Channel Estimation and Data Detection Analysis of Massive MIMO With 1-Bit ADCs

Italo Atzeni , *Member, IEEE*, and Antti Tölli , *Senior Member, IEEE*

Abstract— We present an analytical framework for the channel estimation and the data detection in massive multiple-input multiple-output uplink systems with 1-bit analog-to-digital converters (ADCs) and i.i.d. Rayleigh fading. First, we provide closed-form expressions of the mean squared error (MSE) of the channel estimation considering the state-of-the-art linear minimum MSE estimator and the class of scaled least-squares estimators. For the data detection, we provide closed-form expressions of the expected value and the variance of the estimated symbols when maximum ratio combining is adopted, which can be exploited to efficiently implement minimum distance detection and, potentially, to design the set of transmit symbols. Our analytical findings explicitly depend on key system parameters such as the signal-to-noise ratio (SNR), the number of user equipments, and the pilot length, thus enabling a precise characterization of the performance of the channel estimation and the data detection with 1-bit ADCs. The proposed analysis highlights a fundamental SNR trade-off, according to which operating at the right noise level significantly enhances the system performance.

Index Terms— 1-bit ADCs, channel estimation, data detection, massive MIMO, performance analysis.

I. INTRODUCTION

THE migration of operating frequencies from first- to fourth-generation wireless systems, i.e., from 800 MHz to the sub-3 GHz range, did not bring major changes in terms of signal propagation. The current fifth generation (5G) features a more pronounced transition in this respect by operating at sub-6 GHz frequencies and, eventually, up to 30 GHz with the objective of boosting the data rates. Following this trend, beyond-5G systems will exploit the large amount of bandwidth available in the mmWave band (i.e., 30 GHz–300 GHz) and raise the operating frequencies up to 1 THz [2]. In this context, maintaining the same signal-to-noise ratio (SNR) over a given distance will require larger antenna arrays and increasingly sharp beamforming to spatially focus the signal power.

Although the short wavelength at mmWave and sub-THz frequencies allows to pack many antennas into a very small area, realizing fully digital, high-resolution massive multiple-input multiple-output (MIMO) arrays remains prohibitive in practice [3], [4].

As in the system model illustrated in Fig. 1, each base station (BS) antenna is generally equipped with a dedicated radio-frequency (RF) chain that includes complex, power-hungry analog-to-digital/digital-to-analog converters (ADCs/DACs) [4]. In this setting, while the transmit power can be made inversely proportional to the number of antennas, the power consumed by each ADC/DAC scales linearly with the sampling rate and exponentially with the number of quantization bits [5]–[9]. Another limiting factor is the amount of raw data exchanged between the remote radio head (RRH) and the base-band unit (BBU), which scales linearly with both the sampling rate and the number of quantization bits [10]–[12]. For these reasons, adopting low-resolution ADCs/DACs with 1 to 4 quantization bits as opposed to the typical 10 or more [13] enables the implementation of massive MIMO arrays comprising hundreds (or even thousands) of antennas, which are necessary to operate in the mmWave and sub-THz bands [10]. In this regard, 1-bit ADCs/DACs are particularly appealing due to their minimal power consumption and complexity since they only evaluate the sign of the input signal [5]. Such a coarse quantization is especially motivated at very high frequencies, where high-order modulations are not essential.

There is a large body of literature on massive MIMO with low-resolution and 1-bit ADCs/DACs, ranging from performance analysis to channel estimation and precoding design. The capacity of the low-resolution and the 1-bit quantized MIMO channel is characterized in [14] and [5], respectively, whereas [15] shows that replacing even a small number of high-resolution ADCs with 1-bit ADCs entails a modest performance loss. The performance-quantization trade-off in orthogonal frequency-division multiplexing (OFDM) uplink systems is studied in [6], which shows that using 4 to 6 quantization bits involves almost no performance loss compared with infinite-resolution ADCs. The spectral efficiency of single-carrier and OFDM uplink systems with 1-bit ADCs is analyzed in [16]. The problem of multi-user detection is considered, e.g., in [17] and [18] for low-resolution and 1-bit ADCs, respectively, whereas [19] focuses on the joint channel estimation and data detection. An efficient iterative method for near maximum likelihood detection with 1-bit ADCs is proposed in [7]. The work in [8] analyzes the channel estimation and the uplink achievable rate with 1-bit ADCs. In addition, it proposes a linear minimum mean squared error (MMSE)
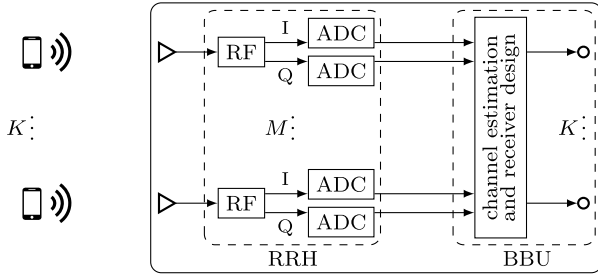
Fig. 1. Fully digital massive MIMO uplink system.

channel estimator based on the Bussgang decomposition, which allows to reformulate the nonlinear quantization function as a linear function with identical first- and second-order statistics [20]: we refer to this estimator as Bussgang linear MMSE (BLM) estimator. A similar analysis is presented in [21] for the downlink direction. The work in [10] extends some of the results derived in [16] and [8] for 1-bit ADCs to the multi-bit case. Specifically, it presents a throughput analysis of uplink systems and proposes a linear channel estimator based on the Bussgang decomposition with low-resolution ADCs. The channel estimation with 1-bit ADCs when the quantization threshold is not known is studied in [22]. The channel estimation exploiting the angular and delay structure is considered in [23] and [24] for low-resolution and 1-bit ADCs, respectively, whereas [25] exploits the temporal correlation for 1-bit ADCs. A recent line of works employs machine learning techniques in scenarios where obtaining accurate channel state information with low-resolution ADCs is impractical (see, e.g., [26], [27]). The benefits of oversampling for 1-bit quantized uplink systems are investigated in [28] and [29]. The performance of linear precoding schemes for downlink systems with 1-bit DACs is analyzed in [9]. A similar analysis with multi-bit DACs is presented in [11] considering both linear and nonlinear precoding, and in [12] considering linear precoding with oversampling in OFDM downlink systems. Lastly, [30] proposes a general optimization framework for downlink precoding with 1-bit DACs and constant envelope assuming quadrature amplitude modulation (QAM) transmit symbols.

A. Contribution

This paper broadens prior analytical studies on the channel estimation and the data detection in massive MIMO uplink systems with 1-bit ADCs. On the one hand, existing works do not provide a precise characterization of the performance of the channel estimation with 1-bit ADCs with respect to key system parameters such as the SNR, the number of user equipments (UEs), and the pilot length. We fill this gap by analyzing the mean squared error (MSE) of the channel estimation along with its asymptotic behavior at high SNR assuming independent and identically distributed (i.i.d.) Rayleigh fading channels among the UEs. In this respect, we consider the BLM estimator in [8] as well as the class of scaled least-squares (LS) estimators, such as the one proposed in [16] (which can be obtained from the former by ignor-

ing the temporal correlation of the quantization distortion). On the other hand, in the context of data detection with 1-bit ADCs, the statistical properties of the estimated symbols have not been characterized by existing works. In this regard, an interesting SNR trade-off was observed in [10], whereby the estimated symbols resulting from transmit symbols with the same phase overlap at high SNR; however, this aspect has not been formally described in the literature. We fill this gap by analyzing the expected value and the variance of the estimated symbols along with their asymptotic behavior at high SNR.[1] Our results on both the channel estimation and the data detection ultimately impact the symbol error rate (SER) and thus provide important practical insights into the design and the implementation of 1-bit quantized systems.

The contributions of this paper are summarized as follows:

- For the channel estimation with 1-bit ADCs, we derive closed-form expressions for the BLM estimator in [8] and for the class of scaled LS estimators, such as the one proposed in [16]. This enables a precise characterization of the performance of the channel estimation with respect to the SNR, the number of UEs, and the pilot length. Furthermore, we show that, in the case of i.i.d. Rayleigh fading channels among the UEs, the BLM estimator can be simplified as a scaled LS estimator with UE-specific scaling factors and that using a common optimized scaling factor for all the UEs entails a negligible performance loss.

- For the data detection with 1-bit ADCs, we characterize the statistical properties of the estimated symbols by deriving closed-form expressions of the expected value and the variance when maximum ratio combining (MRC) is adopted at the BS. These results can be exploited to efficiently implement minimum distance detection (MDD) and, potentially, to design the set of transmit symbols to further improve the data detection performance.

- Building on the proposed analysis, we provide a thorough discussion on the effect of 1-bit quantization on both the channel estimation and the data detection. For each of the two aspects, we describe a fundamental SNR trade-off, according to which operating at the right noise level significantly enhances the system performance. In this respect, the optimal transmit SNR for the channel estimation is shown to decrease as the pilot length increases.

Outline. The rest of the paper is structured as follows. Section II introduces the system model with 1-bit ADCs. Sections III and IV present our performance analysis results on the channel estimation and the data detection, respectively, each including dedicated numerical results and discussions. Finally, Section V summarizes our contributions and draws some concluding remarks.

Notation. $\mathbf{A} = (A_{m,n})$ specifies that $A_{m,n}$ is the (m, n)th entry of matrix $\mathbf{A}$; likewise, $\mathbf{a} = (a_n)$ specifies that a_n

---

[1]This second part of the paper complements the results of [1] with a detailed analysis of the normalized variance of the estimated symbols (also by means of tractable upper bounds), while the first part is entirely new.

is the nth entry of vector $\mathbf{a}$. $(\cdot)^{\mathrm{T}}$, $(\cdot)^{\mathrm{H}}$, and $(\cdot)^*$ represent the transpose, Hermitian transpose, and conjugate operators, respectively. $\mathrm{Re}[\cdot]$ and $\mathrm{Im}[\cdot]$ denote the real part and imaginary part operators, respectively, whereas j is the imaginary unit. $\mathbb{E}[\cdot]$ and $\mathbb{V}[\cdot]$ are the expectation and variance operators, respectively. $\mathbf{I}_N$, $\mathbf{0}_N$, and $\mathbf{1}_N$ denote the N-dimensional identity matrix, all-zero vector, and all-one vector, respectively. $\mathrm{Diag}(\cdot)$ produces a diagonal matrix with the entries of the vector argument on the diagonal, $\mathrm{sgn}(\cdot)$ is the sign function, and $\mathrm{vec}[\cdot]$ is the vectorization operator. The Kronecker product is denoted by $\otimes$ and $\{\cdot\}$ is used to represent sets. Lastly, $\mathcal{CN}(0,1)$ is the complex normal distribution with zero mean and unit variance, whereas $\mathcal{N}(\mathbf{0}_N, \boldsymbol{\Sigma})$ is the real N-variate normal distributions with zero mean and covariance matrix $\boldsymbol{\Sigma}$.

II. SYSTEM MODEL

Consider the scenario depicted in Fig. 1, where a BS with M antennas serves K single-antenna UEs in the uplink. Let $\mathbf{H} \triangleq (H_{m,k}) \in \mathbb{C}^{M \times K}$ denote the uplink channel matrix: assuming i.i.d. Rayleigh fading (as, e.g., in [8], [10]–[12]), the entries of $\mathbf{H}$ are distributed independently as $\mathcal{CN}(0,1)$. Each UE transmits with power ρ and the additive white Gaussian noise (AWGN) at the BS has unit variance: thus, ρ can be interpreted as the transmit SNR. Note that the same transmit SNR is assumed for the two phases of channel estimation and uplink data transmission. Each BS antenna is connected to two 1-bit ADCs, one for the in-phase component and one for the quadrature component of the receive signal. Therefore, according to [10], we introduce the 1-bit quantization function $Q(\cdot): \mathbb{C}^{L \times N} \to \mathcal{Q}$, with

$$Q(\mathbf{A}) \triangleq \sqrt{\frac{\rho K + 1}{2}} \Big(\mathrm{sgn}\big(\mathrm{Re}[\mathbf{A}]\big) + j\, \mathrm{sgn}\big(\mathrm{Im}[\mathbf{A}]\big) \Big) \quad (1)$$

and where $\mathcal{Q} \triangleq \sqrt{\frac{\rho K + 1}{2}} \{\pm 1 \pm j\}^{L \times N}$ is the set containing the scaled symbols of the quadrature phase-shift keying (QPSK) constellation.

A. Channel Estimation

In the channel estimation phase, the UEs simultaneously transmit their uplink pilots of length τ. Let $\mathbf{P} \triangleq (P_{u,k}) \in \mathbb{C}^{\tau \times K}$ denote the pilot matrix whose columns correspond to the pilots used by the UEs, with $|P_{u,k}|^2 = 1, \forall u, k$. We assume $\tau \geq K$ and orthogonal pilots among the UEs, so that $\mathbf{P}^{\mathrm{H}}\mathbf{P} = \tau \mathbf{I}_K$. Hence, the receive signal at the BS prior to quantization is given by

$$\mathbf{Y}_{\mathrm{p}} \triangleq \sqrt{\rho}\mathbf{H}\mathbf{P}^{\mathrm{H}} + \mathbf{Z}_{\mathrm{p}} \in \mathbb{C}^{M \times \tau} \quad (2)$$

where $\mathbf{Z}_{\mathrm{p}} \triangleq (Z_{m,u}) \in \mathbb{C}^{M \times \tau}$ is the AWGN term with entries distributed as $\mathcal{CN}(0,1)$. Then, at the output of the ADCs, we have

$$\mathbf{R}_{\mathrm{p}} \triangleq Q(\mathbf{Y}_{\mathrm{p}}) \in \mathbb{C}^{M \times \tau} \quad (3)$$

with $\mathbf{R}_{\mathrm{p}} = (R_{m,u})$, which is used by the BS to estimate $\mathbf{H}$. Some comments are in order. First, correlating the quantized receive signal $\mathbf{R}_{\mathrm{p}}$ in (3) with $\mathbf{P}$, as done in (5) below,

results in residual pilot contamination even when the pilots are orthogonal (see, e.g., [16]). Second, the pilots should be preferably chosen such that their entries span an interval $\left[\eta, \eta + \frac{\pi}{2}\right]$, with $\eta \in [0, 2\pi]$, so as to accurately estimate the phases (especially at high SNR). This is explained in Appendix A, which provides a detailed discussion on the channel estimation with 1-bit ADCs.

The LS estimator for 1-bit ADCs, which correlates the quantized receive signal $\mathbf{R}_{\mathrm{p}}$ in (3) with $\mathbf{P}$, was first presented in [31]. Then, a linear MMSE estimator based on the Bussgang decomposition (see [20]), which we refer to as BLM estimator, was proposed in [8]. According to this, $\underline{\mathbf{h}} \triangleq \mathrm{vec}[\mathbf{H}] \in \mathbb{C}^{MK \times 1}$ is estimated as[2]

$$\underline{\hat{\mathbf{h}}}_{\mathrm{BLM}} \triangleq \sqrt{\frac{2}{\pi}}\rho \tilde{\mathbf{P}}^{\mathrm{T}} \boldsymbol{\Sigma}_{\mathrm{p}}^{-1} \underline{\mathbf{r}}_{\mathrm{p}} \in \mathbb{C}^{MK \times 1} \quad (4)$$

with $\tilde{\mathbf{P}} \triangleq \mathbf{P} \otimes \mathbf{I}_M \in \mathbb{C}^{M\tau \times MK}$, $\underline{\mathbf{r}}_{\mathrm{p}} \triangleq \mathrm{vec}[\mathbf{R}_{\mathrm{p}}] \in \mathbb{C}^{M\tau \times 1}$, and where $\boldsymbol{\Sigma}_{\mathrm{p}} \triangleq \mathbb{E}[\underline{\mathbf{r}}_{\mathrm{p}}\underline{\mathbf{r}}_{\mathrm{p}}^{\mathrm{H}}] \in \mathbb{C}^{M\tau \times M\tau}$ denotes the covariance matrix of $\underline{\mathbf{r}}_{\mathrm{p}}$. A linear estimator with a simpler structure can be obtained from (4) by ignoring the temporal correlation of the quantization distortion, which implies that the off-diagonal entries of $\boldsymbol{\Sigma}_{\mathrm{p}}$ are zero. Such a scaled LS estimator was proposed in [16] (and later extended to the case of multi-bit ADCs in [10]), whereby $\mathbf{H}$ is estimated as

$$\hat{\mathbf{H}}_{\mathrm{SLS}} \triangleq \sqrt{\Psi}\mathbf{R}_{\mathrm{p}}\mathbf{P} \in \mathbb{C}^{M \times K} \quad (5)$$

where the scaling factor Ψ (common for all the UEs) is defined as

$$\Psi \triangleq \frac{2}{\pi}\rho \left(\frac{2}{\pi}\rho(\tau - K) + \rho K + 1 \right)^{-2}. \quad (6)$$

We point out that Ψ in (6) implicitly depends on the channel distribution. Note that, when $\tau = K$, (5) coincides with (4) since $\boldsymbol{\Sigma}_{\mathrm{p}} = (\rho K + 1)\mathbf{I}_{MK}$; otherwise, (5) accurately approximates (4) at low SNR or when K is large [10]. In Section III, we analyze the performance of the BLM estimator in (4) and of the class of scaled LS estimators, such as the one in (5). Moreover, we highlight the relationship between these two in the case of i.i.d. Rayleigh fading channels among the UEs.

B. Uplink Data Transmission

Let $x_k \in \mathbb{C}$ be the transmit symbol of UE k, with $\mathbb{E}[|x_k|^2] = 1$ and $\mathbf{x} \triangleq (x_k) \in \mathbb{C}^{K \times 1}$. The receive signal at the BS prior to quantization is given by

$$\mathbf{y} \triangleq \sqrt{\rho}\mathbf{H}\mathbf{x} + \mathbf{z} \in \mathbb{C}^{M \times 1} \quad (7)$$

where $\mathbf{z} \triangleq (z_m) \in \mathbb{C}^{M \times 1}$ is the AWGN term with entries distributed as $\mathcal{CN}(0,1)$. Then, at the output of the ADCs, we have

$$\mathbf{r} \triangleq Q(\mathbf{y}) \in \mathbb{C}^{M \times 1} \quad (8)$$

and the BS obtains a soft estimate of $\mathbf{x}$ as

$$\hat{\mathbf{x}} \triangleq \mathbf{V}^{\mathrm{H}}\mathbf{r} \in \mathbb{C}^{K \times 1} \quad (9)$$

[2]Note that, in the case of correlated channels, the BLM estimator in (4) is given by [8, Eq. (60)] and the channel covariance matrix is also embedded into $\boldsymbol{\Sigma}_{\mathrm{p}}$.

where $\mathbf{V} \in \mathbb{C}^{M \times K}$ is the combining matrix adopted at the BS. Finally, the data detection process associates each estimated symbol to a transmit symbol, e.g., via MDD. In Section IV, we focus on characterizing the statistical properties of the estimated symbols when MRC is adopted at the BS.

III. CHANNEL ESTIMATION WITH 1-BIT ADCs

In this section, we are interested in characterizing the performance of the channel estimation with respect to the different parameters when 1-bit ADCs are adopted at each BS antenna (see Section II-A). In doing so, we consider the BLM estimator in (4) and the class of scaled LS estimators, such as the one in (5).

A. MSE of the Channel Estimation

The (normalized) MSE of the channel estimation when the BLM estimator is used is given by

$$\text{MSE}_{\text{BLM}} \triangleq \frac{1}{MK} \mathbb{E}\big[\|\hat{\underline{\mathbf{h}}}_{\text{BLM}} - \underline{\mathbf{h}}\|^2\big] \tag{10}$$

with $\hat{\underline{\mathbf{h}}}_{\text{BLM}}$ defined in (4). In [8, Eq. (17)], the closed-form expression of (10) was derived for the case of $\tau = K$, which gives

$$\text{MSE}_{\text{BLM}} = 1 - \frac{2}{\pi} \frac{\rho K}{\rho K + 1}. \tag{11}$$

Note that the above expression is lower bounded by $1 - \frac{2}{\pi} \simeq 0.363$, which is achieved in the limit of $\rho K \to \infty$. Hence, in realistic scenarios (especially for small values of K), using a pilot length that is equal to the number of UEs results in quite inaccurate channel estimates. In general, τ should be sufficiently large to compensate for the low granularity of the ADCs, as detailed in Appendix A. For this reason, we assume that the pilot matrix $\mathbf{P}$ is chosen such that $\mathbf{P}\mathbf{P}^{\text{H}}$ is circulant[3] and derive a closed-form expression of (10) that is valid for any value of τ and K.

Theorem 1: Suppose that the BLM estimator in (4) is used and that $\mathbf{P}$ is chosen such that $\mathbf{P}\mathbf{P}^{\text{H}}$ is circulant. Then, the MSE of the channel estimation in (10) is given by

$$\text{MSE}_{\text{BLM}} = 1 - \frac{2}{\pi} \frac{\rho \tau^2}{\rho K + 1} \frac{1}{K} \sum_{k=1}^{K} \frac{1}{\tau + \delta_k} \tag{12}$$

where we have defined

$$\delta_k \triangleq \sum_{u \neq v} \bigg(\text{Re}[P_{u,k}^* P_{v,k}] \Omega \bigg(\frac{\rho \sum_{i=1}^{K} \text{Re}[P_{u,i} P_{v,i}^*]}{\rho K + 1} \bigg)$$
$$- \text{Im}[P_{u,k}^* P_{v,k}] \Omega \bigg(\frac{\rho \sum_{i=1}^{K} \text{Im}[P_{u,i} P_{v,i}^*]}{\rho K + 1} \bigg) \bigg) \tag{13}$$

with

$$\Omega(w) \triangleq \frac{2}{\pi} \arcsin(w). \tag{14}$$

Proof: See Appendix B. ∎

The result of Theorem 1 enables a precise characterization of the performance of the BLM estimator with respect to the transmit SNR ρ, the number of UEs K, and the pilot length τ. The parameter δ_k in (13) is roughly proportional to $\tau(\tau - 1)$ and, for a fixed τ, decreases with K. When $\tau = K$, (12) recovers the expression in (11): in fact, $\tau = K$ implies $\delta_k = 0$, $\forall k$, since $\sum_{i=1}^{K} P_{u,i} P_{v,i}^* = 0$, $\forall u \neq v$. Moreover, the choice of the pilot matrix $\mathbf{P}$ affects the MSE of the channel estimation through the parameters $\{\delta_1, \ldots, \delta_K\}$. Lastly, we point out that, if $\mathbf{P}\mathbf{P}^{\text{H}}$ is not circulant, MSE_{BLM} can be computed via the more involved and less insightful expression in (68) (see Appendix B).

We now show that, in the case of i.i.d. Rayleigh fading channels among the UEs and when $\mathbf{P}\mathbf{P}^{\text{H}}$ is circulant, the BLM estimator can be simplified as a scaled LS estimator with UE-specific scaling factors.

Corollary 1: Suppose that $\mathbf{P}$ is chosen such that $\mathbf{P}\mathbf{P}^{\text{H}}$ is circulant. Then, the BLM estimator in (4) can be simplified as

$$\hat{\mathbf{H}}_{\text{BLM}} \triangleq \mathbf{R}_{\text{p}} \mathbf{P} \mathbf{\Psi}^{\frac{1}{2}} \in \mathbb{C}^{M \times K} \tag{15}$$

where we have defined $\mathbf{\Psi} \triangleq \text{Diag}\big([\Psi_1, \ldots, \Psi_K]\big) \in \mathbb{R}^{K \times K}$, with

$$\Psi_k \triangleq \frac{2}{\pi} \frac{\rho \tau^2}{(\rho K + 1)^2 (\tau + \delta_k)^2}. \tag{16}$$

Proof: See Appendix C. ∎

The result of Corollary 1 states that, under the above assumptions, the BLM estimator can be implemented in a way that avoids the inversion of and the multiplication with the $M\tau$-dimensional matrix $\mathbf{\Sigma}_{\text{p}}$, where $M\tau$ can be quite large at mmWave and sub-THz frequencies. In particular, $\mathbf{P}$ diagonalizes $\mathbf{\Sigma}_{\text{p}}$ when $\mathbf{P}\mathbf{P}^{\text{H}}$ is circulant, which greatly simplifies the structure of (4). Lastly, we point out that, in the case of correlated channels, Corollary 1 does not generally hold as the channel covariance matrix is embedded into $\mathbf{\Sigma}_{\text{p}}$ and the latter is not diagonalized by $\mathbf{P}$.

Let us move our focus to the class of scaled LS estimators, such as the one in (5); recall that, unlike the simplified expression of the BLM estimator in (15), scaled LS estimators are characterized by a common scaling factor for all the UEs. In this regard, we first derive the closed-form expression of the MSE of the channel estimation for an arbitrary scaling factor.

Theorem 2: Suppose that the scaled LS estimator in (5) is used with arbitrary Ψ. Then, the MSE of the channel estimation is given by

$$\text{MSE}_{\text{SLS}} \triangleq \frac{1}{MK} \mathbb{E}\big[\|\hat{\mathbf{H}}_{\text{SLS}} - \mathbf{H}\|_{\text{F}}^2\big] \tag{17}$$

$$= 1 + (\rho K + 1)\Psi(\tau + \Delta) - 2\sqrt{\frac{2}{\pi} \rho \Psi \tau} \tag{18}$$

where we have defined

$$\Delta \triangleq \frac{1}{K} \sum_{k=1}^{K} \delta_k \tag{19}$$

with δ_k defined in (13).

---

[3]For example, this condition is satisfied when $\mathbf{P}$ is a partial discrete Fourier transform (DFT) matrix, i.e., $\mathbf{P}$ is composed of any K columns of the τ-dimensional DFT matrix.

Proof: The expression in (18) is obtained from the proof of Corollary 1 by replacing the UE-specific scaling factors $\{\Psi_1, \ldots, \Psi_K\}$ in (99) with the common scaling factor Ψ. ∎ In particular, when Ψ defined in (6) is used, (18) becomes

$$\text{MSE}_{\text{SLS}} = 1 - \frac{2}{\pi}\rho\left(\frac{2}{\pi}\rho(\tau - K) + \rho K + 1\right)^{-2}$$
$$\times \left(\frac{4}{\pi}\rho\tau(\tau - K) + (\rho K + 1)(\tau - \Delta)\right). \quad (20)$$

Note that, when $\tau = K$, (20) recovers the expression of the MSE of the BLM estimator in (11).

Now, if we consider the scaling factor Ψ as a tuning parameter, we can minimize MSE_{SLS} in (18) by optimizing over Ψ. As a result, we obtain the optimal estimator within the class of scaled LS estimators.

Corollary 2: Suppose that the scaled LS estimator

$$\hat{\mathbf{H}}'_{\text{SLS}} \triangleq \sqrt{\Psi'}\mathbf{R}_{\text{p}}\mathbf{P} \in \mathbb{C}^{M \times K} \quad (21)$$

is used, where Ψ' is obtained by minimizing (18) with respect to Ψ and is defined as

$$\Psi' \triangleq \frac{2}{\pi}\frac{\rho\tau^2}{(\rho K + 1)^2(\tau + \Delta)^2} \quad (22)$$

with Δ defined in (19). Then, the MSE of the channel estimation is given by

$$\text{MSE}'_{\text{SLS}} \triangleq \frac{1}{MK}\mathbb{E}\left[\|\hat{\mathbf{H}}'_{\text{SLS}} - \mathbf{H}\|_{\text{F}}^2\right] \quad (23)$$
$$= 1 - \frac{2}{\pi}\frac{\rho\tau^2}{(\rho K + 1)(\tau + \Delta)} \quad (24)$$
$$\leq (20). \quad (25)$$

Proof: Since (18) is a convex function of Ψ, Ψ' in (22) can be obtained by setting $\frac{\text{d}}{\text{d}\Psi}(18) = 0$. Then, replacing Ψ with Ψ' in (18) yields the expression in (24). ∎ Note that Ψ' in (22) implicitly depends on the channel distribution as does Ψ in (6). The result of Corollary 2 shows that the optimal scaled LS estimator (in terms of MSE of the channel estimation) is not the one that simply ignores the temporal correlation of the quantization distortion from the BLM estimator (see (5)–(6)); instead, a simple optimization over the scaling factor can significantly improve the channel estimation accuracy. When $\tau = K$, the optimal scaled LS estimator in (21) coincides with the estimator in (5), and, in turn, with the BLM estimator in (4): in fact, $\tau = K$ implies that Ψ' in (22) reduces to Ψ in (6) and (24) recovers the expression in (11). On the other hand, the estimator in (21) shall be always preferred to the estimator in (5) when $\tau > K$ and the performance gap between the two widens with $\tau - K$. The improved performance of (21) over (5) is also suggested by the resemblance between MSE_{BLM} in (12) and MSE'_{SLS} in (24), where the latter can be obtained from the former by replacing δ_k with Δ, $\forall k$. Remarkably, in Section III-C, we show that the optimal scaled LS estimator in (21) entails a negligible performance loss with respect to the BLM estimator.

It is of particular interest to study the asymptotic behavior of the MSE of the channel estimation at high SNR.

Corollary 3: From Theorems 1 and 2 and from Corollary 2, in the limit of $\rho \to \infty$, we have

$$\lim_{\rho \to \infty} \text{MSE}_{\text{BLM}} = 1 - \frac{2}{\pi}\frac{\tau^2}{K^2}\sum_{k=1}^{K}\frac{1}{\tau + \bar{\delta}_k}, \quad (26)$$

$$\lim_{\rho \to \infty} \text{MSE}_{\text{SLS}} = 1 - \frac{2}{\pi}\left(\frac{2}{\pi}(\tau - K) + K\right)^{-2}$$
$$\times \left(\frac{4}{\pi}\tau(\tau - K) + K(\tau - \bar{\Delta})\right), \quad (27)$$

$$\lim_{\rho \to \infty} \text{MSE}'_{\text{SLS}} = 1 - \frac{2}{\pi}\frac{\tau^2}{K(\tau + \bar{\Delta})} \quad (28)$$

where we have defined

$$\bar{\delta}_k \triangleq \sum_{u \neq v}\left(\text{Re}[P_{u,k}^*P_{v,k}]\Omega\left(\frac{\sum_{i=1}^{K}\text{Re}[P_{u,i}P_{v,i}^*]}{K}\right)\right.$$
$$\left. - \text{Im}[P_{u,k}^*P_{v,k}]\Omega\left(\frac{\sum_{i=1}^{K}\text{Im}[P_{u,i}P_{v,i}^*]}{K}\right)\right) \quad (29)$$

and

$$\bar{\Delta} \triangleq \frac{1}{K}\sum_{k=1}^{K}\bar{\delta}_k. \quad (30)$$

The results of Corollary 3 show that arbitrarily increasing the transmit SNR is detrimental for the performance of the channel estimation since the right amount of noise is necessary to recover the difference in amplitude between channel entries (see Appendix A). This is in sheer contrast with the case of infinite-resolution ADCs, where boosting ρ produces the same beneficial noise-averaging effect as increasing τ. In the next section, we also discuss the asymptotic behavior at low SNR.

B. Tractable Upper Bounds

The MSE expressions derived so far depend on the specific pilot choice through the parameters $\{\delta_1, \ldots, \delta_K\}$ in (13) or Δ in (19). To gain more practical insights, we now consider the single-UE case (i.e., $K = 1$) and derive tractable upper bounds that are independent of the pilot choice. We begin by pointing out that, when $K = 1$, we have $\delta_k = \Delta$ and, thus, MSE_{BLM} in (12) is equal to MSE'_{SLS} in (24). Let $\mathbf{p} \triangleq (p_u) \in \mathbb{C}^{\tau \times 1}$ denote the pilot used by the UE. In this setting, Δ in (19) can be simplified as

$$\Delta = \sum_{u \neq v}\left(\text{Re}[p_u^*p_v]\Omega\left(\frac{\rho\text{Re}[p_up_v^*]}{\rho + 1}\right)\right.$$
$$\left. - \text{Im}[p_u^*p_v]\Omega\left(\frac{\rho\text{Im}[p_up_v^*]}{\rho + 1}\right)\right) \quad (31)$$
$$\leq \tau(\tau - 1)\Omega\left(\frac{\rho}{\rho + 1}\right) \quad (32)$$

where the upper bound in (32) is obtained by fixing $\mathbf{p}$ such that $p_u \in \{\pm\beta, \pm j\beta\}$, $\forall u$, with $\beta \in \mathbb{C}$ and $|\beta|^2 = 1$; in the rest of the paper, when referring to this case, we will simply use $\mathbf{p} = \mathbf{1}_\tau$. Indeed, for a given pilot length τ, such a structure of $\mathbf{p}$ represents the worst possible pilot choice since it maximizes the MSE of the channel estimation in (12) and (20). As detailed in Appendix A, this effect is particularly

detrimental at high SNR and, in the limit of $\rho \to \infty$, each channel entry is reduced to a scaled symbol of the QPSK constellation regardless of the value of τ.

Hence, plugging (32) into (12) and (26) yields

$$\text{MSE}_{\text{BLM}} = 1 - \frac{2}{\pi} \frac{\rho\tau}{(\rho+1)\left(1 + (\tau-1)\Omega\left(\frac{\rho}{\rho+1}\right)\right)} \quad (33)$$

and

$$\lim_{\rho \to \infty} \text{MSE}_{\text{BLM}} = 1 - \frac{2}{\pi} \quad (34)$$

respectively. In addition, considering (33) in the limit of $\tau \to \infty$, we have

$$\lim_{\tau \to \infty} \text{MSE}_{\text{BLM}} = 1 - \frac{2}{\pi} \frac{\rho}{(\rho+1)\Omega\left(\frac{\rho}{\rho+1}\right)}. \quad (35)$$

Likewise, plugging (32) into (20) and (27) yields

$$\text{MSE}_{\text{SLS}} = 1 - \frac{2}{\pi}\rho\tau\left(\frac{2}{\pi}\rho(\tau-1) + \rho + 1\right)^{-2}\left(\frac{4}{\pi}\rho(\tau-1)\right.$$
$$\left. + (\rho+1)\left(1 - (\tau-1)\Omega\left(\frac{\rho}{\rho+1}\right)\right)\right) \quad (36)$$

and

$$\lim_{\rho \to \infty} \text{MSE}_{\text{SLS}} = 1 - \frac{2}{\pi}\tau\left(\frac{2}{\pi}(\tau-1) + 1\right)^{-2}\left(\frac{4}{\pi}(\tau-1) - \tau + 2\right) \quad (37)$$

respectively. Moreover, considering (36) in the limit of $\tau \to \infty$, we have

$$\lim_{\tau \to \infty} \text{MSE}_{\text{SLS}} = \frac{\pi}{2}\frac{\rho+1}{\rho}\Omega\left(\frac{\rho}{\rho+1}\right) - 1. \quad (38)$$

Some comments are in order. First, when $\tau = 1$, both (33) and (36) recover the expression in (11) with $K = 1$. Second, (34) does not depend on τ since, in the absence of noise, estimating the channel repeatedly over the same pilot symbol does not bring any benefit. Third, it can be demonstrated that (33) and (36) are quasiconvex functions of ρ and, as such, they have a unique minimum. This defines a clear SNR trade-off, according to which operating at the right noise level enhances the channel estimation accuracy. In particular, as discussed in Appendix A, we have that:

- At low SNR, the channel estimates are corrupted by the strong noise;
- At high SNR, the difference in amplitude between channel entries cannot be recovered.

In general, when $\tau > 1$, the value of ρ that minimizes (33), denoted by $\rho^\star$, satisfies

$$\frac{2}{\pi}\frac{\rho^\star}{\sqrt{1+2\rho^\star}} - \Omega\left(\frac{\rho^\star}{\rho^\star+1}\right) = \frac{1}{\tau-1}. \quad (39)$$

Since the left-hand side of (39) monotonically increases with the transmit SNR, $\rho^\star$ decreases as the left-hand side of (39) decreases, i.e., as τ increases. This means that using longer pilots allows to operate at lower SNR as the noise can be averaged out more efficiently. This interdependence between ρ and τ can be also observed from (35): in the limit of $\tau \to \infty$, since $\lim_{w \to 0}\frac{w}{\arcsin(w)} = 1$, it follows that $\text{MSE}' \to 0$ as $\rho \to$

0. Lastly, it is shown in Section III-C that the upper bounds obtained by fixing $\mathbf{p} = \mathbf{1}_\tau$ are remarkably tight at low SNR and up to the region around the optimal value of ρ. Therefore, the above observations also apply to the general case.

C. Numerical Results and Discussion

We now focus on the performance evaluation of the channel estimation with 1-bit ADCs with respect to the different parameters based on the analytical results presented in Sections III-A and III-B. In this regard, when $K > 1$, we choose a pilot matrix $\mathbf{P}$ composed of the first K columns of the τ-dimensional DFT matrix. On the other hand, when $K = 1$, we use the pilots $\mathbf{p} = \mathbf{p}^\star$ and $\mathbf{p} = \mathbf{1}_\tau$, where

$$\mathbf{p}^\star \triangleq [1, e^{-j\frac{\pi}{2\tau}}, e^{-j2\frac{\pi}{2\tau}}, \ldots, e^{-j(\tau-1)\frac{\pi}{2\tau}}]^{\text{T}} \in \mathbb{C}^{\tau \times 1} \quad (40)$$

denotes the vector whose entries are equispaced on the first quadrant of the unit circle: these represent the best and the worst possible pilot choices, respectively (see Section III-B and Appendix A for more details on the pilot choice). We thus consider the expressions of MSE_{BLM} derived in (12), resulting from the BLM estimator in (4) (see [8]), MSE_{SLS} derived in (20), resulting from the scaled LS estimator in (5) (see [16]), and MSE'_{SLS} derived in (24), resulting from the optimal scaled LS estimator in (21). These are compared with Monte Carlo simulations with 10^6 independent channel realizations. For the latter, we fix $M = K$ to generate the channel matrices, although the value of M does not affect the analytical and numerical results in any way.

Fig. 2 illustrates the MSE of the channel estimation against the transmit SNR ρ, also including the asymptotic MSE expressions in (26)–(28). Fig. 2(a) considers $K = 4$ and $\tau = 32$ and shows that MSE_{BLM} is 5.3% lower than MSE_{SLS} at high SNR. Remarkably, the performance loss associated with MSE'_{SLS} with respect to MSE_{BLM} is negligible and can only be noticed by examining the relative MSE difference in Fig. 2(b), which reaches its maximum of about 4×10^{-4} in the limit of $\rho \to \infty$. Hence, in the case of i.i.d. Rayleigh fading channels among the UEs, the scaled LS estimator with a common optimized scaling factor for all the UEs essentially achieves the same accuracy as the BLM estimator. Lastly, we highlight the SNR trade-off described in Sections III-A and III-B as well as in Appendix A, whereby the MSE of the channel estimation exhibits a valley at about $\rho = 3$ dB. Fig. 2(c) considers $\tau = 128$ and different values of K. Here, the gap between MSE_{BLM} and MSE_{SLS} widens as $\tau - K$ increases, reaching 8.5% for $K = 8$ at high SNR. In this respect, at high SNR, MSE_{SLS} for $K = 8$ surpasses its counterpart for $K = 16$: in fact, (26)–(28) are not monotonically increasing with K due to the fact that $\bar{\delta}_k$ in (29) is a decreasing function of K. Moreover, the SNR trade-off appears more evident for small values of K. Fig. 2(d) considers the single-UE case, showing that the upper bounds in (33) and (36) obtained by fixing $\mathbf{p} = \mathbf{1}_\tau$ are remarkably tight at low SNR and up to the region around the optimal transmit SNR. Note that the optimal value of ρ with $\mathbf{p} = \mathbf{1}_\tau$ satisfies the condition in (39) and gives an accurate approximation of the optimal value of ρ with $\mathbf{p} = \mathbf{p}^\star$.

Fig. 3 plots the MSE of the channel estimation against the pilot length τ. The transmit SNR is fixed to $\rho = 10$ dB in
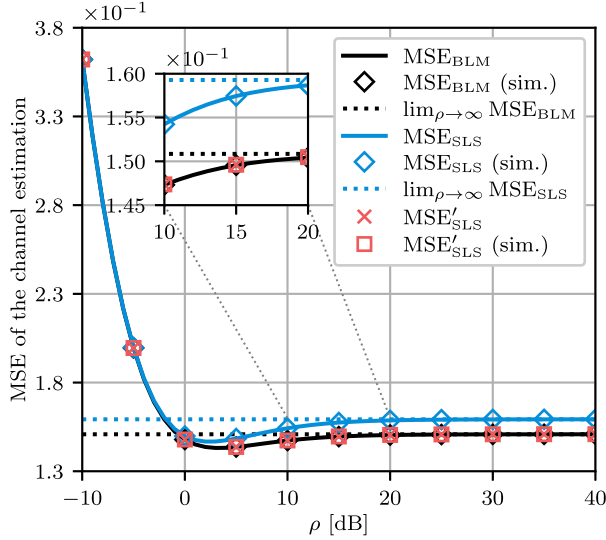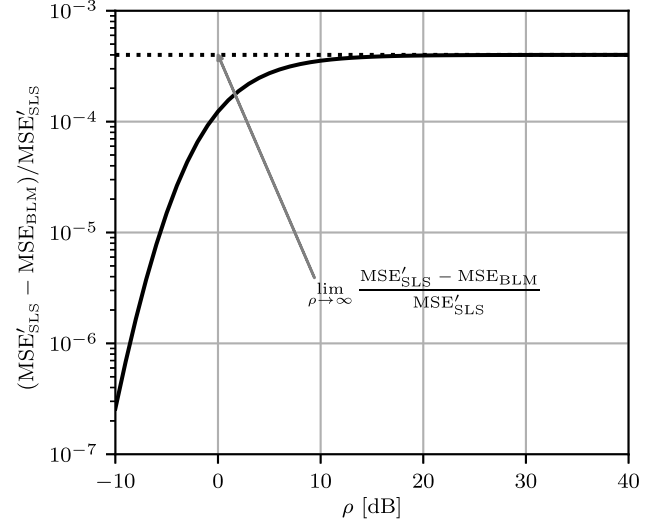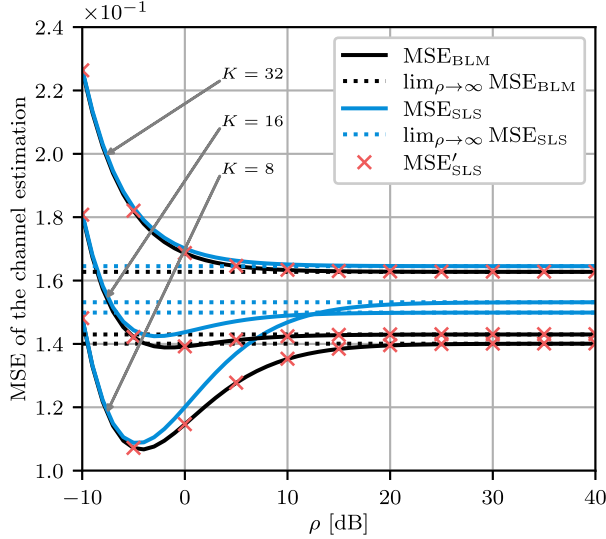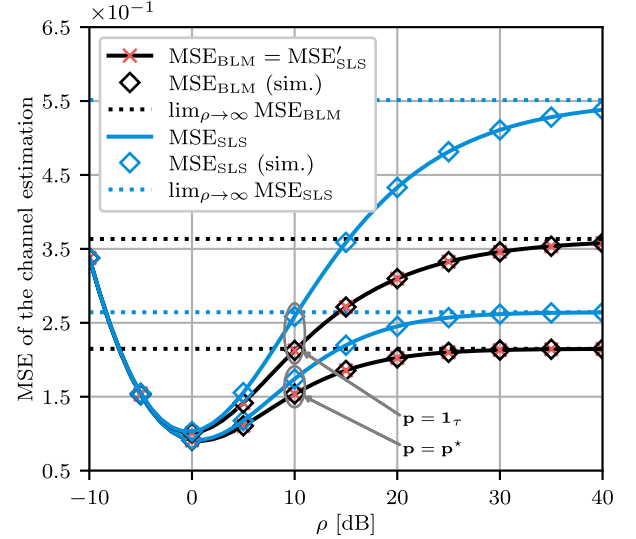
Fig. 2. MSE of the channel estimation against the transmit SNR.

Fig. 3(a)–(c), whereas Fig. 3(d) considers the optimized transmit SNR (we recall that ρ should be reduced as τ increases to enhance the channel estimation accuracy). Fig. 3(a) considers $K = 4$ and shows that $\mathrm{MSE_{BLM}}$ is 10% lower than $\mathrm{MSE_{SLS}}$ at $\tau = 128$. Furthermore, as in Fig. 2, $\mathrm{MSE'_{SLS}}$ closely matches $\mathrm{MSE_{BLM}}$ for any value of τ, which means that the optimal scaled LS estimator essentially achieves the same accuracy as the BLM estimator. Fig. 3(b) considers different values of K, showing that the gap between $\mathrm{MSE_{BLM}}$ and $\mathrm{MSE_{SLS}}$ widens as $\tau - K$ increases and reaches 7.8% for $K = 8$ and $\tau = 128$. Fig. 3(c) examines the case where the number of UEs grows together with the pilot length. In this setting, the gap between $\mathrm{MSE_{BLM}}$ and $\mathrm{MSE_{SLS}}$ is roughly constant and increases with the ratio $\frac{\tau}{K}$, reaching about 5% for $\frac{\tau}{K} = 8$. Lastly, Fig. 3(d) considers the single-UE case and the upper bound on $\mathrm{MSE_{BLM}}$ in (33) obtained by fixing $\mathbf{p} = \mathbf{1}_\tau$, which is optimized over the transmit SNR for each τ. As discussed in Section III-B, the optimal value of ρ satisfies the condition in (39) and decreases as τ increases.

IV. DATA DETECTION WITH 1-BIT ADCs AND MRC

In this section, we are interested in characterizing the performance of the data detection with respect to the different parameters when 1-bit ADCs are adopted at each BS antenna (see Section II-B). In this regard, we consider the scenario where the BS uses the BLM estimator in (4) in the channel estimation phase and the MRC receiver in the data detection phase (see also [8]). Assuming that $\mathbf{PP}^{\mathrm{H}}$ is circulant and building on Corollary 1, the combining matrix is given by $\mathbf{V} = \hat{\mathbf{H}}_{\mathrm{BLM}}$ and we can write the soft estimate in (9) as

$$\hat{\mathbf{x}} = \mathbf{\Psi}^{\frac{1}{2}} \mathbf{P}^{\mathrm{H}} \mathbf{R}_{\mathrm{p}}^{\mathrm{H}} \mathbf{r} \tag{41}$$

$$= \frac{\rho K + 1}{2} \mathbf{\Psi}^{\frac{1}{2}} \mathbf{P}^{\mathrm{H}} \Big(\mathrm{sgn}\big(\mathrm{Re}[\sqrt{\rho}\mathbf{HP}^{\mathrm{H}} + \mathbf{Z}_{\mathrm{p}}]\big)$$
$$+ j\, \mathrm{sgn}\big(\mathrm{Im}[\sqrt{\rho}\mathbf{HP}^{\mathrm{H}} + \mathbf{Z}_{\mathrm{p}}]\big) \Big)^{\mathrm{H}}$$
$$\times \Big(\mathrm{sgn}\big(\mathrm{Re}[\sqrt{\rho}\mathbf{Hx} + \mathbf{z}]\big) + j\, \mathrm{sgn}\big(\mathrm{Im}[\sqrt{\rho}\mathbf{Hx} + \mathbf{z}]\big) \Big). \tag{42}$$

(a) $K = 4$, $\rho = 10$ dB.

(b) $K \in \{8, 16, 32\}$, $\rho = 10$ dB.

(c) $K \in \left\{\frac{\tau}{8}, \frac{\tau}{4}, \frac{\tau}{2}\right\}$, $\rho = 10$ dB.
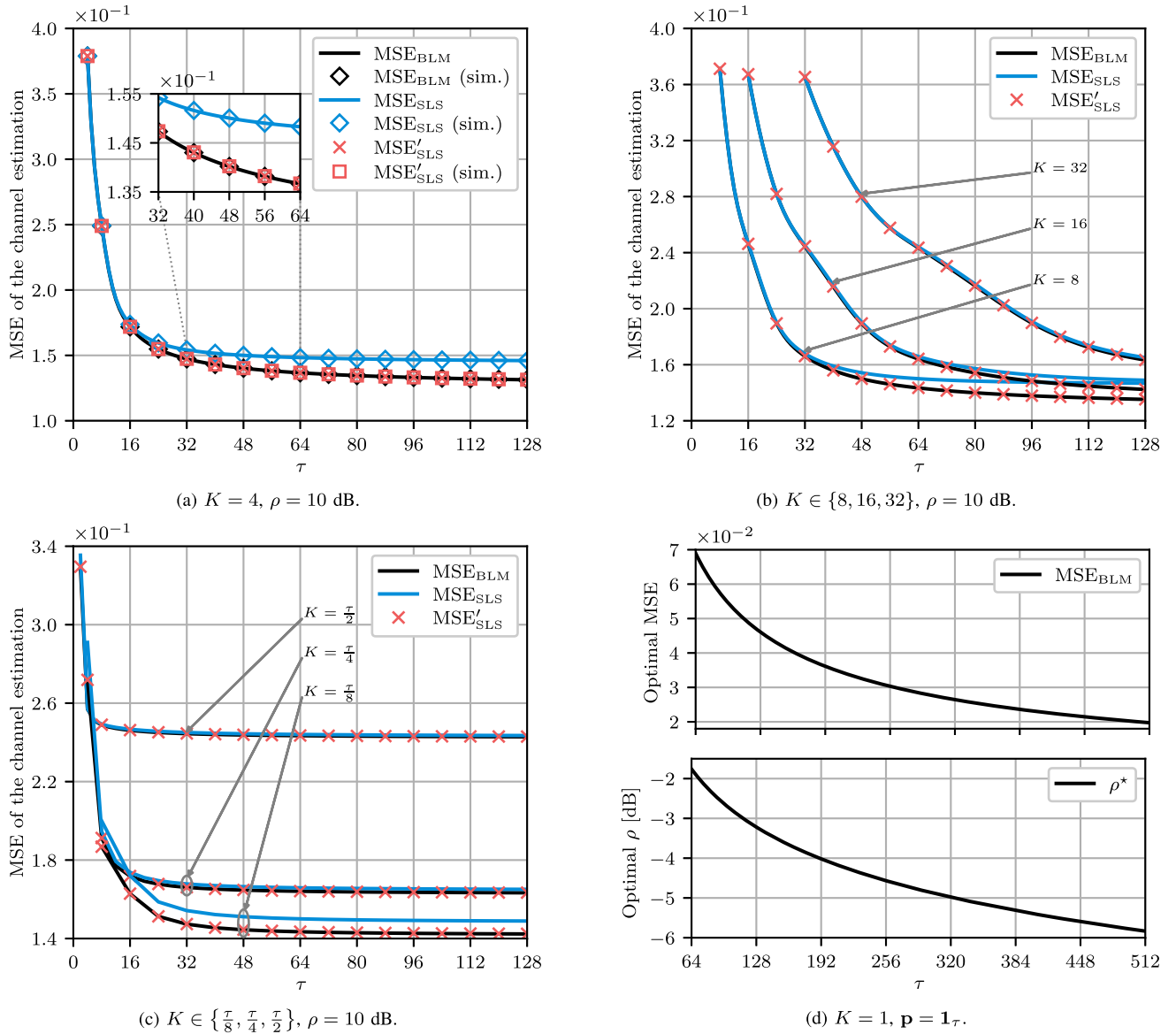
(d) $K = 1$, $\mathbf{p} = \mathbf{1}_\tau$.

Fig. 3.   MSE of the channel estimation against the pilot length.

In the following, we focus on the single-UE case (i.e., $K = 1$) and characterize the statistical properties of the estimated symbols. We recall that the BLM estimator is equivalent to the optimal scaled LS estimator in (21) when $K = 1$, as detailed in Section III-B. Note that, in a multi-UE massive MIMO context with infinite-resolution ADCs, MRC asymptotically becomes the optimal receive strategy as the number of BS antennas increases. However, when the MRC receiver results from the quantized channel estimation, it cannot be perfectly aligned with the channel matrix, resulting in residual multi-UE interference. Hence, the following analysis of the single-UE case does not consider this interference; nonetheless, this can be straightforwardly included at the expense of more involved and less insightful expressions, which is left for future work.

A. Expected Value and Variance of the Estimated Symbols

Let $x \in \mathcal{S}$ be the transmit symbol of the UE, where $\mathcal{S} \triangleq \{s_1, \ldots, s_L\}$ denotes the set of transmit symbols, with $s_\ell \in \mathbb{C}$, $\forall \ell$; for instance, $\mathcal{S}$ may correspond to the QPSK or 16-QAM

constellation. To facilitate the data detection process at the BS, for each transmit symbol $s_\ell \in \mathcal{S}$, we are interested in deriving the closed-form expression of the expected value of the resulting estimated symbol $\hat{s}_\ell$.

Theorem 3: Assuming $K = 1$ and MRC, for each transmit symbol $s_\ell \in \mathcal{S}$, the expected value of the resulting estimated symbol $\hat{s}_\ell$, denoted by $\mathsf{E}_\ell \triangleq \mathbb{E}[\hat{s}_\ell]$, is given by

$$\mathsf{E}_\ell = \sqrt{\frac{2}{\pi}\rho M} \frac{\tau}{\tau + \Delta} \sum_{u=1}^{\tau} p_u^* \left(\Omega\left(\frac{\rho \mathrm{Re}[p_u s_\ell]}{\sqrt{(\rho+1)(\rho|s_\ell|^2+1)}} \right) \right.$$
$$\left. + j\, \Omega\left(\frac{\rho \mathrm{Im}[p_u s_\ell]}{\sqrt{(\rho+1)(\rho|s_\ell|^2+1)}} \right) \right) \quad (43)$$

with Δ given in (31).

Proof: See Appendix D.   ∎

The result of Theorem 3 can be exploited to efficiently implement MDD. In this context, each estimated symbol $\hat{x}$ resulting from transmitting $x \in \mathcal{S}$ can be readily mapped to one of the expected values $\{\mathsf{E}_1, \ldots, \mathsf{E}_L\}$, which are derived as

in (43) without any prior Monte Carlo computation, according to the minimum distance criterion. To further simplify the process and avoid computing the distance between $\hat{x}$ and each E_ℓ, one can construct the Voronoi tessellation based on $\{\mathsf{E}_1, \ldots, \mathsf{E}_L\}$ (and, possibly, other available information) obtaining well-defined detection regions.

Now, for each transmit symbol $s_\ell \in \mathcal{S}$, we derive the closed-form expression of the variance of the resulting estimated symbol $\hat{s}_\ell$.

Theorem 4: Assuming $K = 1$ and MRC, for each transmit symbol $s_\ell \in \mathcal{S}$, the variance of the resulting estimated symbol $\hat{s}_\ell$, denoted by $\mathsf{V}_\ell \triangleq \mathbb{V}[\hat{s}_\ell]$, is given by

$$\mathsf{V}_\ell = \frac{2}{\pi} \rho M \frac{\tau^2}{\tau + \Delta} - \frac{1}{M} |\mathsf{E}_\ell|^2 \qquad (44)$$

with Δ given in (31) and where E_ℓ is derived in closed form in (43).

Proof: See Appendix E. ∎

The result of Theorem 4 quantifies the dispersion of the estimated symbols about their expected value, which results from the 1-bit quantization applied to both the channel estimation (through the MRC receiver) and the uplink data transmission. This dispersion is not isotropic and assumes different shapes for different transmit symbols, as illustrated in Section IV-C (see also [32]). Some additional comments are in order. First, V_ℓ reduces as $|s_\ell|$ increases due to the negative term on the right-hand side of (44): this is somewhat intuitive since the transmit symbols that lie further from the origin are less subject to noise. Second, although V_ℓ increases linearly with the number of BS antennas M, the normalized variance $\frac{\mathsf{V}_\ell}{|\mathsf{E}_\ell|^2}$ (which expresses the relative dispersion of the estimated symbols about their expected value) is inversely proportional to M. Third, the combined results of Theorems 3 and 4 can be exploited to design the set of transmit symbols $\mathcal{S}$ by jointly minimizing the relative dispersion and the overlap between different symbols after the estimation, which is left for future work. Lastly, in the context of MDD via Voronoi tessellation described above, one can utilize the variance derived as in (44) to further refine the detection regions [1].

It is of particular interest to study the asymptotic behavior of the expected value and variance of the estimated symbols at high SNR.

Corollary 4: From Theorems 3 and 4, in the limit of $\rho \to \infty$, we have

$$\lim_{\rho \to \infty} \frac{\mathsf{E}_\ell}{\sqrt{\rho}} = \sqrt{\frac{2}{\pi}} M \frac{\tau}{\tau + \bar{\Delta}} \sum_{u=1}^{\tau} p_u^* \left(\Omega \left(\frac{\mathrm{Re}[p_u s_\ell]}{|s_\ell|} \right) \right.$$
$$\left. + j\, \Omega \left(\frac{\mathrm{Im}[p_u s_\ell]}{|s_\ell|} \right) \right) \quad (45)$$

and

$$\lim_{\rho \to \infty} \frac{\mathsf{V}_\ell}{\rho} = \frac{2}{\pi} M \frac{\tau^2}{\tau + \bar{\Delta}} - \frac{1}{M} \lim_{\rho \to \infty} \frac{|\mathsf{E}_\ell|^2}{\rho} \qquad (46)$$

with $\bar{\Delta}$ defined in (30), which can be simplified for $K = 1$ as

$$\bar{\Delta} = \sum_{u \neq v} \left(\mathrm{Re}[p_u^* p_v] \Omega \left(\mathrm{Re}[p_u p_v^*] \right) - \mathrm{Im}[p_u^* p_v] \Omega \left(\mathrm{Im}[p_u p_v^*] \right) \right). \qquad (47)$$

Corollary 4 formalizes a behavior of the estimated symbols that was observed in [10]. From (45), it emerges that, at high SNR, all the estimated symbols lie on a circle around the origin and their amplitude no longer conveys any information. As a consequence, the estimated symbols resulting from transmit symbols with the same phase become indistinguishable in terms of their expected value, which depends only on $\frac{\mathrm{Re}[s_\ell]}{|s_\ell|}$ and $\frac{\mathrm{Im}[s_\ell]}{|s_\ell|}$. For instance, if $\mathcal{S}$ corresponds to the 16-QAM constellation as in Section IV-C, the inner estimated symbols become indistinguishable from the outer estimated symbols with the same phase. Furthermore, according to (46), these estimated symbols become identical also in terms of variance. In the light of this, blindly minimizing the (normalized) variance of the estimated symbols is not the key to enhancing the system performance. Instead, the variance (which roughly decreases with the transmit SNR) should be minimized alongside the overlap between different symbols after the estimation (which generally increases with the transmit SNR). This determines a clear SNR trade-off, according to which operating at the right noise level enhances the data detection accuracy and thus reduces the SER. In the next section, we also discuss the asymptotic behavior at low SNR.

B. Tractable Upper Bounds

As done in Section III-B for the MSE of the channel estimation, tractable upper bounds on the normalized variance of the estimated symbols, i.e., that do not depend on the specific pilot choice, can be obtained by fixing $\mathbf{p} = \mathbf{1}_\tau$ since such a structure of $\mathbf{p}$ represents the worst possible pilot choice (see Section III-B and Appendix A). Hence, plugging (32) into (44) and (46) yields

$$\frac{\mathsf{V}_\ell}{|\mathsf{E}_\ell|^2} = \frac{1}{M} \frac{1 + (\tau - 1)\Omega\left(\frac{\rho}{\rho+1}\right)}{\tau} \left(\left(\Omega \left(\frac{\rho \mathrm{Re}[s_\ell]}{\sqrt{(\rho+1)(\rho|s_\ell|^2+1)}} \right) \right)^2 \right.$$
$$\left. + \left(\Omega \left(\frac{\rho \mathrm{Im}[s_\ell]}{\sqrt{(\rho+1)(\rho|s_\ell|^2+1)}} \right) \right)^2 \right)^{-1} - \frac{1}{M} \quad (48)$$

and

$$\lim_{\rho \to \infty} \frac{\mathsf{V}_\ell}{|\mathsf{E}_\ell|^2} = \frac{1}{M} \left(\left(\Omega \left(\frac{\mathrm{Re}[s_\ell]}{|s_\ell|} \right) \right)^2 + \left(\Omega \left(\frac{\mathrm{Im}[s_\ell]}{|s_\ell|} \right) \right)^2 \right)^{-1} - \frac{1}{M} \quad (49)$$

respectively. In addition, considering (48) in the limit of $\tau \to \infty$, we have

$$\lim_{\tau \to \infty} \frac{\mathsf{V}_\ell}{|\mathsf{E}_\ell|^2} = \frac{1}{M} \Omega \left(\frac{\rho}{\rho+1} \right) \left(\left(\Omega \left(\frac{\rho \mathrm{Re}[s_\ell]}{\sqrt{(\rho+1)(\rho|s_\ell|^2+1)}} \right) \right)^2 \right.$$
$$\left. + \left(\Omega \left(\frac{\rho \mathrm{Im}[s_\ell]}{\sqrt{(\rho+1)(\rho|s_\ell|^2+1)}} \right) \right)^2 \right)^{-1} - \frac{1}{M}. \quad (50)$$

Some comments are in order. First, the normalized variance of the estimated symbols can be made arbitrarily close to zero by increasing the number of BS antennas M. Second, (49) does not depend on τ since, in the absence of noise, estimating the channel repeatedly over the same pilot symbol does not
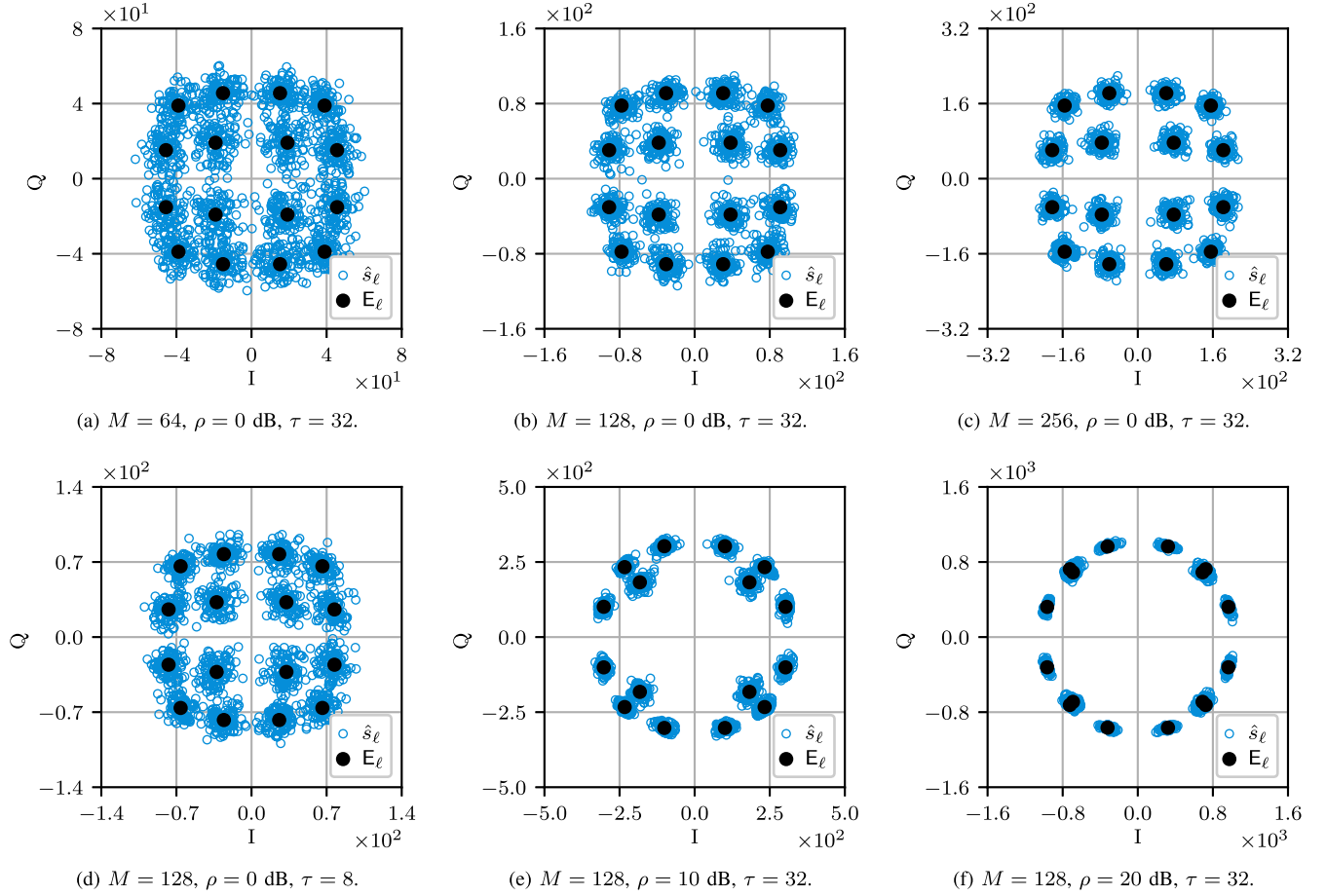
Fig. 4. Estimated symbols with the MRC receiver, with 16-QAM transmit symbols and $K = 1$. The expected value of the estimated symbols is computed as in (43).

bring any benefit. Third, it can be demonstrated that (48) is a quasiconvex function of ρ and, as such, it has a unique minimum that defines a further SNR trade-off. It is shown in Section IV-C that this SNR trade-off, which is inherited from the channel estimation phase through the MRC receiver, is not as significant as the one described in Corollary 4. In fact, the normalized variance of the estimated symbols roughly decreases with ρ; on the other hand, the difference in amplitude between symbols cannot be recovered if ρ is too high. Lastly, while for the channel estimation a reduction of the transmit SNR can be compensated by increasing the pilot length (see Section III-B), a low transmit SNR in the uplink data transmission phase inevitably results in a high normalized variance of the estimated symbols (see, e.g., Fig. 6).[4] In this respect, we point out that the system performance can be further enhanced by optimizing the transmit SNR separately for the two phases of channel estimation and uplink data transmission.

C. Numerical Results and Discussion

We now focus on the performance evaluation of the data detection with 1-bit ADCs with respect to the different parameters using the analytical results presented in Sections IV-A

[4]For instance, assuming $s_\ell = 1$, it is straightforward to observe from (48) that $\lim_{\rho \to 0} \frac{V_\ell}{|E_\ell|^2} = \infty$.

and IV-B. In this regard, we assume that the BS uses the BLM estimator in (4), which is equivalent to the optimal scaled LS estimator in (21) when $K = 1$, in the channel estimation phase and the MRC receiver in the data detection phase. We thus consider the expressions of E_ℓ and V_ℓ derived in (43) and (44), respectively, for the single-UE case (i.e., $K = 1$). As in Section III-C, we use the pilots $\mathbf{p} = \mathbf{p}^\star$, with $\mathbf{p}^\star$ defined in (40), and $\mathbf{p} = \mathbf{1}_\tau$. Moreover, we specifically analyze the scenario where the set of transmit symbols $\mathcal{S}$ corresponds to the 16-QAM constellation, i.e., $\mathcal{S} = \frac{1}{\sqrt{10}}\{ \pm 1 \pm j, \pm 1 \pm j\,3, \pm 3 \pm j, \pm 3 \pm j\,3\}$, which is normalized such that $\frac{1}{L}\sum_{\ell=1}^{L} |s_\ell|^2 = 1$; however, we remark that our analytical framework is valid for any choice of $\mathcal{S}$.

Fig. 4 plots the estimated symbols for different settings, where each 16-QAM symbol is transmitted over 10^2 independent channel realizations and $\mathbf{p} = \mathbf{p}^\star$ is used in the channel estimation phase. The expected value of the estimated symbols is computed as in Theorem 3: this matches the sample average of the estimated symbols for each 16-QAM transmit symbol and can be used to efficiently implement MDD. Comparing Fig. 4(a)–(c), which consider the same transmit SNR and pilot length, the relative dispersion of the estimated symbols about their expected value reduces as the number of BS antennas grows from $M = 64$ to $M = 256$. In fact, a higher granularity in the antenna domain allows
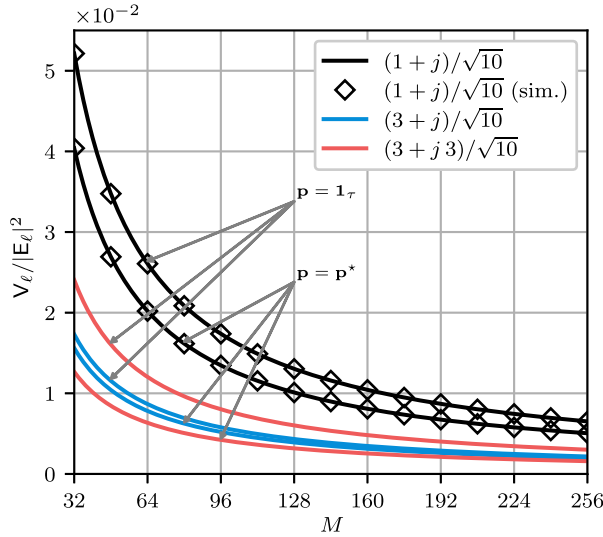
Fig. 5. Normalized variance of the estimated symbols against the number of BS antennas, with 16-QAM transmit symbols, $K = 1$, $\rho = 10$ dB, and $\tau = 32$.

Fig. 6. Normalized variance of the estimated symbols against the transmit SNR, with 16-QAM transmit symbols, $K = 1$, $M = 128$, and $\tau = 32$.

to sum the contribution of a larger number of independent channel entries. On the other hand, comparing Fig. 4(b) and (d), which consider the same number of BS antennas and transmit SNR, the relative dispersion of the estimated symbols about their expected value slightly intensifies as we decrease the pilot length from $\tau = 32$ to $\tau = 8$. This stems from the overall diminished accuracy of the channel estimate used to compute the MRC receiver for each channel realization. Lastly, comparing Fig. 4(b) and (e)–(f), which consider the same number of BS antennas and pilot length, the estimated symbols resulting from the 16-QAM transmit symbols with the same phase, i.e., $\pm \frac{1}{\sqrt{10}}(1 \pm j)$ and $\pm \frac{1}{\sqrt{10}}(3 \pm j\,3)$, get closer as the transmit SNR increases from $\rho = 0$ dB to $\rho = 10$ dB and they almost fully overlap when $\rho = 20$ dB. This behavior was observed in [10] and is formalized in Corollary 4, according to which such estimated symbols become identical in terms of both their expected value and variance at high SNR. In this respect, the SNR trade-off described in Sections IV-A and IV-B is quite evident: while the normalized variance of the estimated symbols roughly decreases with ρ, the difference in amplitude between symbols cannot be recovered if ρ is too high. For the 16-QAM, this produces a SER of about 25% since there are four pairs of indistinguishable estimated symbols (see also Fig. 8). In summary, having independent phases between the channel entries and operating at the right noise level are crucial to accurately estimate the phases and the amplitudes, respectively; we refer to Appendix A and to the related discussion in [10] for more details.

Let us now examine the behavior of the variance of the estimated symbols derived in Theorem 4, which we compare with Monte Carlo simulations with 10^6 independent channel realizations. Fig. 5 considers $\rho = 10$ dB and $\tau = 32$, showing how the normalized variance of the estimated symbols $\frac{V_\ell}{|E_\ell|^2}$ decreases with the number of BS antennas M. The transmit symbols $\pm \frac{1}{\sqrt{10}}(1 \pm j)$, having the smallest power within the 16-QAM constellation, exhibit the most severe dispersion of the estimated symbols about their expected value. Furthermore, the upper bound in (48) becomes more accurate
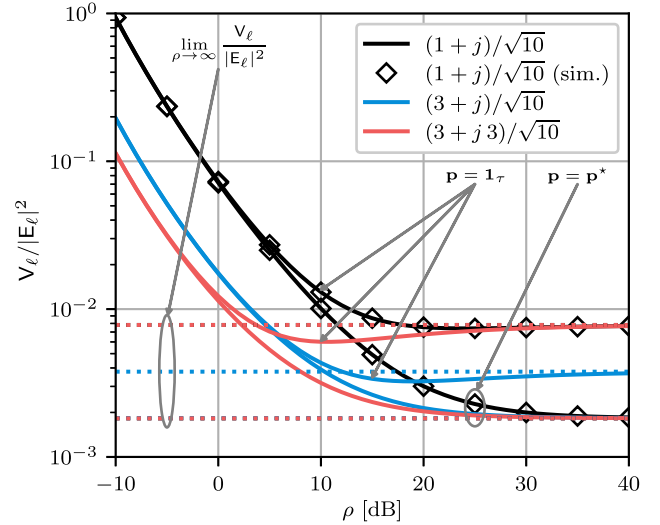
as M grows. Fig. 6 considers $M = 128$ and $\tau = 32$, showing that $\frac{V_\ell}{|E_\ell|^2}$ generally diminishes with the transmit SNR ρ except for the SNR trade-off exhibited with $\mathbf{p} = \mathbf{1}_\tau$; here, the asymptotic expressions in (46) and (49) are also included. Despite this trend, we recall that the difference in amplitude between symbols cannot be recovered if ρ is too high, as discussed in the previous paragraph for Fig. 4: thus, arbitrarily increasing the transmit SNR is detrimental for the system performance. Lastly, Fig. 7 considers $M = 128$ and $\rho = 10$ dB, showing how $\frac{V_\ell}{|E_\ell|^2}$ reduces with the pilot length τ.

We conclude this section by investigating the combined impact of the channel estimation and the data detection with 1-bit ADCs on the system performance in terms of SER, which we compute numerically via Monte Carlo simulations with 10^6 independent channel realizations. In this context, the symbols are decoded by means of MDD aided by the result of Theorem 3. Fig. 8 illustrates the SER against the transmit SNR ρ, with $M = 128$ and $\tau = 32$. Here, the SNR trade-off appears quite evident, whereby the SER decreases until it reaches its minimum at about $\rho = 5$ dB (where the upper bound obtained with $\mathbf{p} = \mathbf{1}_\tau$ proves to be remarkably tight) before escalating again. Then, the SER asymptotically reaches 25% at high SNR, where the inner estimated symbols of the 16-QAM constellation become indistinguishable from the outer estimated symbols with the same phase (see also Fig. 4(f)). We remark that the SER can be further reduced by optimizing the transmit SNR separately for the two phases of channel estimation and uplink data transmission, which is left for future work.

V. CONCLUSION

This paper presents an analytical framework for the channel estimation and the data detection in massive MIMO uplink systems with 1-bit ADCs. First, we provide a precise characterization of the MSE of the channel estimation with respect to different parameters. In addition, we show that, for i.i.d. Rayleigh fading, the BLM estimator can be simplified as a scaled LS estimator with UE-specific scaling factors and that using a common optimized scaling factor for all the
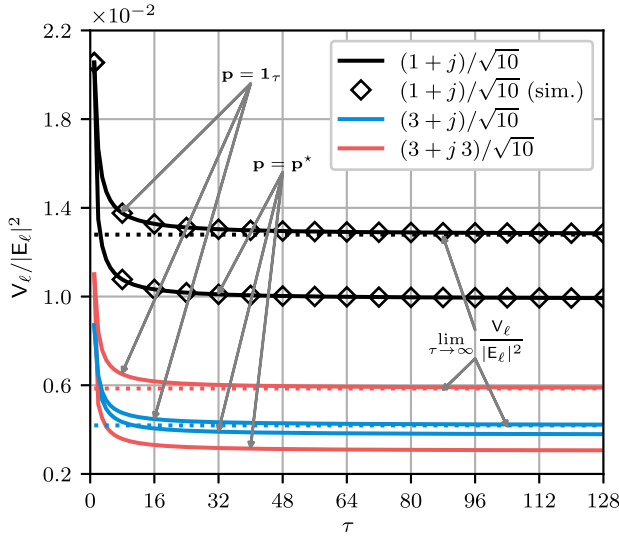
Fig. 7. Normalized variance of the estimated symbols against the pilot length, with 16-QAM transmit symbols, $K = 1$, $M = 128$, and $\rho = 10$ dB.

Fig. 8. SER against the transmit SNR, with 16-QAM transmit symbols, $K = 1$, $M = 128$, and $\tau = 32$.

UEs entails no noticeable performance loss. For the data detection, we characterize the expected value and the variance of the estimated symbols when MRC is adopted. These results can be exploited to efficiently implement MDD and to properly design the set of transmit symbols. The proposed analysis gives important practical insights into the design and the implementation of 1-bit quantized systems. In particular, it highlights a fundamental SNR trade-off, according to which arbitrarily increasing the transmit SNR is detrimental for the system performance. In this respect, the optimal transmit SNR for the channel estimation is shown to decrease as the pilot length increases.

Future work will consider extensions of the proposed analytical framework to more realistic channel models (for the channel estimation) and to the multi-UE case (for the data detection), as well as a SER optimal design of the set of transmit symbols capitalizing on our data detection analysis.

APPENDIX A
FUNDAMENTALS OF CHANNEL ESTIMATION WITH 1-BIT ADCs

Assuming $K = 1$, let $\mathbf{h} \triangleq (h_m) \in \mathbb{C}^{M \times 1}$ and $\mathbf{p} \triangleq (p_u) \in \mathbb{C}^{\tau \times 1}$ denote the uplink channel vector and the pilot, respectively, of the UE. When a scaled LS estimator (such as the one in (5)) is used, the channel estimate $\hat{\mathbf{h}} \triangleq (\hat{h}_m)$ is obtained as

$$\hat{\mathbf{h}} = \sqrt{\Psi} Q \left(\begin{bmatrix} \sqrt{\rho}h_1 p_1^* + Z_{1,1} & \cdots & \sqrt{\rho}h_1 p_\tau^* + Z_{1,\tau} \\ \vdots & \ddots & \vdots \\ \sqrt{\rho}h_M p_1^* + Z_{M,1} & \cdots & \sqrt{\rho}h_M p_\tau^* + Z_{M,\tau} \end{bmatrix} \right) \begin{bmatrix} p_1 \\ \vdots \\ p_\tau \end{bmatrix} \tag{51}$$

with

$$\hat{h}_m = \sqrt{\frac{\rho+1}{2}} \Psi \sum_{u=1}^{\tau} p_u \big(\mathrm{sgn}(\mathrm{Re}[\sqrt{\rho}h_m p_u^* + Z_{m,u}]) + j\,\mathrm{sgn}(\mathrm{Im}[\sqrt{\rho}h_m p_u^* + Z_{m,u}]) \big). \tag{52}$$
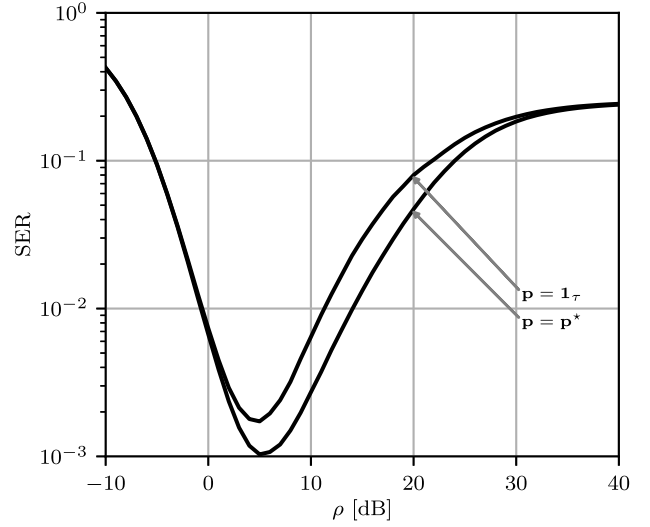
Let $h_m = \alpha_m e^{j\theta_m}$, with $\vartheta_m \triangleq (\theta_m \bmod \frac{\pi}{2})$, and let $p_u = e^{j\phi_u}$ (recall that $|p_u|^2 = 1$, $\forall u$). Assuming $\rho \to \infty$, the phase of h_m can be estimated from $Q(h_m \mathbf{p}^{\mathrm{H}})\mathbf{p}$ as detailed in (53)–(54), shown at the top of the next page, i.e., $Q(e^{j(\theta_m - \phi_u)})$ shifts quadrant according to the phase of p_u. Assuming that the entries of $\mathbf{p}$ span the unit circle, in the limit of $\tau \to \infty$, we obtain (55)–(56), shown at the top of the next page, where (56) follows from

$$\int_{\vartheta_m - \frac{\pi}{2}}^{\vartheta_m} e^{j\phi} \mathrm{d}\phi = (1 - j)e^{j\vartheta_m}, \tag{57}$$

$$\int_{\vartheta_m}^{\vartheta_m + \frac{\pi}{2}} e^{j\phi} \mathrm{d}\phi = (1 + j)e^{j\vartheta_m}, \tag{58}$$

$$\int_{\vartheta_m + \frac{\pi}{2}}^{\vartheta_m + \pi} e^{j\phi} \mathrm{d}\phi = (-1 + j)e^{j\vartheta_m}, \tag{59}$$

$$\int_{\vartheta_m + \pi}^{\vartheta_m + \frac{3\pi}{2}} e^{j\phi} \mathrm{d}\phi = (-1 - j)e^{j\vartheta_m}. \tag{60}$$

Finally, from (56), we have

$$(1 - j)\big(\mathrm{sgn}(\mathrm{Re}[h_m]) + j\,\mathrm{sgn}(\mathrm{Im}[h_m])\big)e^{j\vartheta_m}$$

$$= \begin{cases} 2\,e^{j\vartheta_m} & \text{if } \theta_m \in \left[0, \frac{\pi}{2}\right] \text{ (i.e., if } \theta_m = \vartheta_m), \\ 2j\,e^{j\vartheta_m} & \text{if } \theta_m \in \left[\frac{\pi}{2}, \pi\right] \text{ (i.e., if } \theta_m = \vartheta_m + \frac{\pi}{2}), \\ -2e^{j\vartheta_m} & \text{if } \theta_m \in \left[\pi, \frac{3\pi}{2}\right] \text{ (i.e., if } \theta_m = \vartheta_m + \pi), \\ -2j\,e^{j\vartheta_m} & \text{if } \theta_m \in \left[\frac{3\pi}{2}, 2\pi\right] \text{ (i.e., if } \theta_m = \vartheta_m + \frac{3\pi}{2}) \end{cases} \tag{61}$$

$$= 2e^{j\theta_m} \tag{62}$$

which yields

$$\lim_{\tau \to \infty} \frac{1}{\tau} \sum_{u=1}^{\tau} Q(h_m p_u^*)p_u = \frac{4}{\pi}\sqrt{\frac{\rho+1}{2}}e^{j\theta_m}. \tag{63}$$

$$Q(h_m p_u^*) p_u = Q\big(e^{j\,(\theta_m - \phi_u)}\big) e^{j\,\phi_u} \tag{53}$$

$$
= \begin{cases}
\sqrt{\dfrac{\rho+1}{2}}\big(\operatorname{sgn}(\operatorname{Re}[h_m]) + j\,\operatorname{sgn}(\operatorname{Im}[h_m])\big)e^{j\,\phi_u} & \text{if } \phi_u \in \big[\vartheta_m - \tfrac{\pi}{2}, \vartheta_m\big], \\[12pt]
\sqrt{\dfrac{\rho+1}{2}}\big(\operatorname{sgn}(\operatorname{Im}[h_m]) - j\,\operatorname{sgn}(\operatorname{Re}[h_m])\big)e^{j\,\phi_u} & \text{if } \phi_u \in \big[\vartheta_m, \vartheta_m + \tfrac{\pi}{2}\big], \\[12pt]
\sqrt{\dfrac{\rho+1}{2}}\big(-\operatorname{sgn}(\operatorname{Re}[h_m]) - j\,\operatorname{sgn}(\operatorname{Im}[h_m])\big)e^{j\,\phi_u} & \text{if } \phi_u \in \big[\vartheta_m + \tfrac{\pi}{2}, \vartheta_m + \pi\big], \\[12pt]
\sqrt{\dfrac{\rho+1}{2}}\big(-\operatorname{sgn}(\operatorname{Im}[h_m]) + j\,\operatorname{sgn}(\operatorname{Re}[h_m])\big)e^{j\,\phi_u} & \text{if } \phi_u \in \big[\vartheta_m + \pi, \vartheta_m + \tfrac{3\pi}{2}\big]
\end{cases}
\tag{54}
$$

$$
\lim_{\tau\to\infty} \frac{1}{\tau}\sum_{u=1}^{\tau} Q(h_m p_u^*) p_u = \frac{1}{2\pi}\sqrt{\frac{\rho+1}{2}}\Bigg(\big(\operatorname{sgn}(\operatorname{Re}[h_m]) + j\,\operatorname{sgn}(\operatorname{Im}[h_m])\big)\int_{\vartheta_m - \frac{\pi}{2}}^{\vartheta_m} e^{j\,\phi}\mathrm{d}\phi
$$
$$
+ \big(\operatorname{sgn}(\operatorname{Im}[h_m]) - j\,\operatorname{sgn}(\operatorname{Re}[h_m])\big)\int_{\vartheta_m}^{\vartheta_m + \frac{\pi}{2}} e^{j\,\phi}\mathrm{d}\phi + \big(-\operatorname{sgn}(\operatorname{Re}[h_m]) - j\,\operatorname{sgn}(\operatorname{Im}[h_m])\big)
$$
$$
\times \int_{\vartheta_m + \frac{\pi}{2}}^{\vartheta_m + \pi} e^{j\,\phi}\mathrm{d}\phi + \big(-\operatorname{sgn}(\operatorname{Im}[h_m]) + j\,\operatorname{sgn}(\operatorname{Re}[h_m])\big)\int_{\vartheta_m + \pi}^{\vartheta_m + \frac{3\pi}{2}} e^{j\,\phi}\mathrm{d}\phi\Bigg)
\tag{55}
$$
$$
= \frac{2}{\pi}\sqrt{\frac{\rho+1}{2}}(1-j)\big(\operatorname{sgn}(\operatorname{Re}[h_m]) + j\,\operatorname{sgn}(\operatorname{Im}[h_m])\big)e^{j\,\vartheta_m}
\tag{56}
$$

Hence, the phase of h_m can be estimated accurately if τ is sufficiently large and the pilot symbols span the unit circle. Nonetheless, from (53)–(54), it is straightforward to see that $Q\big(e^{j\,(\theta_m - \phi_u)}\big)e^{j\,\phi_u} = Q\big(e^{j\,(\theta_m - \phi_u \mp \frac{\pi}{2})}\big)e^{j\,(\phi_u \pm \frac{\pi}{2})} = Q\big(e^{j\,(\theta_m - \phi_u \mp \pi)}\big)e^{j\,(\phi_u \pm \pi)}$, i.e., shifting the phase of the pilot symbol by a multiple of $\frac{\pi}{2}$ does not add any information about the phase of h_m when $\rho \to \infty$. As a consequence, the best possible pilot choice features equispaced and non-repeating phases on an interval $\big[\eta, \eta + \frac{\pi}{2}\big]$, with $\eta \in [0, 2\pi)$ (one such choice is $\mathbf{p}^\star$ in (40)). On the other hand, the worst possible pilot choice is given by fixing $\mathbf{p}$ such that $p_u \in \{\pm\beta, \pm j\,\beta\}$, $\forall u$, with $\beta \in \mathbb{C}$ and $|\beta|^2 = 1$ (one such choice is $\mathbf{p} = \mathbf{1}_\tau$). Note that fixing $\mathbf{p} = \mathbf{1}_\tau$ with $\rho \to \infty$ would reduce each channel entry to a scaled symbol of the QPSK constellation regardless of the value of τ, whereas with finite ρ the phase of h_m can be still estimated by exploiting the independent noise realizations over the pilot symbols.

The right-hand side of (63) does not include any information about the amplitude of h_m due to the assumption that $\rho \to \infty$. Assuming now finite ρ and, for simplicity, $\mathbf{p} = \mathbf{1}_\tau$, the amplitude of h_m can be estimated from $Q\big(\sqrt{\rho}h_m\mathbf{1}_\tau^\mathrm{T} + [Z_{m,1}, \ldots, Z_{m,\tau}]\big)\mathbf{1}_\tau$, where

$$
Q(\sqrt{\rho}h_m + Z_{m,u}) = \sqrt{\frac{\rho+1}{2}}\Big(\operatorname{sgn}\big(\sqrt{\rho}\operatorname{Re}[h_m] + \operatorname{Re}[Z_{m,u}]\big)
$$
$$
+ j\,\operatorname{sgn}\big(\sqrt{\rho}\operatorname{Im}[h_m] + \operatorname{Im}[Z_{m,u}]\big)\Big). \tag{64}
$$

In the limit of $\tau \to \infty$, we have

$$
\lim_{\tau\to\infty} \frac{1}{\tau}\sum_{u=1}^{\tau} Q(\sqrt{\rho}h_m + Z_{m,u}) = \sqrt{\frac{\rho+1}{2}}\Big(\operatorname{erf}\big(\sqrt{\rho}\operatorname{Re}[h_m]\big)
$$
$$
+ j\,\operatorname{erf}\big(\sqrt{\rho}\operatorname{Im}[h_m]\big)\Big) \tag{65}
$$

where $\operatorname{erf}(w) \triangleq \frac{2}{\sqrt{\pi}}\int_0^w e^{-t^2}\mathrm{d}t$ denotes the error function. Since $\operatorname{erf}(w)$ is approximately linear for $w \in [-1, 1]$, the difference in amplitude between channel entries can be estimated accurately if their real and imaginary parts lie in $\big[-\frac{1}{\sqrt{\rho}}, \frac{1}{\sqrt{\rho}}\big]$ and τ is sufficiently large, and this holds despite choosing $\mathbf{p} = \mathbf{1}_\tau$. On the other hand, if τ is not sufficiently large at low SNR, the channel estimates are corrupted by the strong noise. Hence, the estimation of the amplitude of h_m benefits from operating at the right noise level.

APPENDIX B
PROOF OF THEOREM 1

We begin by writing (10) as

$$
\mathrm{MSE}_{\mathrm{BLM}} = \frac{1}{MK}\Big(\mathbb{E}[\underline{\mathbf{h}}^\mathrm{H}\underline{\mathbf{h}}] + \mathbb{E}[\hat{\underline{\mathbf{h}}}_{\mathrm{BLM}}^\mathrm{H}\hat{\underline{\mathbf{h}}}_{\mathrm{BLM}}] - 2\mathbb{E}\big[\operatorname{Re}[\hat{\underline{\mathbf{h}}}_{\mathrm{BLM}}^\mathrm{H}\underline{\mathbf{h}}]\big]\Big)
\tag{66}
$$
$$
= 1 + \frac{1}{MK}\bigg(\frac{2}{\pi}\rho\operatorname{tr}\big(\tilde{\mathbf{P}}^\mathrm{T}\boldsymbol{\Sigma}_{\mathrm{p}}^{-1}\mathbb{E}[\mathbf{r}_{\mathrm{p}}\mathbf{r}_{\mathrm{p}}^\mathrm{H}]\boldsymbol{\Sigma}_{\mathrm{p}}^{-1}\tilde{\mathbf{P}}^*\big)
$$
$$
- 2\sqrt{\frac{2}{\pi}}\rho\operatorname{Re}\Big[\operatorname{tr}\big(\mathbb{E}[\underline{\mathbf{h}}\mathbf{r}_{\mathrm{p}}^\mathrm{H}]\boldsymbol{\Sigma}_{\mathrm{p}}^{-1}\tilde{\mathbf{P}}^*\big)\Big]\bigg)
\tag{67}
$$
$$
= 1 - \frac{1}{MK}\frac{2}{\pi}\rho\operatorname{tr}\big(\tilde{\mathbf{P}}^\mathrm{T}\boldsymbol{\Sigma}_{\mathrm{p}}^{-1}\tilde{\mathbf{P}}^*\big)
\tag{68}
$$

where (68) follows from applying $\mathbb{E}[\underline{\mathbf{h}}\mathbf{r}_{\mathrm{p}}^\mathrm{H}] = \sqrt{\frac{2}{\pi}}\rho\tilde{\mathbf{P}}^\mathrm{T}$; note that a similar MSE expression appears in [33, Eq. (48)]. Now, as detailed in Appendix B-A, we can express the covariance matrix of $\underline{\mathbf{r}}_{\mathrm{p}}$ as

$$
\boldsymbol{\Sigma}_{\mathrm{p}} = (\rho K + 1)\boldsymbol{\Phi} \otimes \mathbf{I}_M
\tag{69}
$$

where we have defined $\boldsymbol{\Phi} \triangleq (\Phi_{u,v}) \in \mathbb{C}^{\tau \times \tau}$, with

$$\Phi_{u,v} \triangleq \begin{cases} 1 & \text{if } u = v, \\ \Omega\left(\frac{\rho \sum_{k=1}^{K} \text{Re}[P_{u,k}P_{v,k}^*]}{\rho K+1}\right) - j\,\Omega\left(\frac{\rho \sum_{k=1}^{K} \text{Im}[P_{u,k}P_{v,k}^*]}{\rho K+1}\right) \\ & \text{if } u \neq v. \end{cases} \tag{70}$$

Furthermore, let $\mathbf{p}_k \in \mathbb{C}^{\tau \times 1}$ denote the kth column of $\mathbf{P}$. Hence, we can write

$$\tilde{\mathbf{P}}^{\text{T}} \boldsymbol{\Sigma}_{\text{p}}^{-1} \tilde{\mathbf{P}}^* = \frac{1}{\rho K+1} (\mathbf{P}^{\text{T}} \boldsymbol{\Phi}^{-1} \mathbf{P}^*) \otimes \mathbf{I}_M \tag{71}$$

$$= \frac{\tau^2}{\rho K+1} (\mathbf{P}^{\text{T}} \boldsymbol{\Phi} \mathbf{P}^*)^{-1} \otimes \mathbf{I}_M \tag{72}$$

where (72) results from the fact that, if $\mathbf{P}$ is chosen such that $\mathbf{P}\mathbf{P}^{\text{H}}$ is circulant, $\boldsymbol{\Phi}$ is also circulant and, as a consequence, so is its inverse: in this case, $\mathbf{P}$ diagonalizes both $\boldsymbol{\Phi}$ and its inverse, which implies $\mathbf{p}_k^{\text{T}} \boldsymbol{\Phi} \mathbf{p}_i^* = 0, \forall k \neq i$. Finally, plugging (72) into (68) yields

$$\text{MSE}_{\text{BLM}} = 1 - \frac{1}{MK} \frac{2}{\pi} \frac{\rho\tau^2}{\rho K+1} \text{tr}\left((\mathbf{P}^{\text{T}} \boldsymbol{\Phi} \mathbf{P}^*)^{-1} \otimes \mathbf{I}_M\right) \tag{73}$$

$$= 1 - \frac{1}{K} \frac{2}{\pi} \frac{\rho\tau^2}{\rho K+1} \text{tr}\left((\mathbf{P}^{\text{T}} \boldsymbol{\Phi} \mathbf{P}^*)^{-1}\right) \tag{74}$$

$$= 1 - \frac{2}{\pi} \frac{\rho\tau^2}{\rho K+1} \frac{1}{K} \sum_{k=1}^{K} \frac{1}{\mathbf{p}_k^{\text{T}} \boldsymbol{\Phi} \mathbf{p}_k^*} \tag{75}$$

and the expression in (12) is obtained by observing that $\mathbf{p}_k^{\text{T}} \boldsymbol{\Phi} \mathbf{p}_k^* = \tau + \delta_k$, with δ_k defined in (13). ∎

A. Derivations of (69)

In this section, we derive the closed-form expression of $\boldsymbol{\Sigma}_{\text{p}}$. To this end, we introduce the following definitions:

$$A_{m,u} \triangleq \text{sgn}\left(\text{Re}\left[\sqrt{\rho} \sum_{k=1}^{K} H_{m,k}P_{u,k}^* + Z_{m,u}\right]\right) \tag{76}$$

$$= \text{sgn}\left(\sqrt{\rho} \sum_{k=1}^{K} \left(\text{Re}[H_{m,k}]\text{Re}[P_{u,k}]\right.\right.$$
$$\left.\left. + \text{Im}[H_{m,k}]\text{Im}[P_{u,k}]\right) + \text{Re}[Z_{m,u}]\right), \tag{77}$$

$$B_{m,u} \triangleq \text{sgn}\left(\text{Im}\left[\sqrt{\rho} \sum_{k=1}^{K} H_{m,k}P_{u,k}^* + Z_{m,u}\right]\right) \tag{78}$$

$$= \text{sgn}\left(\sqrt{\rho} \sum_{k=1}^{K} \left(-\text{Re}[H_{m,k}]\text{Im}[P_{u,k}]\right.\right.$$
$$\left.\left. + \text{Im}[H_{m,k}]\text{Re}[P_{u,k}]\right) + \text{Im}[Z_{m,u}]\right). \tag{79}$$

Moreover, we present the following proposition, which will be also used in Appendix D.

Proposition 1: Let $\boldsymbol{\zeta} \sim \mathcal{N}(0, \gamma \mathbf{I}_N)$, with $\gamma > 0$. For $\mathbf{a}_1, \mathbf{a}_2 \in \mathbb{R}^{N \times 1}$, we have

$$\mathbb{E}\left[\text{sgn}(\mathbf{a}_1^{\text{T}} \boldsymbol{\zeta})\text{sgn}(\mathbf{a}_2^{\text{T}} \boldsymbol{\zeta})\right] = \Omega\left(\frac{\mathbf{a}_1^{\text{T}} \mathbf{a}_2}{\|\mathbf{a}_1\| \|\mathbf{a}_2\|}\right). \tag{80}$$

Proof: To obtain the expression in (80), we first observe that

$$\mathbb{E}\left[\text{sgn}(X_1)\text{sgn}(X_2)\right]$$
$$= \mathbb{P}[X_1 > 0, X_2 > 0] + \mathbb{P}[X_1 < 0, X_2 < 0]$$
$$- \mathbb{P}[X_1 > 0, X_2 < 0] - \mathbb{P}[X_1 < 0, X_2 > 0]. \tag{81}$$

For $X_1 = \mathbf{a}_1^{\text{T}} \boldsymbol{\zeta}$ and $X_2 = \mathbf{a}_2^{\text{T}} \boldsymbol{\zeta}$, the first term on the right-hand side of (81) can be obtained building on [34] as

$$\mathbb{P}\left[\mathbf{a}_1^{\text{T}} \boldsymbol{\zeta} > 0, \mathbf{a}_2^{\text{T}} \boldsymbol{\zeta} > 0\right] = \frac{1}{4}\left(1 + \Omega\left(\frac{\mathbf{a}_1^{\text{T}} \mathbf{a}_2}{\|\mathbf{a}_1\| \|\mathbf{a}_2\|}\right)\right) \tag{82}$$

where $\frac{\mathbf{a}_1^{\text{T}} \mathbf{a}_2}{\|\mathbf{a}_1\| \|\mathbf{a}_2\|}$ represents the correlation coefficient between X_1 and X_2, and the other terms can be derived following similar steps. ∎

Now, we can write the $((u-1)M+m, (v-1)M+n)$th entry of $\underline{\mathbf{r}}_{\text{p}} \underline{\mathbf{r}}_{\text{p}}^{\text{H}}$ as

$$R_{m,u}R_{n,v}^* = \frac{\rho K+1}{2}(A_{m,u} + j\,B_{m,u})(A_{n,v} + j\,B_{n,v})^* \tag{83}$$

$$= \frac{\rho K+1}{2}\left(A_{m,u}A_{n,v} + B_{m,u}B_{n,v}\right.$$
$$\left. + j\,(B_{m,u}A_{n,v} - A_{m,u}B_{n,v})\right) \tag{84}$$

with $R_{m,u}R_{m,u}^* = \rho K+1$. Hence, the expected value of (84) is given by (85), shown at the top of the next page. This results from

$$\mathbb{E}[A_{m,u}A_{m,v}] = \mathbb{E}[B_{m,u}B_{m,v}] \tag{86}$$

$$= \Omega\left(\frac{\rho \sum_{k=1}^{K} \text{Re}[P_{u,k}P_{v,k}^*]}{\rho K+1}\right), \tag{87}$$

$$\mathbb{E}[A_{m,u}B_{m,v}] = \Omega\left(\frac{\rho \sum_{k=1}^{K} \text{Im}[P_{u,k}P_{v,k}^*]}{\rho K+1}\right) \tag{88}$$

which are derived by applying Proposition 1. For instance, $\mathbb{E}[A_{m,u}A_{m,v}]$ in (86)–(87) can be obtained by plugging (89)–(91), shown at the top of the next page, into (80), which gives (92), shown at the top of the next page. Finally, the expression in (69) readily follows from (85).

APPENDIX C
PROOF OF COROLLARY 1

We show that the estimator in (15) provides the same MSE of the channel estimation as the BLM estimator in (4). To this end, we write $\text{vec}[\hat{\mathbf{H}}_{\text{BLM}}] = \tilde{\boldsymbol{\Psi}}^{\frac{1}{2}} \tilde{\mathbf{P}}^{\text{T}} \underline{\mathbf{r}}_{\text{p}}$, where we have defined $\tilde{\boldsymbol{\Psi}} \triangleq \boldsymbol{\Psi} \otimes \mathbf{I}_M \in \mathbb{C}^{MK \times MK}$. Following similar steps as in the proof of Theorem 1, we obtain

$$\text{MSE}_{\text{BLM}} = 1 + \frac{1}{MK}\left(\text{tr}\left(\tilde{\boldsymbol{\Psi}}^{\frac{1}{2}} \tilde{\mathbf{P}}^{\text{T}} \mathbb{E}[\mathbf{r}_{\text{p}} \mathbf{r}_{\text{p}}^{\text{H}}] \tilde{\mathbf{P}}^* \tilde{\boldsymbol{\Psi}}^{\frac{1}{2}}\right)\right.$$
$$\left. - 2\text{Re}\left[\text{tr}\left(\mathbb{E}[\underline{\mathbf{h}}\mathbf{r}_{\text{p}}^{\text{H}}] \tilde{\mathbf{P}}^* \tilde{\boldsymbol{\Psi}}^{\frac{1}{2}}\right)\right]\right) \tag{93}$$

$$= 1 + \frac{1}{MK}\left(\text{tr}\left(\tilde{\boldsymbol{\Psi}}^{\frac{1}{2}} \tilde{\mathbf{P}}^{\text{T}} \boldsymbol{\Sigma}_{\text{p}} \tilde{\mathbf{P}}^* \tilde{\boldsymbol{\Psi}}^{\frac{1}{2}}\right)\right.$$
$$\left. - 2\sqrt{\frac{2}{\pi}}\rho\,\text{tr}\left(\tilde{\mathbf{P}}^{\text{T}} \tilde{\mathbf{P}}^* \tilde{\boldsymbol{\Psi}}^{\frac{1}{2}}\right)\right). \tag{94}$$

$$\mathbb{E}[R_{m,u}R_{n,v}^*] = \begin{cases} (\rho K + 1)\Big(\Omega\Big(\frac{\rho \sum_{k=1}^K \mathrm{Re}[P_{u,k}P_{v,k}^*]}{\rho K+1}\Big) - j\,\Omega\Big(\frac{\rho \sum_{k=1}^K \mathrm{Im}[P_{u,k}P_{v,k}^*]}{\rho K+1}\Big)\Big) & \text{if } m=n, u\neq v, \\ 0 & \text{if } m \neq n \end{cases} \tag{85}$$

$$\boldsymbol{\zeta} = \big[\mathrm{Re}[H_{m,1}],\dots,\mathrm{Re}[H_{m,K}],\mathrm{Im}[H_{m,1}],\dots,\mathrm{Im}[H_{m,K}],\mathrm{Re}[Z_{m,u}],\mathrm{Re}[Z_{m,v}]\big]^{\mathrm{T}} \sim \mathcal{N}\Big(\mathbf{0}_{(2K+2)},\tfrac{1}{2}\mathbf{I}_{(2K+2)}\Big), \tag{89}$$

$$\mathbf{a}_1 = \big[\sqrt{\rho}\mathrm{Re}[P_{u,1}],\dots,\sqrt{\rho}\mathrm{Re}[P_{u,K}],\sqrt{\rho}\mathrm{Im}[P_{u,1}],\dots,\sqrt{\rho}\mathrm{Im}[P_{u,K}],1,0\big]^{\mathrm{T}} \in \mathbb{R}^{(2K+2)\times 1}, \tag{90}$$

$$\mathbf{a}_2 = \big[\sqrt{\rho}\mathrm{Re}[P_{v,1}],\dots,\sqrt{\rho}\mathrm{Re}[P_{v,K}],\sqrt{\rho}\mathrm{Im}[P_{v,1}],\dots,\sqrt{\rho}\mathrm{Im}[P_{v,K}],0,1\big]^{\mathrm{T}} \in \mathbb{R}^{(2K+2)\times 1} \tag{91}$$

$$\mathbb{E}\bigg[\mathrm{sgn}\bigg(\sqrt{\rho}\sum_{i=1}^K \big(\mathrm{Re}[H_{m,i}]\mathrm{Re}[P_{u,i}] + \mathrm{Im}[H_{m,i}]\mathrm{Im}[P_{u,i}]\big) + \mathrm{Re}[Z_{m,u}]\bigg)$$

$$\times \mathrm{sgn}\bigg(\sqrt{\rho}\sum_{i=1}^K \big(\mathrm{Re}[H_{m,i}]\mathrm{Re}[P_{v,i}] + \mathrm{Im}[H_{m,i}]\mathrm{Im}[P_{v,i}]\big) + \mathrm{Re}[Z_{m,v}]\bigg)\bigg]$$

$$= \Omega\bigg(\frac{\rho \sum_{k=1}^K (\mathrm{Re}[P_{u,k}]\mathrm{Re}[P_{v,k}] + \mathrm{Im}[P_{u,k}]\mathrm{Im}[P_{v,k}])}{(\rho \sum_{k=1}^K (\mathrm{Re}[P_{u,k}]^2 + \mathrm{Im}[P_{u,k}]^2) + 1)^{\frac{1}{2}}(\rho \sum_{k=1}^K (\mathrm{Re}[P_{v,k}]^2 + \mathrm{Im}[P_{v,k}]^2) + 1)^{\frac{1}{2}}}\bigg) \tag{92}$$

Hence, we can write

$$\tilde{\boldsymbol{\Psi}}^{\frac{1}{2}}\tilde{\mathbf{P}}^{\mathrm{T}}\boldsymbol{\Sigma}_{\mathrm{p}}\tilde{\mathbf{P}}^*\tilde{\boldsymbol{\Psi}}^{\frac{1}{2}} = (\rho K + 1)(\boldsymbol{\Psi}^{\frac{1}{2}}\mathbf{P}^{\mathrm{T}}\boldsymbol{\Phi}\mathbf{P}^*\boldsymbol{\Psi}^{\frac{1}{2}}) \otimes \mathbf{I}_M, \tag{95}$$

$$\tilde{\mathbf{P}}^{\mathrm{T}}\tilde{\mathbf{P}}^*\tilde{\boldsymbol{\Psi}}^{\frac{1}{2}} = (\mathbf{P}^{\mathrm{T}}\mathbf{P}^*\boldsymbol{\Psi}^{\frac{1}{2}}) \otimes \mathbf{I}_M \tag{96}$$

and (94) becomes

$$\mathrm{MSE}_{\mathrm{BLM}} = 1 + \frac{1}{MK}\bigg((\rho K + 1)\mathrm{tr}\big((\boldsymbol{\Psi}^{\frac{1}{2}}\mathbf{P}^{\mathrm{T}}\boldsymbol{\Phi}\mathbf{P}^*\boldsymbol{\Psi}^{\frac{1}{2}}) \otimes \mathbf{I}_M\big)$$

$$- 2\sqrt{\frac{2}{\pi}}\rho\,\mathrm{tr}\big((\mathbf{P}^{\mathrm{T}}\mathbf{P}^*\boldsymbol{\Psi}^{\frac{1}{2}}) \otimes \mathbf{I}_M\big)\bigg) \tag{97}$$

$$= 1 + \frac{1}{K}\bigg((\rho K + 1)\mathrm{tr}(\boldsymbol{\Psi}^{\frac{1}{2}}\mathbf{P}^{\mathrm{T}}\boldsymbol{\Phi}\mathbf{P}^*\boldsymbol{\Psi}^{\frac{1}{2}})$$

$$- 2\sqrt{\frac{2}{\pi}}\rho\,\mathrm{tr}(\mathbf{P}^{\mathrm{T}}\mathbf{P}^*\boldsymbol{\Psi}^{\frac{1}{2}})\bigg) \tag{98}$$

$$= 1 + \frac{1}{K}\sum_{k=1}^K \bigg((\rho K + 1)\Psi_k(\tau + \delta_k) - 2\sqrt{\frac{2}{\pi}\Psi_k}\rho\tau\bigg). \tag{99}$$

Since each term in the summation of (99) is a convex function of Ψ_k, the expression of the UE-specific scaling factor in (16) can be obtained by setting $\frac{\mathrm{d}}{\mathrm{d}\Psi_k}(99) = 0$. Finally, plugging (16) into (99) yields the MSE of the BLM estimator in (12). ∎

APPENDIX D
PROOF OF THEOREM 3

We begin by introducing the following definitions:[5]

$$a_{m,u} \triangleq \mathrm{sgn}\big(\mathrm{Re}[\sqrt{\rho}h_m p_u^* + Z_{m,u}]\big) \tag{100}$$

$$= \mathrm{sgn}\big(\sqrt{\rho}(\mathrm{Re}[h_m]\mathrm{Re}[p_u] + \mathrm{Im}[h_m]\mathrm{Im}[p_u])$$

$$+ \mathrm{Re}[Z_{m,u}]\big), \tag{101}$$

$$b_{m,u} \triangleq \mathrm{sgn}\big(\mathrm{Im}[\sqrt{\rho}h_m p_u^* + Z_{m,u}]\big) \tag{102}$$

$$= \mathrm{sgn}\big(\sqrt{\rho}(-\mathrm{Re}[h_m]\mathrm{Im}[p_u] + \mathrm{Im}[h_m]\mathrm{Re}[p_u])$$

$$+ \mathrm{Im}[Z_{m,u}]\big) \tag{103}$$

[5]Note that (100)–(103) are equivalent to (76)–(79) for $K = 1$.

and

$$c_m \triangleq \mathrm{sgn}\big(\mathrm{Re}[\sqrt{\rho}h_m s_\ell + z_m]\big) \tag{104}$$

$$= \mathrm{sgn}\big(\sqrt{\rho}(\mathrm{Re}[h_m]\mathrm{Re}[s_\ell] - \mathrm{Im}[h_m]\mathrm{Im}[s_\ell]) + \mathrm{Re}[z_m]\big), \tag{105}$$

$$d_m \triangleq \mathrm{sgn}\big(\mathrm{Im}[\sqrt{\rho}h_m s_\ell + z_m]\big) \tag{106}$$

$$= \mathrm{sgn}\big(\sqrt{\rho}(\mathrm{Re}[h_m]\mathrm{Im}[s_\ell] + \mathrm{Im}[h_m]\mathrm{Re}[s_\ell]) + \mathrm{Im}[z_m]\big). \tag{107}$$

From (42), we can write the estimated symbol as

$$\hat{s}_\ell = \frac{\rho + 1}{2}\sqrt{\Psi'}\sum_{m=1}^M\sum_{u=1}^\tau p_u^*(a_{m,u} + j\,b_{m,u})^*(c_m + j\,d_m) \tag{108}$$

$$= \frac{\rho + 1}{2}\sqrt{\Psi'}\sum_{m=1}^M\sum_{u=1}^\tau p_u^*\big(a_{m,u}c_m + b_{m,u}d_m$$

$$+ j\,(a_{m,u}d_m - b_{m,u}c_m)\big). \tag{109}$$

Hence, the expression in (43) results from

$$\mathbb{E}[a_{m,u}c_m] = \mathbb{E}[b_{m,u}d_m] \tag{110}$$

$$= \Omega\bigg(\frac{\rho\mathrm{Re}[p_u s_\ell]}{\sqrt{(\rho+1)(\rho|s_\ell|^2 + 1)}}\bigg), \tag{111}$$

$$\mathbb{E}[a_{m,u}d_m] = -\mathbb{E}[b_{m,u}c_m] \tag{112}$$

$$= \Omega\bigg(\frac{\rho\mathrm{Im}[p_u s_\ell]}{\sqrt{(\rho+1)(\rho|s_\ell|^2 + 1)}}\bigg) \tag{113}$$

which are derived again by applying Proposition 1. For instance, $\mathbb{E}[a_{m,u}c_m]$ in (110)–(111) can be obtained by plugging

$$\boldsymbol{\zeta} = \big[\mathrm{Re}[h_m],\mathrm{Im}[h_m],\mathrm{Re}[Z_{m,u}],\mathrm{Re}[z_m]\big]^{\mathrm{T}} \sim \mathcal{N}\Big(\mathbf{0}_4,\tfrac{1}{2}\mathbf{I}_4\Big), \tag{114}$$

$$\mathbf{a}_1 = \big[\sqrt{\rho}\mathrm{Re}[p_u],\sqrt{\rho}\mathrm{Im}[p_u],1,0\big]^{\mathrm{T}} \in \mathbb{R}^{4\times 1}, \tag{115}$$

$$\mathbf{a}_2 = \big[\sqrt{\rho}\mathrm{Re}[s_\ell],-\sqrt{\rho}\mathrm{Im}[s_\ell],0,1\big]^{\mathrm{T}} \in \mathbb{R}^{4\times 1} \tag{116}$$

$$\mathbb{E}\Big[\text{sgn}\Big(\sqrt{\rho}\big(\text{Re}[h_m]\text{Re}[p_u] + \text{Im}[h_m]\text{Im}[p_u]\big) + \text{Re}[Z_{m,u}]\Big)\text{sgn}\Big(\sqrt{\rho}\big(\text{Re}[h_m]\text{Re}[s_\ell] - \text{Im}[h_m]\text{Im}[s_\ell]\big) + \text{Re}[z_m]\Big)\Big]$$
$$= \Omega\bigg(\frac{\rho(\text{Re}[p_u]\text{Re}[s_\ell] - \text{Im}[p_u]\text{Im}[s_\ell])}{\sqrt{\rho(\text{Re}[p_u]^2 + \text{Im}[p_u]^2) + 1}\sqrt{\rho(\text{Re}[s_\ell]^2 + \text{Im}[s_\ell]^2) + 1}}\bigg) \tag{117}$$

$$|\mathsf{E}_\ell|^2 = (\rho+1)^2 \Psi' M^2 \sum_{u=1}^{\tau} \sum_{v=1}^{\tau} \bigg(\text{Re}[p_u^* p_v]\Big(\Omega\Big(\frac{\rho\text{Re}[p_u s_\ell]}{\sqrt{(\rho+1)(\rho|s_\ell|^2 + 1)}}\Big)\Omega\Big(\frac{\rho\text{Re}[p_v s_\ell]}{\sqrt{(\rho+1)(\rho|s_\ell|^2 + 1)}}\Big)$$
$$+ \Omega\Big(\frac{\rho\text{Im}[p_u s_\ell]}{\sqrt{(\rho+1)(\rho|s_\ell|^2 + 1)}}\Big)\Omega\Big(\frac{\rho\text{Im}[p_v s_\ell]}{\sqrt{(\rho+1)(\rho|s_\ell|^2 + 1)}}\Big)\Big) + \text{Im}[p_u^* p_v]\Big(\Omega\Big(\frac{\rho\text{Re}[p_u s_\ell]}{\sqrt{(\rho+1)(\rho|s_\ell|^2 + 1)}}\Big)$$
$$\times \Omega\Big(\frac{\rho\text{Im}[p_v s_\ell]}{\sqrt{(\rho+1)(\rho|s_\ell|^2 + 1)}}\Big) - \Omega\Big(\frac{\rho\text{Im}[p_u s_\ell]}{\sqrt{(\rho+1)(\rho|s_\ell|^2 + 1)}}\Big)\Omega\Big(\frac{\rho\text{Re}[p_v s_\ell]}{\sqrt{(\rho+1)(\rho|s_\ell|^2 + 1)}}\Big)\Big)\bigg) \tag{119}$$

$$\text{Re}[\hat{s}_\ell]^2 = \frac{(\rho+1)^2}{4}\Psi'\bigg(\sum_{m=1}^{M}\sum_{u=1}^{\tau}\big(\text{Re}[p_u](a_{m,u}c_m + b_{m,u}d_m) + \text{Im}[p_u](a_{m,u}d_m - b_{m,u}c_m)\big)\bigg)^2 \tag{120}$$
$$= \frac{(\rho+1)^2}{4}\Psi'\bigg(\sum_{m=1}^{M}\sum_{u=1}^{\tau}\big(\text{Re}[p_u](a_{m,u}c_m + b_{m,u}d_m) + \text{Im}[p_u](a_{m,u}d_m - b_{m,u}c_m)\big)^2$$
$$+ \sum_{m=1}^{M}\sum_{u \neq v}\big(\text{Re}[p_u](a_{m,u}c_m + b_{m,u}d_m) + \text{Im}[p_u](a_{m,u}d_m - b_{m,u}c_m)\big)\big(\text{Re}[p_v](a_{m,v}c_m + b_{m,v}d_m)$$
$$+ \text{Im}[p_v](a_{m,v}d_m - b_{m,v}c_m)\big) + \sum_{m \neq n}\sum_{u=1}^{\tau}\sum_{v=1}^{\tau}\big(\text{Re}[p_u](a_{m,u}c_m + b_{m,u}d_m) + \text{Im}[p_u](a_{m,u}d_m - b_{m,u}c_m)\big)$$
$$\times \big(\text{Re}[p_v](a_{n,v}c_n + b_{n,v}d_n) + \text{Im}[p_v](a_{n,v}d_n - b_{n,v}c_n)\big)\bigg) \tag{121}$$

$$|\hat{s}_\ell|^2 = \frac{(\rho+1)^2}{4}\Psi'\bigg(4M\tau + 2\sum_{m=1}^{M}\sum_{u \neq v}\big(\text{Re}[p_u^* p_v](a_{m,u}a_{m,v} + b_{m,u}b_{m,v}) - \text{Im}[p_u^* p_v](a_{m,u}b_{m,v} - b_{m,u}a_{m,v})\big)$$
$$+ \sum_{m \neq n}\sum_{u=1}^{\tau}\sum_{v=1}^{\tau}\Big(\text{Re}[p_u^* p_v]\big((a_{m,u}c_m + b_{m,u}d_m)(a_{n,v}c_n + b_{n,v}d_n) + (a_{m,u}d_m - b_{m,u}c_m)(a_{n,v}d_n - b_{n,v}c_n)\big)$$
$$+ \text{Im}[p_u^* p_v]\big((a_{m,u}c_m + b_{m,u}d_m)(a_{n,v}d_n - b_{n,v}c_n) - (a_{m,u}d_m - b_{m,u}c_m)(a_{n,v}c_n + b_{n,v}d_n)\big)\Big)\bigg) \tag{122}$$

$$\mathbb{E}\big[|\hat{s}_\ell|^2\big] = (\rho+1)^2 \Psi' M\bigg(\tau + \sum_{u \neq v}\Big(\text{Re}[p_u^* p_v]\Omega\Big(\frac{\rho\text{Re}[p_u p_v^*]}{\rho+1}\Big) - \text{Im}[p_u^* p_v]\Omega\Big(\frac{\rho\text{Im}[p_u p_v^*]}{\rho+1}\Big)\Big)$$
$$+ (M-1)\sum_{u=1}^{\tau}\sum_{v=1}^{\tau}\text{Re}[p_u^* p_v]\Big(\Omega\Big(\frac{\rho\text{Re}[p_u s_\ell]}{\sqrt{(\rho+1)(\rho|s_\ell|^2+1)}}\Big)\Omega\Big(\frac{\rho\text{Re}[p_v s_\ell]}{\sqrt{(\rho+1)(\rho|s_\ell|^2+1)}}\Big) + \Omega\Big(\frac{\rho\text{Im}[p_u s_\ell]}{\sqrt{(\rho+1)(\rho|s_\ell|^2+1)}}\Big)$$
$$\times \Omega\Big(\frac{\rho\text{Im}[p_v s_\ell]}{\sqrt{(\rho+1)(\rho|s_\ell|^2+1)}}\Big)\Big) + (M-1)\sum_{u=1}^{\tau}\sum_{v=1}^{\tau}\text{Im}[p_u^* p_v]\Big(\Omega\Big(\frac{\rho\text{Re}[p_u s_\ell]}{\sqrt{(\rho+1)(\rho|s_\ell|^2+1)}}\Big)$$
$$\times \Omega\Big(\frac{\rho\text{Im}[p_v s_\ell]}{\sqrt{(\rho+1)(\rho|s_\ell|^2+1)}}\Big) - \Omega\Big(\frac{\rho\text{Im}[p_u s_\ell]}{\sqrt{(\rho+1)(\rho|s_\ell|^2+1)}}\Big)\Omega\Big(\frac{\rho\text{Re}[p_v s_\ell]}{\sqrt{(\rho+1)(\rho|s_\ell|^2+1)}}\Big)\Big)\bigg) \tag{123}$$

into (80), which gives (117), shown at the top of the page. ∎

APPENDIX E
PROOF OF THEOREM 4

The variance of the estimated symbol $\hat{s}_\ell$ can be written as

$$\mathsf{V}_\ell = \mathbb{E}\big[|\hat{s}_\ell|^2\big] - |\mathsf{E}_\ell|^2 \tag{118}$$

with $\mathbb{E}\big[|\hat{s}_\ell|^2\big] = \mathbb{E}\big[\text{Re}[\hat{s}_\ell]^2\big] + \mathbb{E}\big[\text{Im}[\hat{s}_\ell]^2\big]$ and where $|\mathsf{E}_\ell|^2$ is given by (119), shown at the top of the page (cf. (43)). Furthermore, recalling the definitions in (100)–(107) and building on (108)–(109), we can write $\text{Re}[\hat{s}_\ell]^2$ as in (120)–(121), shown at the top of the page, and $\text{Im}[\hat{s}_\ell]^2$ can be obtained following similar steps. Then, summing up (121) and $\text{Im}[\hat{s}_\ell]^2$ yields $|\hat{s}_\ell|^2$ in (122), shown at the top of the page. Now, we have that $\xi_{m,u} \in \{a_{m,u}, b_{m,u}, c_m, d_m\}$ and

$\xi_{n,v} \in \{a_{n,v}, b_{n,v}, c_n, d_n\}$ are independent random variables if $m \neq n$ regardless of the indices u, v. This implies that $\mathbb{E}[a_{m,u} c_m b_{n,v} d_n] = \mathbb{E}[a_{m,u} c_m] \mathbb{E}[b_{n,v} d_n]$ and the same holds for the other products in the second summation of (122). Hence, the expected value of (122) is given by (123), shown at the top of the previous page, which follows from (110)–(113) and from[6]

$$\mathbb{E}[a_{m,u} a_{m,v}] = \mathbb{E}[b_{m,u} b_{m,v}] \tag{124}$$

$$= \Omega\left(\frac{\rho \mathrm{Re}[p_u p_v^*]}{\rho + 1}\right), \tag{125}$$

$$\mathbb{E}[a_{m,u} b_{m,v}] = \Omega\left(\frac{\rho \mathrm{Im}[p_u p_v^*]}{\rho + 1}\right). \tag{126}$$

Finally, the expression in (44) is obtained by plugging (119) and (123) into (118). ∎

ACKNOWLEDGMENT

REFERENCES

[1] I. Atzeni and A. Tölli, "Uplink data detection analysis of 1-bit quantized massive MIMO," in *Proc. IEEE Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Lucca, Italy, Sep. 2021, pp. 1–5.

[2] N. Rajatheva *et al.* (Jun. 2020). *White Paper on Broadband Connectivity in 6G*. [Online]. Available: http://jultika.oulu.fi/files/isbn9789526226798.pdf

[3] F. Rusek *et al.*, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.

[4] M. Xiao *et al.*, "Millimeter wave communications for future mobile networks," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 9, pp. 1909–1935, Sep. 2017.

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

[6] C. Studer and G. Durisi, "Quantized massive MU-MIMO-OFDM uplink," *IEEE Trans. Commun.*, vol. 64, no. 6, pp. 2387–2399, Jun. 2016.

[7] J. Choi, J. Mo, and R. W. Heath, Jr., "Near maximum-likelihood detector and channel estimator for uplink multiuser massive MIMO systems with one-bit ADCs," *IEEE Trans. Commun.*, vol. 64, no. 5, pp. 2005–2018, Mar. 2016.

[8] Y. Li, C. Tao, G. Seco-Granados, A. Mezghani, A. L. Swindlehurst, and L. Liu, "Channel estimation and performance analysis of one-bit massive MIMO systems," *IEEE Trans. Signal Process.*, vol. 65, no. 15, pp. 4075–4089, Aug. 2017.

[9] A. K. Saxena, I. Fijalkow, and A. L. Swindlehurst, "Analysis of one-bit quantized precoding for the multiuser massive MIMO downlink," *IEEE Trans. Signal Process.*, vol. 65, no. 17, pp. 4624–4634, Sep. 2017.

[10] S. Jacobsson, G. Durisi, M. Coldrey, U. Gustavsson, and C. Studer, "Throughput analysis of massive MIMO uplink with low-resolution ADCs," *IEEE Trans. Wireless Commun.*, vol. 16, no. 6, pp. 4038–4051, Jun. 2017.

[11] S. Jacobsson, G. Durisi, M. Coldrey, T. Goldstein, and C. Studer, "Quantized precoding for massive MU-MIMO," *IEEE Trans. Commun.*, vol. 65, no. 11, pp. 4670–4684, Nov. 2017.

[12] S. Jacobsson, G. Durisi, M. Coldrey, and C. Studer, "Linear precoding with low-resolution DACs for massive MU-MIMO-OFDM downlink," *IEEE Trans. Wireless Commun.*, vol. 18, no. 3, pp. 1595–1609, Mar. 2019.

[13] J. Singh, O. Dabeer, and U. Madhow, "On the limits of communication with low-precision analog-to-digital conversion at the receiver," *IEEE Trans. Commun.*, vol. 57, no. 12, pp. 3629–3639, Dec. 2009.

[14] A. Mezghani and J. A. Nossek, "Capacity lower bound of MIMO channels with output quantization and correlated noise," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Cambridge, MA, USA, Jul. 2012, pp. 1–5.

[15] N. Liang and W. Zhang, "Mixed-ADC massive MIMO," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 983–997, Sep. 2016.

[16] C. Mollén, J. Choi, E. G. Larsson, and R. W. Heath, Jr., "Uplink performance of wideband massive MIMO with one-bit ADCs," *IEEE Trans. Wireless Commun.*, vol. 16, no. 1, pp. 87–100, Jan. 2017.

[17] S. Wang, Y. Li, and J. Wang, "Multiuser detection in massive spatial modulation MIMO with low-resolution ADCs," *IEEE Trans. Wireless Commun.*, vol. 14, no. 4, pp. 2156–2168, Apr. 2015.

[18] Y.-S. Jeon, N. Lee, S.-N. Hong, and R. W. Heath, Jr., "One-bit sphere decoding for uplink massive MIMO systems with one-bit ADCs," *IEEE Trans. Wireless Commun.*, vol. 17, no. 7, pp. 4509–4521, Jul. 2018.

[19] C. K. Wen, C. J. Wang, S. Jin, K. K. Wong, and P. Ting, "Bayes-optimal joint channel-and-data estimation for massive MIMO with low-precision ADCs," *IEEE Trans. Signal Process.*, vol. 64, no. 10, pp. 2541–2556, May 2016.

[20] J. J. Bussgang, "Crosscorrelation functions of amplitude-distorted Gaussian signals," Res. Lab. Electron., Massachusetts Inst. Technol., Cambridge, MA, USA, Tech. Rep. 216, Mar. 1952.

[21] Y. Li, C. Tao, A. L. Swindlehurst, A. Mezghani, and L. Liu, "Downlink achievable rate analysis in massive MIMO systems with one-bit DACs," *IEEE Commun. Lett.*, vol. 21, no. 7, pp. 1669–1672, Jul. 2017.

[22] M. S. Stein, S. Bar, J. A. Nossek, and J. Tabrikian, "Performance analysis for channel estimation with 1-bit ADC and unknown quantization threshold," *IEEE Trans. Signal Process.*, vol. 66, no. 10, pp. 2557–2571, May 2018.

[23] J. Mo, P. Schniter, and R. W. Heath, Jr., "Channel estimation in broadband millimeter wave MIMO systems with few-bit ADCs," *IEEE Trans. Signal Process.*, vol. 66, no. 5, pp. 1141–1154, Mar. 2018.

[24] S. Rao, A. Mezghani, and A. L. Swindlehurst, "Channel estimation in one-bit massive MIMO systems: Angular versus unstructured models," *IEEE J. Sel. Topics Signal Process.*, vol. 13, no. 5, pp. 1017–1031, Sep. 2019.

[25] H. Kim and J. Choi, "Channel estimation for spatially/temporally correlated massive MIMO systems with one-bit ADCs," *EURASIP J. Wireless Commun. Netw.*, vol. 2019, no. 1, pp. 267–281, Dec. 2019.

[26] Y.-S. Jeon, N. Lee, and H. V. Poor, "Robust data detection for MIMO systems with one-bit ADCs: A reinforcement learning approach," *IEEE Trans. Wireless Commun.*, vol. 19, no. 3, pp. 1663–1676, Mar. 2020.

[27] L. V. Nguyen, D. T. Ngo, N. H. Tran, A. L. Swindlehurst, and D. H. N. Nguyen, "Supervised and semi-supervised learning for MIMO blind detection with low-resolution ADCs," *IEEE Trans. Wireless Commun.*, vol. 19, no. 4, pp. 2427–2442, Apr. 2020.

[28] A. B. Üçüncü and A. Ö. Yilmaz, "Oversampling in one-bit quantized massive MIMO systems and performance analysis," *IEEE Trans. Wireless Commun.*, vol. 17, no. 12, pp. 7952–7964, Dec. 2018.

[29] Z. Shao, L. T. N. Landau, and R. C. De Lamare, "Channel estimation for large-scale multiple-antenna systems using 1-bit ADCs and oversampling," *IEEE Access*, vol. 8, pp. 85243–85256, 2020.

[30] M. Shao, Q. Li, W.-K. Ma, and A. M.-C. So, "A framework for one-bit and constant-envelope precoding over multiuser massive MISO channels," *IEEE Trans. Signal Process.*, vol. 67, no. 20, pp. 5309–5324, Oct. 2019.

[31] C. Risi, D. Persson, and E. G. Larsson, "Massive MIMO with 1-bit ADC," 2014, *arXiv:1404.7736*.

[32] D. Abdelhameed, K. Umebayashi, A. Al-Tahmeesschi, I. Atzeni, and A. Tölli, "Enhanced signal detection for massive SIMO communications with 1-bit ADCs," in *Proc. IEEE Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Lucca, Italy, Sep. 2021, pp. 1–5.

[33] S. Rao, G. Seco-Granados, H. Pirzadeh, J. A. Nossek, and A. L. Swindlehurst, "Massive MIMO channel estimation with low-resolution spatial sigma-delta ADCs," *IEEE Access*, vol. 9, pp. 109320–109334, 2021.

[34] I. G. Abrahamson, "Orthant probabilities for the quadrivariate normal distribution," *Ann. Math. Statist.*, vol. 35, no. 4, pp. 1685–1703, Dec. 1964.

---

[6]Note that (124)–(126) are equivalent to (86)–(88) for $K = 1$.
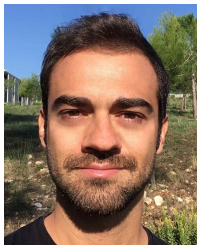
Italo Atzeni (Member, IEEE) received the M.Sc. degree (Hons.) in telecommunications engineering from the University of Cagliari, Italy, in 2010, and the Ph.D. degree (Hons.) in signal theory and communications from the Polytechnic University of Catalonia–BarcelonaTech, Spain, in 2014. From 2014 to 2017, he was a Researcher with the Mathematical and Algorithmic Sciences Laboratory, Paris Research Center, Huawei Technologies, France. From 2017 to 2018, he was a Research Associate with the Communication Systems Department, EURECOM, France. He has previously held a research appointment at The Hong Kong University of Science and Technology, Hong Kong, in 2013. Since 2019, he has been with the Centre for Wireless Communications, University of Oulu, Finland, where he is currently an Assistant Professor. His primary research interests are in communication and information theory, statistical signal processing, convex and distributed optimization theory, and their applications to low-complexity, energy-efficient solutions for multi-antenna communications. He received the Best Paper Award in the Wireless Communications Symposium at IEEE ICC 2019. He was granted the MSCA-IF Grant for the project Device-Centric Low-Complexity High-Frequency Networks (DELIGHT) in 2020. He serves as an Editor for the IEEE COMMUNICATIONS LETTERS.

Antti Tölli (Senior Member, IEEE) received the Dr.Sc.(Tech.) degree in electrical engineering from the University of Oulu, Oulu, Finland, in 2008. From 1998 to 2003, he worked with Nokia Networks as a Research Engineer and the Project Manager both in Finland and Spain. In May 2014, he was granted a Five Year (2014–2019) Academy Research Fellow post by the Academy of Finland. During the academic year 2015–2016, he visited EURECOM, Sophia Antipolis, France, while from August 2018 to June 2019 he was visiting the University of California Santa Barbara, USA. He is currently an Associate Professor with the Centre for Wireless Communications (CWC), University of Oulu. He has authored numerous papers in peer-reviewed international journals and conferences and several patents all in the area of signal processing and wireless communications. His research interests include radio resource management and transceiver design for broadband wireless communications with a special emphasis on distributed interference management in heterogeneous wireless networks. From 2017 to 2021, he served as an Associate Editor for IEEE TRANSACTIONS ON SIGNAL PROCESSING.