108899, 2023.

[15] ——, "Covariance recovery for one-bit sampled non-stationary signals with time-varying sampling thresholds," *IEEE Transactions on Signal Processing*, vol. 70, pp. 5222–5236, 2022.

[16] Y.-H. Xiao, L. Huang, D. Ramírez, C. Qian, and H. C. So, "Covariance matrix recovery from one-bit data with non-zero quantization thresholds: Algorithm and performance analysis," *IEEE Transactions on Signal Processing*, vol. 71, pp. 4060–4076, 2023.

[17] C.-L. Liu and P. Vaidyanathan, "One-bit normalized scatter matrix estimation for complex elliptically symmetric distributions," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2020, pp. 9130–9134.

[18] S. Dirksen, J. Maly, and H. Rauhut, "Covariance estimation under one-bit quantization," *The Annals of Statistics*, vol. 50, no. 6, pp. 3538–3562, 2022.

[19] S. Dirksen and J. Maly, "Tuning-free one-bit covariance estimation using data-driven dithering," *IEEE Transactions on Information Theory*, 2024.

[20] J. Chen, M. K. Ng, and D. Wang, "Quantizing heavy-tailed data in statistical estimation:(near) minimax rates, covariate quantization, and uniform recovery," *IEEE Transactions on Information Theory*, vol. 70, no. 3, pp. 2003–2038, 2023.

[21] H. Xu and Z. Yang, "Bit efficient toeplitz covariance estimation," *arXiv preprint arXiv:2412.12678*, 2024.

[22] S. Sedighi, M. B. Shankar, M. Soltanalian, and B. Ottersten, "DoA estimation using low-resolution multi-bit sparse array measurements," *IEEE Signal Processing Letters*, vol. 28, pp. 1400–1404, 2021.

[23] F. Schwab, "Two-bit correlators: Miscellaneous results," VLBA Correlator Memo 75, National Radio Astronomy Observatory, Tech. Rep., 1986.

[24] D. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[25] P. T. Boufounos and R. G. Baraniuk, "1-bit compressive sensing," in *Proc. the 42nd Annu. Conf. Inf. Sci and Syst.*, pp. 16–21, 2008.

[26] K. Knudson, R. Saab, and R. Ward, "One-bit compressive sensing with norm estimation," *IEEE Transactions on Information Theory*, vol. 62, no. 5, pp. 2748–2758, 2016.

[27] U. S. Kamilov, A. Bourquard, A. Amini, and M. Unser, "One-bit measurements with adaptive thresholds," *IEEE Signal Processing Letters*, vol. 19, no. 10, pp. 607–610, 2012.

[28] A. G. Zebadua, P.-O. Amblard, E. Moisan, and O. J. J. Michel, "Compressed and quantized correlation estimators," *IEEE Transactions on Signal Processing*, vol. 65, no. 1, pp. 56–68, 2017.

[29] A. Ameri, A. Bose, J. Li, and M. Soltanalian, "One-bit radar processing with time-varying sampling thresholds," *IEEE Transactions on Signal Processing*, vol. 67, no. 20, pp. 5297–5308, 2019.

[30] C. Xu and L. Jacques, "Quantized compressive sensing with RIP matrices: the benefit of dithering," *Information and Inference: A Journal of the IMA*, vol. 9, no. 3, pp. 543–586, 11 2019.

[31] X. Lu, W. Liu, and A. Alomainy, "A 1.5-bit quantization scheme and its application to sparse array direction estimation," in *2024 19th International Symposium on Wireless Communication Systems (ISWCS)*, Rio de Janeiro, Brazil, July 2024, pp. 1–5.

[32] R. Price, "A useful theorem for nonlinear devices having gaussian inputs," *IRE Transactions on Information Theory*, vol. 4, no. 2, pp. 69–72, 1958.

[33] W. Kibble, "An extension of a theorem of mehler's on hermite polynomials," in *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 41, no. 1. Cambridge University Press, 1945, pp. 12–15.

[34] M. Abramowitz and I. A. Stegun, *Handbook of mathematical functions with formulas, graphs, and mathematical tables*. US Government printing office, 1968, vol. 55.

[35] X. Huang and B. Liao, "One-bit MUSIC," *IEEE Signal Processing Letters*, vol. 26, no. 7, pp. 961–965, 2019.

[36] J. S. Dai, "Euler–rodrigues formula variations, quaternion conjugation and intrinsic connections," *Mechanism and Machine Theory*, vol. 92, pp. 144–152, 2015.

[37] O. Bar-Shalom and A. J. Weiss, "DOA estimation using one-bit quantized measurements," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 38, no. 3, pp. 868–884, 2002.

Xicheng Lu received the B.Sc. degree in electrical engineering from Fuzhou University, China, in 2020, and the M.Sc. degree in electronic and electrical engineering from the University of Sheffield, U.K., in 2022. He is currently pursuing the Ph.D. degree with the School of Electronic Engineering and Computer Science, Queen Mary University of London, U.K. His research interests include low-bit quantization, phase retrieval, and array signal processing.

Wei Liu (S'01-M'04-SM'10) received his BSc and LLB degrees from Peking University, China, MPhil from the University of Hong Kong, and PhD from the University of Southampton, UK. He then worked as a postdoc first at Southampton and later at Imperial College London. From 2005 to 2023, he was a Lecturer/Senior Lecturer at the Department of Electronic and Electrical Engineering, University of Sheffield, UK and from 2023 to 2024, a Reader at the School of Electronic Engineering and Computer Science, Queen Mary University of London (Visiting Professor from 2024 to 2027). Since 2024, he has been a Professor at the Department of Electrical and Electronic Engineering, Hong Kong Polytechnic University. He has published 470+ journal and conference papers, six book chapters, and two research monographs titled "Wideband Beamforming: Concepts and Techniques" (Wiley, 2010) and "Low-Cost Smart Antennas" (Wiley, 2019), respectively. His research interests are mainly focused on sensor array and multichannel signal processing and its various applications, such as robotics and autonomous systems, human computer interface, radar, sonar, and wireless communications.

He is a member of the Applied Signal Processing Systems Technical Committee (Vice-Chair for 2025-2026) of the IEEE Signal Processing Society (SPS), the Digital Signal Processing Technical Committee (Chair for 2022-2024) of the IEEE Circuits and Systems Society, and the IEEE SPS Education Board (2024-2026, Chair of its Educational Conference Program Committee), and a former member of the Sensor Array and Multichannel Signal Processing Technical Committee of the IEEE SPS (Chair for 2021-2022), the IEEE SPS Technical Directions Board (2021-2022), and the IEEE SPS Conference Board and its Executive Subcommittee (2022-2023). He also acted as an associate editor for IEEE Trans. on Signal Processing, IEEE Access, and Journal of the Franklin Institute, and an Executive Associate Editor-in-Chief of the journal Frontiers of Information Technology and Electronic Engineering; currently he is an associate editor for IEEE Transactions on Circuits and Systems I: Regular Papers and IEEE Antennas and Wireless Propagation Letters. He is an IEEE Distinguished Lecturer for the Aerospace and Electronic Systems Society (2023-2026).

Akram Alomainy (CEng, FRSA, FHEA, SMIEEE) is a leading academic in antennas, applied electromagnetics, and body-centric wireless communications. He is Deputy Dean for Postgraduate Research in the Faculty of Science and Engineering at Queen Mary University of London (QMUL), Lead of the Centre for Electronics, and Head of the Antennas and Electromagnetics Research Group. He oversees postgraduate research strategy across five schools, directing doctoral training, funding allocation, and research culture initiatives, while leading a team of over 70 staff and researchers supported by state-of-the-art facilities, including a £3.5 million antenna measurement laboratory. His research spans antennas, electromagnetic materials, terahertz communications, wearable and implantable devices, and nanoscale wireless networks, with applications in healthcare, defence, communications, and the creative industries. He has authored/co-authored 5 books, 6 book chapters, and 500+ papers (13,000+ citations), and secured £16M+ in competitive funding, including £6M+ as PI. His awards include the QMUL Research Excellence Award (2022), Education Excellence Award (2019), Abu Dhabi Research Excellence Awards (2017, 2018), and the Isambard Kingdom Brunel Award (2011). He is a Fellow of the Royal Society of Arts and a Chartered Engineer, serves on editorial boards including IEEE Access and IEEE Antennas and Wireless Propagation Letters, and has chaired international research panels. He is also a Visiting Fellow at Goldsmiths, University of London.