

Data-Aided Regularization of Direct-Estimate Combiner in Distributed MIMO Systems

Bikshapathi Gouda, Italo Atzeni, and Antti Tölli

Centre for Wireless Communications, University of Oulu, Finland

E-mail: {bikshapathi.gouda, italo.atzeni, antti.tolli}@oulu.fi

Abstract—This paper explores the data-aided regularization of the direct-estimate combiner in the uplink of a distributed multiple-input multiple-output system. The network-wide combiner can be computed directly from the pilot signal received at each access point, eliminating the need for explicit channel estimation. However, the sample covariance matrix of the received pilot signal that is used in its computation may significantly deviate from the actual covariance matrix when the number of pilot symbols is limited. To address this, we apply a regularization to the sample covariance matrix using a shrinkage coefficient based on the received data signal. Initially, the shrinkage coefficient is determined by minimizing the difference between the sample covariance matrices obtained from the received pilot and data signals. Given the limitations of this approach in interference-limited scenarios, the shrinkage coefficient is iteratively optimized using the sample mean squared error of the hard-decision symbols, which is more closely related to the actual system's performance, e.g., the symbol error rate (SER). Numerical results demonstrate that the proposed regularization of the direct-estimate combiner significantly enhances the SER, particularly when the number of pilot symbols is limited.

Index Terms—Direct-estimate combiner, distributed MIMO, interference rejection, regularization.

I. INTRODUCTION

Distributed multiple-input multiple-output (MIMO) systems consist of geographically separated access points (APs) connected to a central unit (CU) via fronthaul links for data and channel state information (CSI) exchange. The APs coherently serve the user equipments (UEs) by utilizing a network-wide precoder or combiner with the objective of improving the spectral efficiency and providing a uniform service across the network [1]–[3]. Focusing on the uplink, the network-wide combiner can be computed directly from the pilot signal received at each AP without the need for explicit channel estimation. This is referred to as *direct-estimate combiner*, which inherently accounts for potential interference by utilizing the sample covariance matrix of the received pilot signal [2]–[4]. In the absence of statistical CSI, the least-squares (LS) method can be used to compute the direct-estimate combiner. When the number of pilot symbols is limited, a regularization of the sample covariance matrix can be employed to enhance the system's performance. For instance, the use of exponentially weighted regularization across multiple estimates

was investigated in [4] to improve the accuracy of the direct-estimate precoder and combiner under time-varying channels.

Many studies have focused on accurately estimating the interference-plus-noise covariance matrix using received signal samples. Shrinkage techniques have been used to adjust the diagonal elements with a weighted factor determined by the shrinkage coefficient [5], [6]. For instance, the impact of different shrinkage coefficients is analyzed in [7]. Additionally, methods like leave-one-out cross-validation can be employed to further enhance the accuracy of the sample covariance matrix [8], [9]. However, the estimation accuracy of the covariance matrix is usually not directly related to metrics such as the symbol error rate (SER) or signal-to-interference-plus-noise ratio (SINR). To address this, an expected likelihood criterion was proposed in [10] to optimize the shrinkage coefficient, demonstrating effectiveness across a range of SINR values with varying numbers of UEs and interference sources. Furthermore, the shrinkage coefficient was specifically adjusted to maximize the SINR in [11].

This paper considers a distributed MIMO system serving multiple UEs in the uplink. The network-wide combiner is computed at the CU via direct estimation by minimizing the LS objective over the received pilot signal. Due to the limited number of pilot symbols, the sample covariance matrix used in the computation of the direct-estimate combiner may be inaccurate, leading to a degradation of the system's performance. To address this, we apply a regularization to the sample covariance matrix by adjusting its diagonal elements with a shrinkage coefficient. Assuming that the interference statistics do not change during both the pilot and data transmission phases, the shrinkage coefficient is optimized using the received data signal as follows. Initially, a closed-form expression of the shrinkage coefficient is derived by minimizing the difference between the sample covariance matrices obtained from the received pilot and data signals. However, recognizing the limitations of this approach in interference-limited scenarios, we propose an iterative optimization of the shrinkage coefficient that minimizes the sample mean squared error (MSE) of the hard-decision symbols. In this approach, the shrinkage coefficient is iteratively updated using a gradient-based method along with the hard-decision symbols. Numerical results demonstrate that the regularization of the direct-estimate combiner significantly enhances the SER in interference-limited scenarios and with a small number of pilot symbols.

II. SYSTEM MODEL

We consider a set of APs $\mathcal{B} \triangleq \{1, \dots, B\}$, each equipped with M antennas, serving a set of single-antenna UEs $\mathcal{K} \triangleq \{1, \dots, K\}$ in the uplink. During the data transmission phase, the received signal at AP b is given by

$$\mathbf{Y}_b^d \triangleq \sum_{k \in \mathcal{K}} \sqrt{\rho_k} \mathbf{h}_{b,k} \mathbf{d}_k^H + \mathbf{Z}_b^d \in \mathbb{C}^{M \times \tau^d}, \quad (1)$$

where $\rho_k \in \mathbb{R}_+$ is the transmit power of UE k , $\mathbf{h}_{b,k} \in \mathbb{C}^{M \times 1}$ is the uplink channel between UE k and AP b , $\mathbf{d}_k \in \mathbb{C}^{\tau^d \times 1}$ is the data symbol vector transmitted by UE k (with $\mathbb{E}[\|\mathbf{d}_k\|^2] = \tau^d$), and $\mathbf{Z}_b^d \in \mathbb{C}^{M \times \tau^d}$ represents the combination of interference from out-of-cluster UEs (i.e., UEs served by other APs) and additive white Gaussian noise (AWGN) at AP b . Accordingly, the aggregated received data signal across all the APs is

$$\begin{aligned} \mathbf{Y}^d &\triangleq [(\mathbf{Y}_1^d)^T, \dots, (\mathbf{Y}_B^d)^T]^T \\ &= \sum_{k \in \mathcal{K}} \sqrt{\rho_k} \mathbf{h}_k \mathbf{d}_k^H + \mathbf{Z}^d \in \mathbb{C}^{BM \times \tau^d}, \end{aligned} \quad (2)$$

with $\mathbf{h}_k \triangleq [\mathbf{h}_{1,k}^T, \dots, \mathbf{h}_{B,k}^T]^T \in \mathbb{C}^{BM \times 1}$ and $\mathbf{Z}^d \triangleq [(\mathbf{Z}_1^d)^T, \dots, (\mathbf{Z}_B^d)^T]^T \in \mathbb{C}^{BM \times \tau^d}$, where each column of \mathbf{Z}^d is assumed to follow a circularly-symmetric complex normal distribution with zero mean and covariance matrix $\Psi \in \mathbb{C}^{BM \times BM}$. The aggregated received data signal in (2) is obtained at the CU by combining the received signals in (1) forwarded by each AP via fronthaul links.

The soft estimate of \mathbf{d}_k can be obtained at the CU by combining \mathbf{Y}^d with the combiner $\mathbf{w}_k \in \mathbb{C}^{BM \times 1}$ as

$$\hat{\mathbf{d}}_k \triangleq (\mathbf{Y}^d)^H \mathbf{w}_k \in \mathbb{C}^{\tau^d \times 1}. \quad (3)$$

Since the computation of the combiner \mathbf{w}_k requires CSI at the CU, we assume that each UE transmits pilots for the channel estimation along with the data in each coherence interval, as illustrated in Figure 1. Let $\mathbf{p}_k \in \mathbb{C}^{\tau^p \times 1}$ represent the pilot vector transmitted by UE k (with $\|\mathbf{p}_k\|^2 = \tau^p$). During the pilot transmission phase, the received signal at AP b is given by

$$\mathbf{Y}_b^p \triangleq \sum_{k \in \mathcal{K}} \sqrt{\rho_k} \mathbf{h}_{b,k} \mathbf{p}_k^H + \mathbf{Z}_b^p \in \mathbb{C}^{M \times \tau^p}, \quad (4)$$

where \mathbf{Z}_b^p represents the combination of interference from out-of-cluster UEs and AWGN at AP b (similar to \mathbf{Z}_b^d in the data transmission phase). Consequently, the aggregated received pilot signal across all the APs is

$$\begin{aligned} \mathbf{Y}^p &\triangleq [(\mathbf{Y}_1^p)^T, \dots, (\mathbf{Y}_B^p)^T]^T \\ &= \sum_{k \in \mathcal{K}} \sqrt{\rho_k} \mathbf{h}_k \mathbf{p}_k^H + \mathbf{Z}^p \\ &= \mathbf{H} \Omega_\rho^{\frac{1}{2}} \mathbf{P}^H + \mathbf{Z}^p \in \mathbb{C}^{BM \times \tau^p}, \end{aligned} \quad (5)$$

with $\mathbf{H} \triangleq [\mathbf{h}_1, \dots, \mathbf{h}_K] \in \mathbb{C}^{BM \times K}$, $\Omega_\rho \triangleq \text{Diag}([\rho_1, \dots, \rho_K]) \in \mathbb{R}^{K \times K}$, $\mathbf{P} \triangleq [\mathbf{p}_1, \dots, \mathbf{p}_K] \in \mathbb{C}^{\tau^p \times K}$, and $\mathbf{Z}^p \triangleq [(\mathbf{Z}_1^p)^T, \dots, (\mathbf{Z}_B^p)^T]^T \in \mathbb{C}^{BM \times \tau^p}$. Assuming that the interference characteristics remain constant during both the pilot and data transmission phases, each column of \mathbf{Z}^p follows



Figure 1. Uplink transmission of the pilot and data signals within the coherence interval.

the same distribution as the columns of \mathbf{Z}^d . In the following, we discuss the computation of the direct-estimate combiner and its regularization using the received pilot and data signals.

III. REGULARIZATION OF THE DIRECT-ESTIMATE COMBINER

In the absence of statistical CSI, the aggregated direct-estimate combiner for all the UEs, denoted by $\mathbf{W} \triangleq [\mathbf{w}_1, \dots, \mathbf{w}_K] \in \mathbb{C}^{BM \times K}$, can be obtained by minimizing the LS objective over the received pilot signal, i.e.,

$$\underset{\mathbf{W}}{\text{argmin}} \|\mathbf{W}^H \mathbf{Y}^p - \mathbf{P}^H\|_F^2 = \frac{1}{\tau^p} \mathbf{Q}^{-1} \mathbf{Y}^p \mathbf{P}, \quad (6)$$

with \mathbf{Y}^p in (5) and where $\mathbf{Q} \triangleq \frac{1}{\tau^p} \mathbf{Y}^p (\mathbf{Y}^p)^H \in \mathbb{C}^{BM \times BM}$ is the sample covariance matrix of the received pilot signal, which converges to the actual covariance matrix $\mathbf{H} \Omega_\rho \mathbf{H}^H + \Psi$ as $\tau^p \rightarrow \infty$. However, with a limited sample size (i.e., with small τ^p), \mathbf{Q} may not accurately approximate the actual covariance matrix [5]. In this setting, a regularization of the sample covariance matrix using shrinkage techniques can be applied to improve the estimation accuracy. Specifically, we adjust the diagonal elements of \mathbf{Q} with a shrinkage coefficient $\alpha \in (0, 1]$ as [5]

$$\begin{aligned} \mathbf{R}(\alpha) &\triangleq (1 - \alpha) \mathbf{Q} + \alpha \frac{\text{tr}(\mathbf{Q})}{BM} \mathbf{I}_{BM}, \\ &= \mathbf{Q} + \alpha \mathbf{S} \in \mathbb{C}^{BM \times BM}, \end{aligned} \quad (7)$$

with $\mathbf{S} \triangleq \frac{\text{tr}(\mathbf{Q})}{BM} \mathbf{I}_{BM} - \mathbf{Q} \in \mathbb{C}^{BM \times BM}$. In (7), the regularization shrinks the sample covariance matrix towards a diagonal matrix to improve the estimation accuracy based on the received pilot signal.¹ Finally, with the regularized sample covariance matrix, the direct-estimate combiner can be rewritten as

$$\mathbf{W} = \frac{1}{\tau^p} \mathbf{R}(\alpha)^{-1} \mathbf{Y}^p \mathbf{P}. \quad (8)$$

Several methods have been proposed in the literature for computing the shrinkage coefficient to enhance the estimation accuracy of the covariance matrix [5]. A simple and well-known technique is based on the oracle estimator, where the shrinkage coefficient is obtained by minimizing the difference between the regularized sample covariance matrix of the received pilot signal and the actual covariance matrix, i.e.,

$$\begin{aligned} \underset{\alpha}{\text{argmin}} \|\mathbf{R}(\alpha) - (\mathbf{H} \Omega_\rho \mathbf{H}^H + \Psi)\|_F^2, \\ = \frac{1}{\text{tr}(\mathbf{S} \mathbf{S}^H)} \text{Re} \left[\text{tr} \left(\left(\mathbf{H} \Omega_\rho \mathbf{H}^H + \Psi - \mathbf{Q} \right) \mathbf{S} \right) \right]. \end{aligned} \quad (9)$$

¹The shrinkage coefficient can vary for each AP; however, for simplicity, we use a single coefficient across all the APs and optimize it at the CU. Alternatively, it is also possible to estimate the shrinkage coefficient at each AP in a distributed fashion using local CSI.

However, this approach is obviously impractical as it requires the knowledge of \mathbf{H} and Ψ . Therefore, we propose an alternative approach to optimize the shrinkage coefficient based on the received data signal in Section III-A. Additionally, the criterion in (9) based on the received pilot signal may not directly optimize metrics such as the SER or SINR [10]. Hence, we propose to optimize the shrinkage coefficient by iteratively minimizing the sample MSE of the hard-decision symbols in Section III-B.

A. Regularization Based on the Received Data Signal

The oracle estimator in (9) is infeasible due to the required knowledge of \mathbf{H} and Ψ . However, since the number of data symbols is generally larger than the number of pilot symbols, i.e., $\tau^d \geq \tau^p$, we can utilize the sample covariance matrix of the received data signal $\frac{1}{\tau^d} \mathbf{Y}^d (\mathbf{Y}^d)^H$, which also converges to $\mathbf{H} \Omega_p \mathbf{H}^H + \Psi$ as $\tau^d \rightarrow \infty$. Consequently, the shrinkage coefficient can be obtained by minimizing the difference between the regularized sample covariance matrix of the received pilot signal and the sample covariance matrix of the received data signal, i.e.,

$$\begin{aligned} & \underset{\alpha}{\operatorname{argmin}} \left\| \mathbf{R}(\alpha) - \frac{1}{\tau^d} \mathbf{Y}^d (\mathbf{Y}^d)^H \right\|_F^2 \\ &= \frac{1}{\operatorname{tr}(\mathbf{S} \mathbf{S}^H)} \operatorname{Re} \left[\operatorname{tr} \left(\left(\frac{1}{\tau^d} \mathbf{Y}^d (\mathbf{Y}^d)^H - \mathbf{Q} \right) \mathbf{S} \right) \right]. \end{aligned} \quad (10)$$

Indeed, the sample covariance matrix of the received data signal can also be utilized directly to enhance the estimation accuracy of the covariance matrix by employing alternative objective functions, as detailed in [12]. However, the estimation accuracy may be poor in interference-limited scenarios. Moreover, improving the estimation accuracy of the covariance matrix may not directly translate into improvements in metrics such as SER and SINR. Therefore, we propose an iterative optimization of the shrinkage coefficient based on the sample MSE of the hard-decision symbols, which is more closely related to the system's performance in terms of SER and SINR.

B. Iterative Regularization Based on the Hard-Decision Symbols

In this method, the shrinkage coefficient in (7) is iteratively updated using the hard-decision symbols. Specifically, at each iteration and for a fixed shrinkage coefficient α , the soft-estimated symbols of UE k are computed as in (3) using the direct-estimate combiner of UE k , which corresponds to the k th column of \mathbf{W} in (8). The hard-decision symbols are then obtained by mapping the soft-estimated symbols to the set of transmitted data symbols using the minimum-distance criterion, which offers lower complexity than the optimal maximum-likelihood detector and is expressed as

$$\bar{\mathbf{d}}_k \triangleq \underset{\mathbf{d}_k \in \mathcal{D}_k^{\tau^d \times 1}}{\operatorname{argmin}} \|\hat{\mathbf{d}}_k - \mathbf{d}_k\|_2^2, \quad (11)$$

where \mathcal{D}_k is the set of all possible transmitted symbols of UE k . Then, using $\bar{\mathbf{d}}_k$, the shrinkage coefficient α is optimized by minimizing the sample MSE between the soft-estimated

Algorithm 1 Iterative update of the shrinkage coefficient

Data: $\mathbf{P} \in \mathbb{C}^{\tau^p \times K}$, $\mathbf{Y}^p \in \mathbb{C}^{BM \times \tau^p}$, and $\mathbf{Y}^d \in \mathbb{C}^{BM \times \tau^d}$.

Initialization: $\alpha^{(0)} = 0$, $i = 0$, and β .

Until a predefined termination criterion is satisfied, **do**:

- 1) $i \leftarrow i + 1$.
- 2) Compute \mathbf{W} as in (8) with $\alpha = \alpha^{(i-1)}$.
- 3) Compute $\hat{\mathbf{d}}_k$ as in (3) for all the UEs.
- 4) Compute $\bar{\mathbf{d}}_k$ as in (11) for all the UEs.
- 5) Compute $\frac{\partial \epsilon(\alpha)}{\partial \alpha}$ as in (13) and update $\alpha^{(i)}$ as in (14).

End

symbols in (3) and the hard-decision symbols in (11). To this end, we express the average sample MSE across all the UEs as

$$\begin{aligned} \epsilon(\alpha) &\triangleq \frac{1}{K \tau^d} \sum_{k \in \mathcal{K}} \|(\mathbf{Y}^d)^H \mathbf{w}_k - \bar{\mathbf{d}}_k\|_2^2 \\ &= \frac{1}{K \tau^d} \operatorname{tr} \left(\frac{1}{(\tau^p)^2} \mathbf{P}^H (\mathbf{Y}^p)^H \mathbf{R}(\alpha)^{-1} \mathbf{Y}^d (\mathbf{Y}^d)^H \mathbf{R}(\alpha)^{-1} \mathbf{Y}^p \mathbf{P} \right. \\ &\quad \left. - \frac{2}{\tau^p} \operatorname{Re} [\mathbf{P}^H (\mathbf{Y}^p)^H \mathbf{R}(\alpha)^{-1} \mathbf{Y}^d \bar{\mathbf{D}}^H] + \bar{\mathbf{D}} \bar{\mathbf{D}}^H \right), \end{aligned} \quad (12)$$

with $\bar{\mathbf{D}} \triangleq [\bar{\mathbf{d}}_1, \dots, \bar{\mathbf{d}}_K]^T \in \mathbb{C}^{\tau^d \times K}$. In (12), minimizing $\epsilon(\alpha)$ with respect to α can be interpreted as minimizing the cross-validation loss with respect to the ridge parameter in ridge regression [13]. Since $\epsilon(\alpha)$ is not convex with respect to α , we aim to achieve a locally optimal α using a gradient-based method. For this purpose, we write the derivative of $\epsilon(\alpha)$ with respect to α as

$$\begin{aligned} \frac{\partial}{\partial \alpha} \epsilon(\alpha) &= -\frac{2}{K \tau^d} \operatorname{Re} \left[\operatorname{tr} \left(\mathbf{P}^H (\mathbf{Y}^p)^H \mathbf{R}(\alpha)^{-1} \mathbf{S} \mathbf{R}(\alpha)^{-1} \mathbf{Y}^d \right. \right. \\ &\quad \left. \left. \times \left(\frac{1}{(\tau^p)^2} (\mathbf{Y}^d)^H \mathbf{R}(\alpha)^{-1} \mathbf{Y}^p \mathbf{P} + \frac{1}{\tau^p} \bar{\mathbf{D}}^H \right) \right) \right]. \end{aligned} \quad (13)$$

At each iteration i , α can be updated in the direction of the negative derivative as

$$\alpha^{(i)} = \alpha^{(i-1)} - \beta \frac{\partial}{\partial \alpha} \epsilon(\alpha) \Big|_{\alpha=\alpha^{(i-1)}}, \quad (14)$$

where $\beta > 0$ is the step size chosen to promote faster convergence. The iterative update of the shrinkage coefficient α based on the hard-decision symbols is outlined in Algorithm 1, where an example of predefined termination criterion is that the value of α does not change significantly between consecutive iterations.

We point out that $\mathbf{R}(\alpha)$ in (13) does not need to be inverted at each iteration. In fact, since $\mathbf{R}(\alpha)$ retains the same eigenvectors as \mathbf{Q} for any α , one can compute the eigenvalue decomposition of \mathbf{Q} at the beginning of the algorithm and update only the eigenvalues at each iteration, simplifying the inversion of $\mathbf{R}(\alpha)$ to that of a diagonal matrix. Additionally, a stochastic gradient-based method can be employed to compute the derivative in (13) based on a subset of received data symbols, further reducing the computational complexity.

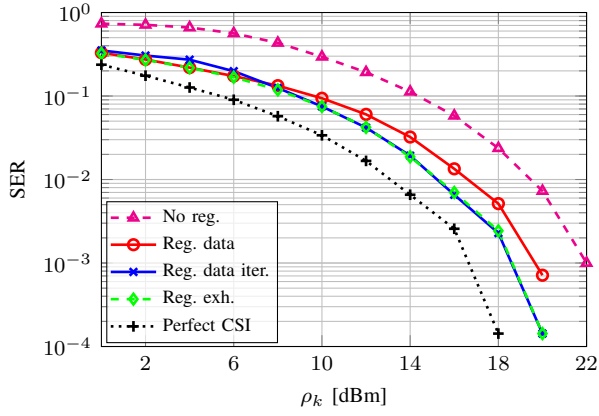


Figure 2. SER of QPSK versus UE transmit power for $K = 6$, $\tau^p = 8$, and $\tau^d = 1000$, without interference.

IV. NUMERICAL RESULTS

We consider a distributed MIMO system with $B = 2$ APs, each equipped with $M = 4$ antennas and placed 100 m apart, serving $K = 6$ single-antenna UEs. Assuming uncorrelated Rayleigh fading, each channel is generated as $\text{vec}(\mathbf{H}_{b,k}) \sim \mathcal{CN}(0, \zeta_{b,k} \mathbf{I}_{MN})$, with $\zeta_{b,k} \triangleq -30.5 - 36.7 \log_{10}(r_{b,k})$ [dB], where $r_{b,k}$ denotes the distance between AP b and UE k [2]. The AWGN power at the APs is fixed to -95 dBm. The proposed regularization methods in Section III-A and Section III-B (Algorithm 1) are referred to as *Reg. data* and *Reg. data iter.*, respectively. These methods are compared with: *i*) no regularization ($\alpha = 0$), referred to as *No reg.*; *ii*) regularization optimized through exhaustive line search to minimize the MSE, referred to as *Reg. exh.*; and *iii*) the ideal case with perfect CSI, referred to as *Perfect CSI*.

In Figure 2, the SER for QPSK is plotted against the UE transmit power, without any external interference source. Both *Reg. data* and *Reg. data iter.* outperform *No reg.*, with *Reg. data iter.* achieving a gain of 3–4 dB compared to *No reg.* and performing closely to *Reg. exh.*. However, at low UE transmit powers, the performance of *Reg. data iter.* is inferior to *Reg. data* due to errors in the hard-decision symbols. As the UE transmit power increases, the SER performance of *Reg. data iter.* surpasses that of *Reg. data*, demonstrating considerable improvements for a fixed UE transmit power.

In Figure 3, the SER for QPSK is plotted against the UE transmit power, with an external interference source. The power of the interference source is set to be 5 dB lower than the UE transmit power, and it transmits random symbols during both the pilot and data transmission phases. The behavior of *No reg.*, *Reg. data iter.*, and *Reg. exh.* is similar to the results shown in Figure 2. However, regularization using *Reg. data*, which aligns the sample covariance matrices obtained from the received pilot and data signals, yields a suboptimal shrinkage coefficient. This is because the signals from random interference sources during the pilot phase exhibit higher correlation due to the limited sample size compared with the data transmission phase. This leads to a degradation in the SER performance, even compared with *No reg.*

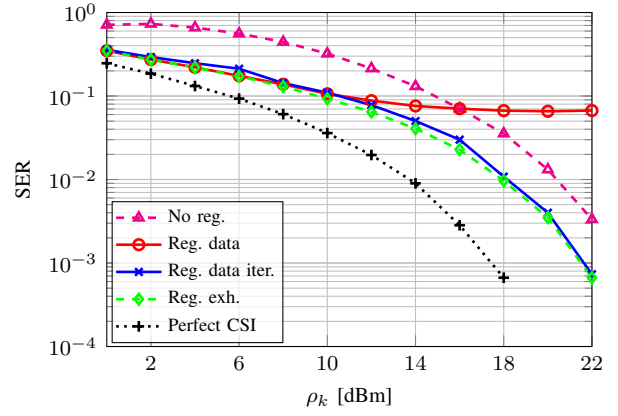


Figure 3. SER of QPSK versus UE transmit power for $K = 6$, $\tau^p = 8$, and $\tau^d = 1000$, with interference.

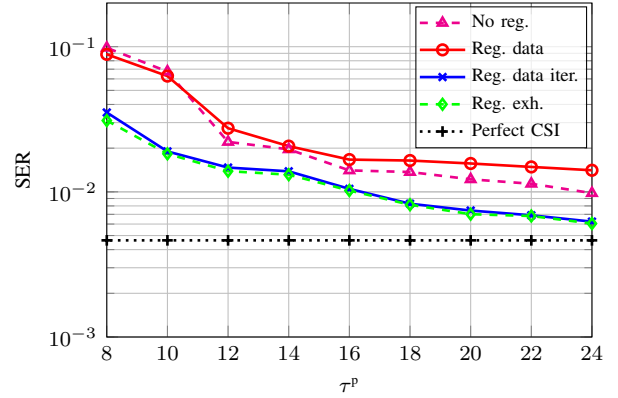


Figure 4. SER of QPSK versus pilot length for $K = 6$, $\rho_k = 15$ dBm, and $\tau^d = 1000$, with interference.

In Figure 4, the SER for QPSK is plotted against the pilot length with an external interference source. The interference characteristics are similar to those used in Figure 3. As the pilot length increases, the performance of *No reg.*, *Reg. data iter.*, *Reg. exh.*, and *Reg. data* improves, approaching the SER of *Perfect CSI*. However, a significant SER gap remains between *Reg. data iter.* and both *No reg.* and *Reg. data*. Furthermore, the performance of *Reg. data iter.* is close to that of *Reg. exh.* for any pilot length.

V. CONCLUSIONS

Considering the uplink of a distributed MIMO system, we investigated the data-aided regularization of the direct-estimate combiner to address the inaccurate estimation of the covariance matrix used in its computation due to the limited number of pilot symbols. Specifically, we applied a shrinkage-based regularization method and initially optimized the shrinkage coefficient by minimizing the difference between the covariance matrices obtained from the received pilot and data signals. Recognizing the limitations of this approach in interference-limited scenarios, we proposed an iterative method to optimize the shrinkage coefficient by minimizing the sample MSE of the hard-decision symbols. Numerical results demonstrated that the proposed methods significantly improve the SER, particularly under interference-limited conditions and with a small number of pilot symbols.

REFERENCES

- [1] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-free massive MIMO versus small cells," *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 1834–1850, 2017.
- [2] I. Atzeni, B. Gouda, and A. Tölili, "Distributed precoding design via over-the-air signaling for cell-free massive MIMO," *IEEE Trans. Wireless Commun.*, vol. 20, no. 2, pp. 1201–1216, 2021.
- [3] B. Gouda, I. Atzeni, and A. Tölili, "Pilot-aided distributed multi-group multicast precoding design for cell-free massive MIMO," *IEEE Trans. Wireless Commun.*, vol. 23, no. 8, pp. 9282–9298, 2024.
- [4] C. Shi, R. A. Berry, and M. L. Honig, "Bi-directional training for adaptive beamforming and power control in interference networks," *IEEE Trans. Signal Process.*, vol. 62, no. 3, pp. 607–618, 2014.
- [5] Y. Chen, "Regularized estimation of high-dimensional covariance matrices." Ph.D. dissertation, University of Michigan, 2011.
- [6] E. Raninen and E. Ollila, "Coupled regularized sample covariance matrix estimator for multiple classes," *IEEE Trans. Signal Process.*, vol. 69, pp. 5681–5692, 2021.
- [7] A. Davydov, V. Sergeev, B. Mondal, A. Papathanassiou, and A. Sengupta, "Robust MMSE-IRC for uplink massive MIMO aided C-RAN network," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, 2020.
- [8] J. Tong, Q. Guo, J. Xi, Y. Yu, and P. J. Schreier, "Choosing the diagonal loading factor for linear signal estimation using cross validation," in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Process. (ICASSP)*, 2016.
- [9] J. Tong, P. J. Schreier, Q. Guo, S. Tong, J. Xi, and Y. Yu, "Shrinkage of covariance matrices for linear signal estimation using cross-validation," *IEEE Trans. Signal Process.*, vol. 64, no. 11, pp. 2965–2975, 2016.
- [10] A. M. Kuzminskiy, Y. I. Abramovich, P. Xiao, and R. Tafazolli, "Uniform expected likelihood solution for interference rejection combining regularization," in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Process. (ICASSP)*, 2016.
- [11] J. Qian, J. Jin, and H. Wang, "Minimum eigenvalue based covariance matrix estimation with limited samples," in *Proc. IEEE Wireless Commun. and Netw. Conf. (WCNC)*, 2024.
- [12] O. Ledoit and M. Wolf, "Shrinkage estimation of large covariance matrices: Keep it simple, statistician?" *J. Multivariate Analysis*, vol. 186, p. 104796, 2021.
- [13] W. T. Stephenson, Z. Frangella, M. Udell, and T. Broderick, "Can we globally optimize cross-validation loss? Quasiconvexity in ridge regression," in *Proc. Adv. in Neural Inf. Proces. Syst.*, 2021.