Enforcing Statistical Orthogonality in Massive MIMO Systems via Covariance Shaping

Placido Mursia, *Member, IEEE,* Italo Atzeni, *Member, IEEE,* Laura Cottatellucci, *Member, IEEE,*
and David Gesbert, *Fellow, IEEE*

Abstract—This paper tackles the problem of downlink data transmission in massive multiple-input multiple-output (MIMO) systems where user equipments (UEs) exhibit high spatial correlation and channel estimation is limited by strong pilot contamination. Signal subspace separation among UEs is, in fact, rarely realized in practice and is generally beyond the control of the network designer (as it is dictated by the physical scattering environment). In this context, we propose a novel statistical beamforming technique, referred to as *MIMO covariance shaping*, that exploits multiple antennas at the UEs and leverages the realistic non-Kronecker structure of massive MIMO channels to target a suitable shaping of the channel statistics performed at the UE-side. To optimize the covariance shaping strategies, we propose a low-complexity block coordinate descent algorithm that is proved to converge to a limit point of the original nonconvex problem. For the two-UE case, this is shown to converge to a stationary point of the original problem. Numerical results illustrate the sum-rate performance gains of the proposed method with respect to spatial multiplexing in scenarios where the spatial selectivity of the base station is not sufficient to separate closely spaced UEs.

Index terms—Covariance shaping, massive MIMO, multi-user MIMO, pilot contamination, statistical beamforming.

I. INTRODUCTION

Massive multiple-input multiple-output (MIMO) is a multi-antenna technology that has great potential to boost the spectral efficiency (SE) of cellular networks by means of highly directional beamforming and spatial multiplexing of many user equipments (UEs) in the same time-frequency resources. It thus plays a pivotal role in current 5G New Radio (NR) implementations [3]–[5] and is expected to maintain this prominence in future wireless generations [6]–[10]. The benefits of massive MIMO can be ascribed to the large number of antennas available at the base station (BS), which we denote by M. In this context, it is shown in [11] that the achievable SE of downlink/uplink massive MIMO systems is unbounded as M grows large and when minimum mean squared error

P. Mursia is with NEC Laboratories Europe GmbH, Heidelberg, Germany, and the Communication Systems Department, EURECOM, France (email: placido.mursia@neclab.eu). I. Atzeni is with the Centre for Wireless Communications, University of Oulu, Finland (email: italo.atzeni@oulu.fi). L. Cottatellucci is with the Department of Electrical, Electronics, and Communication Engineering, Friedrich-Alexander University Erlangen-Nuremberg, Germany (email: laura.cottatellucci@fau.de). D. Gesbert is with the Communication Systems Department, EURECOM, France (email: david.gesbert@eurecom.fr).

(MMSE) precoding/combining is adopted at the BS, which can asymptotically remove any interference. However, in presence of a large number of UEs and finite BS antennas, the aforementioned approach might still be limited by interference and must rely on accurate instantaneous channel state information (CSI). In crowded scenarios, such as outdoor events, transport hubs, and stadiums, the channels of closely spaced UEs exhibit high spatial correlation, which hinders the capability of the BS to separate such UEs during both the channel estimation phase and the data transmission phase. Moreover, when the channel coherence time is limited, non-orthogonal pilots might be used across the UEs during the channel estimation phase, which results in strong pilot contamination [12], [13].

To overcome these issues and facilitate the operations in the massive MIMO regime, several works have proposed to leverage statistical CSI [14]–[25]. This consists mainly in the channel covariance matrix of each UE, which is essentially dictated by the angle spread spanned by the multipath propagation of the signals impinging on the antenna array. This angle spread is often bounded due to the high spatial resolution of the massive array compared with the limited scattering. As a result, the channel covariance matrices in massive MIMO tend to be low-rank and dominated by few major propagation directions [11], [17], [18]. This particular property can be exploited for several applications such as reducing the feedback overhead in the channel estimation phase [14]–[16] and mitigating the interference in the downlink data transmission phase when the UEs exhibit non-overlapping or orthogonal signal subspaces [17], [18]. For example, statistical CSI can be used to precode signals such that their chosen propagation paths do not interfere in average [19], [20]. In this regard, a robust precoding/decoding design based on the average mean squared error (MSE) matrix is derived in [19] under different CSI conditions, whereas [20] provides a lower bound on the ergodic sum rate for the two-UE setting when the UEs are equipped with a single antenna.

There is also a large body of literature focusing on the hybrid precoding problem, where the beamforming applied at the BS is factorized into an inner and an outer precoding matrix [14]–[16], [21], [22]. Here, the former is based on instantaneous CSI while the latter depends only on the second-order channel statistics. This results in reduced pilot length required for the instantaneous effective channel estimation and lower computational complexity associated with the inner precoding design. Moreover, in the context of frequency-division duplex (FDD) systems, such a hybrid precoding method can be used to reduce the size of the resulting effective channels

by exploiting the near-orthogonality of the angle spreads of different UE groups [14]. In [15], this approach is shown to incur a vanishing loss compared with the case of full CSI when the number of BS antennas grows large, while the same result is obtained in [16] for the case of finite number of BS antennas and suitable UE scheduling. In time-division duplex (TDD) systems, existing works have pointed out how the exploitation of statistical CSI enables efficient pilot reuse across the UEs [23]–[25]. In [23], pilot reuse for massive MIMO transmissions over spatially correlated Rayleigh fading channels is proposed when the angle spreads of the UEs with the same pilots are non-overlapping. In [24], the second-order channel statistics are used to precode the pilots, thus reducing the variance of the channel estimation error by a factor that is proportional to the number of UE antennas. A robust channel estimation method dealing with scenarios in which the angle spreads of the desired and interfering channels are overlapping is proposed in [25] and exploits channel separability in the power domain [26].

Most of the above works rely on specific structures of the channel covariance matrices of the UEs in terms of rank and degree of separation among signal subspaces. Such properties are determined by the locations of the UEs combined with the surrounding scattering environments, where both these factors are generally beyond the control of the network designer. Indeed, in many practical scenarios, the UEs exhibit high spatial correlation, e.g., when they are not sufficiently far apart. Hence, the condition of non-overlapping or orthogonal signal subspaces is rarely satisfied in practice [11]. Note that, in this setting, the number of BS antennas needed to effectively separate interfering UEs with MMSE precoding/combining increases with their spatial correlation, i.e., as the UEs get closer to each other, and may become impractical in crowded scenarios. Moreover, existing works are often based on the assumption of Kronecker channel model owing to its simple analytical formulation. However, such a model has been shown to be an over-simplification of the true nature of MIMO channels [27].[1]

A. Contributions

Building on the fact that most future UEs are envisioned to be equipped with a small-to-moderate number of antennas, this paper proposes a novel statistical beamforming technique at the UEs-side, referred to as *MIMO covariance shaping*.[2] While existing methods assume statistical orthogonality among the UEs as a property given by the physical scattering environment and the Kronecker channel model due to its analytical tractability, MIMO covariance shaping aims at modifying the channel statistics of the UEs with highly overlapping channel covariance matrices in order to enforce signal subspace separation in scenarios where the spatial selectivity of the BS is not sufficient to separate such UEs. The proposed approach consists in preemptively applying a statistical beamforming

at the UE-side during both the uplink pilot-aided channel estimation phase and the downlink data transmission phase, and relies uniquely on statistical CSI. In this context, each UE employs its antennas to excite only a subset of all the possible propagation directions towards the BS such that the spatial correlation among the interfering UEs is minimized while preserving enough useful power for effective data transmission. Hence, MIMO covariance shaping is suitable for both pilot decontamination in TDD systems and statistical precoding/combining. Remarkably, the proposed method exploits the realistic non-Kronecker channel structure, which allows to suitably alter the channel statistics perceived at the BS by designing the transceiver at the UE-side. Therefore, it has the unique advantage of turning the generally inconvenient non-Kronecker nature of massive MIMO channels into a benefit. Numerical results show the sum-rate performance gains with respect to a reference scheme employing multiple antennas at the UE for spatial multiplexing in scenarios where the spatial selectivity of the BS is not sufficient to separate UEs with highly overlapping channel covariance matrices and the channel estimation is limited by strong pilot contamination.

The MIMO covariance shaping framework was initially proposed in our prior works [1], [2] and recently used for the minimization of the outage probability in [30]. This work extends the metric proposed in [1] to all pairs of interfering UEs to design a statistical receive beamforming at the UE along with a statistical precoding at the BS. Differently from [30], we argue that MIMO covariance shaping must be adopted for *both* pilot decontamination during the uplink pilot-aided channel estimation phase and combining during the downlink data transmission phase. By doing so, the BS is able to acquire accurate instantaneous estimates of the effective channels, which are then used for efficient downlink precoding and result in a great enhancement of the network performance.

The contributions of this paper are summarized as follows.
- We present the novel concept of MIMO covariance shaping, which aims at designing a suitable shaping of the channel covariance matrices of the UEs to enforce a full or partial separation of their signal subspaces that would be otherwise highly overlapping.
- We point out that the exploitation of the non-Kronecker nature of massive MIMO channels is crucial to suitably alter the channel statistics perceived at the BS by acting at the UE-side.
- We derive a tractable expression of the ergodic achievable sum rate under MIMO covariance shaping with the objective of characterizing the impact of the proposed framework on the system performance.
- We optimize the covariance shaping strategies by minimizing the variance of the inter-UE interference (as a metric to measure the spatial correlation) among the interfering UEs. To this end, we propose a low-complexity block coordinate descent algorithm that is proved to converge to a limit point of the original nonconvex problem. For the two-UE case, this is shown to converge to a stationary point of the original problem.
- We provide numerical results characterizing several scenarios where MIMO covariance shaping outperforms

---

[1]The Kronecker channel model is justified in scenarios where both transmitter and receiver are surrounded by clusters of scatterers that are very far apart, which is hardly realized in practice [28].

[2]Note that the term "covariance shaping" has been used in the past in entirely different contexts, mainly estimation theory (see, e.g., [29]).

spatial multiplexing. In particular, a superior sum-rate performance is achieved when the spatial selectivity of the BS is not sufficient to separate UEs exhibiting high spatial correlation and the channel estimation is limited by strong pilot contamination.

Outline. The rest of the paper is structured as follows. Section II introduces the system model for a general multi-UE MIMO downlink system. Section III presents the concept of MIMO covariance shaping. Section IV proposes efficient methods to optimize the covariance shaping vectors. Section V presents numerical results to evaluate the performance of the proposed framework. Finally, Section VI concludes the paper.

Notation. Lowercase and uppercase boldface letters denote vectors and matrices, respectively, whereas $(\cdot)^{\mathrm{T}}$, $(\cdot)^{\mathrm{H}}$, and $(\cdot)^{*}$ are the transpose, Hermitian transpose, and conjugate operators, respectively. $\|\cdot\|$ and $\|\cdot\|_{\mathrm{F}}$ represent the Euclidean norm for vectors and the Frobenius norm for matrices, respectively, whereas $\mathbb{E}[\cdot]$ and $\mathbb{V}[\cdot]$ are the expectation and variance operators, respectively. $\mathbf{I}_A$ denotes the A-dimensional identity matrix and $\mathbf{0}$ represents the zero vector or matrix with proper dimension. $\mathrm{tr}(\cdot)$ and $\mathrm{vec}(\cdot)$ are the trace and vectorization operators, respectively. $(a_{mn})_{m,n=1}^{M}$ denotes the matrix of size $M \times M$ whose (m,n)th element is given by a_{mn}, whereas $[a_1, \ldots, a_A]$ represents horizontal concatenation. We use $\otimes$ to denote the Kronecker product and $\mathbf{u}_{\min}[\mathbf{A}]$ to denote the eigenvector corresponding to the minimum eigenvalue of matrix $\mathbf{A}$, with $\|\mathbf{u}_{\min}(\mathbf{A})\| = 1$. $\mathbb{1}_A$ denotes the indicator function, which is equal to 1 if condition A is satisfied and to 0 otherwise, and $\lceil \cdot \rceil$ is the ceiling operator. Lastly, $\mathcal{CN}(\mathbf{0}, \mathbf{A})$ is the circularly symmetric complex Gaussian distribution with zero mean and covariance matrix $\mathbf{A}$.

II. SYSTEM MODEL

In this section, we introduce the channel model for a general multi-UE MIMO downlink system. Then, we describe the uplink pilot-aided channel estimation assuming a TDD setting and channel reciprocity between the uplink and the downlink. Finally, we discuss the system model for the downlink data transmission.

A. Channel Model

Consider a multi-UE MIMO system where a BS equipped with M antennas serves K UEs with N antennas each in the downlink. Let $\mathbf{H}_k \triangleq [\mathbf{h}_{k,1}, \ldots, \mathbf{h}_{k,M}] = [\mathbf{g}_{k,1}^{\mathrm{T}}, \ldots, \mathbf{g}_{k,N}^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{C}^{N \times M}$ denote the downlink channel matrix of UE k, where $\mathbf{h}_{k,m} \in \mathbb{C}^{N \times 1}$ and $\mathbf{g}_{k,n} \in \mathbb{C}^{1 \times M}$ are the channel vectors between the mth BS antenna and UE k and between the BS and the nth antenna of UE k, respectively. We assume a correlated Rayleigh fading channel model where the entries of $\mathbf{H}_k$ satisfy $\mathrm{vec}(\mathbf{H}_k) \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Sigma}_k)$ [31, Ch. 3]. Here, the channel covariance matrix $\boldsymbol{\Sigma}_k \in \mathbb{C}^{NM \times NM}$ has the following general structure:

$$\boldsymbol{\Sigma}_k \triangleq \begin{bmatrix} \boldsymbol{\Sigma}_{k,11} & \boldsymbol{\Sigma}_{k,12} & \ldots & \boldsymbol{\Sigma}_{k,1M} \\ \boldsymbol{\Sigma}_{k,12}^{\mathrm{H}} & \boldsymbol{\Sigma}_{k,22} & & \vdots \\ \vdots & & \ddots & \\ \boldsymbol{\Sigma}_{k,1M}^{\mathrm{H}} & \ldots & & \boldsymbol{\Sigma}_{k,MM} \end{bmatrix} \quad (1)$$

where each block $\boldsymbol{\Sigma}_{k,mn} \triangleq \mathbb{E}[\mathbf{h}_{k,m}\mathbf{h}_{k,n}^{\mathrm{H}}] \in \mathbb{C}^{N \times N}$ represents the cross-covariance matrix between the mth and nth columns of $\mathbf{H}_k$. Lastly, we define the covariance matrix seen at UE k as $\mathbf{R}_k \triangleq \mathbb{E}[\mathbf{H}_k\mathbf{H}_k^{\mathrm{H}}] \in \mathbb{C}^{N \times N}$ and the covariance matrix relative to UE k seen at the BS as $\mathbf{T}_k \triangleq \mathbb{E}[\mathbf{H}_k^{\mathrm{H}}\mathbf{H}_k] \in \mathbb{C}^{M \times M}$, respectively. Observe that, in the case of downlink data transmission, $\mathbf{R}_k$ and $\mathbf{T}_k$ represent the receive and transmit covariance matrices, respectively.

B. Uplink Pilot-Aided Channel Estimation

Assuming a TDD setting and channel reciprocity between the uplink and the downlink (see, e.g., [32, Ch. 1.3.5]), the channel matrices $\{\mathbf{H}_k\}_{k=1}^{K}$ are estimated at the BS using antenna-specific uplink pilots, such that N pilot vectors per UE are required. Let $\mathcal{S}_p \triangleq \{k : \text{UE } k \text{ has pilot } \mathbf{P}_p\}$ be the set of UEs sharing the same pilot matrix $\mathbf{P}_p \in \mathbb{C}^{N \times \tau}$, with $p = 1, \ldots, P$. Here, $P < K$ denotes the number of orthogonal pilot matrices and τ represents the pilot length. The orthogonal pilot matrices satisfy $\{\mathbf{P}_p\mathbf{P}_p^{\mathrm{H}} = \frac{\tau}{N}\mathbf{I}_N\}_{p=1}^{P}$, i.e., N orthogonal pilot vectors are assigned to each UE, and $\{\mathbf{P}_p\mathbf{P}_q^{\mathrm{H}} = \mathbf{0}\}_{p \neq q}$. Note that these conditions imply $\tau \geq PN$. Furthermore, $\mathbf{Y} \in \mathbb{C}^{M \times \tau}$ denotes the receive signal at the BS during the uplink pilot-aided channel estimation phase, which is given by

$$\mathbf{Y} \triangleq \sum_{p=1}^{P} \sum_{k \in \mathcal{S}_p} \sqrt{\rho_{\mathrm{UE}}}\mathbf{H}_k^{\mathrm{H}}\mathbf{P}_p + \mathbf{Z} \quad (2)$$

where ρ_{UE} is the transmit power at the UEs and $\mathbf{Z} \in \mathbb{C}^{M \times \tau}$ is the noise term at the BS with elements independently distributed as $\mathcal{CN}(0, \sigma_{\mathrm{BS}}^2)$. Let $\boldsymbol{\Phi}_{k,nn} \triangleq \mathbb{E}[\mathbf{g}_{k,n}^{\mathrm{T}}\mathbf{g}_{k,n}^{*}] \in \mathbb{C}^{M \times M}$ denote the covariance matrix of $\mathbf{g}_{k,n}$. Then, the MMSE estimate of $\mathbf{g}_{k,n}$, with $k \in \mathcal{S}_p$, is given by (see, e.g., [32, Ch. 3.2])

$$\hat{\mathbf{g}}_{k,n}^{\mathrm{H}} \triangleq \frac{N}{\tau\sqrt{\rho_{\mathrm{UE}}}}\boldsymbol{\Phi}_{k,nn}\mathbf{Q}_{k,nn}^{-1}\mathbf{Y}\mathbf{P}_p^{\mathrm{H}}\mathbf{e}_n \in \mathbb{C}^{M \times 1} \quad (3)$$

where $\mathbf{e}_n \in \mathbb{R}^{N \times 1}$ is the nth column of $\mathbf{I}_N$, $\mathbf{Q}_{k,nn} \triangleq \left(\boldsymbol{\Phi}_{k,nn} + \sum_{j \in \mathcal{S}_p \setminus \{k\}} \boldsymbol{\Phi}_{j,nn} + \frac{N}{\tau\varrho_{\mathrm{UE}}}\mathbf{I}_M\right) \in \mathbb{C}^{M \times M}$ is the normalized covariance matrix of the receive signal after correlation with $\mathbf{P}_p^{\mathrm{H}}\mathbf{e}_n$, and $\varrho_{\mathrm{UE}} \triangleq \frac{\rho_{\mathrm{UE}}}{\sigma_{\mathrm{BS}}^2}$. Note that (3) can be seen as a superposition of channels estimated using the same pilot $\mathbf{P}_p^{\mathrm{H}}\mathbf{e}_n$, which cannot be distinguished by the BS: this phenomenon is referred to as pilot contamination [12]. Finally, the MMSE estimate of $\mathbf{H}_k$ is obtained as $\hat{\mathbf{H}}_k \triangleq [\hat{\mathbf{g}}_{k,1}^{\mathrm{T}}, \ldots, \hat{\mathbf{g}}_{k,N}^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{C}^{N \times M}$.

C. Downlink Data Transmission

In the following, we focus on the downlink data transmission and assume that the BS transmits $L_k \leq N$ independent symbols to UE k. We denote the transmit symbol vector for UE k by $\mathbf{s}_k \in \mathbb{C}^{L_k \times 1}$, with $\mathbb{E}[\mathbf{s}_k\mathbf{s}_k^{\mathrm{H}}] = \mathbf{I}_{L_k}$, and introduce the multi-UE transmit symbol vector $\mathbf{s} \triangleq [\mathbf{s}_1^{\mathrm{T}}, \ldots, \mathbf{s}_K^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{C}^{L \times 1}$, with $L \triangleq \sum_{k=1}^{K} L_k$. Before the transmission, the BS precodes $\mathbf{s}$ using the multi-UE precoding matrix $\mathbf{W} \triangleq [\mathbf{W}_1, \ldots, \mathbf{W}_K] \in \mathbb{C}^{M \times L}$, where $\mathbf{W}_k \triangleq [\mathbf{w}_{k,1}, \ldots, \mathbf{w}_{k,L_k}] \in \mathbb{C}^{M \times L_k}$ is the precoding matrix corresponding to $\mathbf{s}_k$ and $\mathbf{W}$

satisfies the power constrain $\|\mathbf{W}\|_{\mathrm{F}}^2 = 1$. The receive signal at UE k is thus given by

$$\mathbf{y}_k \triangleq \sqrt{\rho_{\mathrm{BS}}}\mathbf{H}_k\mathbf{W}_k\mathbf{s}_k + \sqrt{\rho_{\mathrm{BS}}}\sum_{j\neq k}\mathbf{H}_k\mathbf{W}_j\mathbf{s}_j + \mathbf{z}_k \in \mathbb{C}^{N\times 1} \tag{4}$$

where ρ_{BS} is the transmit power at the BS and $\mathbf{z}_k \sim \mathcal{CN}(0, \sigma_{\mathrm{UE}}^2\mathbf{I}_N)$ is the noise term at the UEs. Finally, UE k decodes $\mathbf{s}_k$ as $\hat{\mathbf{s}}_k \triangleq \mathbf{V}_k^{\mathrm{H}}\mathbf{y}_k \in \mathbb{C}^{L_k\times 1}$, where $\mathbf{V}_k \triangleq [\mathbf{v}_{k,1}, \ldots, \mathbf{v}_{k,L_k}] \in \mathbb{C}^{N\times L_k}$ is the corresponding combining matrix. Assuming perfect CSI at the UEs, the resulting sum rate is given by

$$R \triangleq \sum_{k=1}^K \sum_{\ell=1}^{L_k} \log_2\left(1 + \frac{|\mathbf{v}_{k,\ell}^{\mathrm{H}}\mathbf{H}_k\mathbf{w}_{k,\ell}|^2}{\sum_{(j,s)\neq(k,\ell)}|\mathbf{v}_{k,\ell}^{\mathrm{H}}\mathbf{H}_k\mathbf{w}_{j,s}|^2 + \frac{1}{\varrho_{\mathrm{BS}}}\|\mathbf{v}_{k,\ell}\|^2}\right) \tag{5}$$

with $\varrho_{\mathrm{BS}} \triangleq \frac{\rho_{\mathrm{BS}}}{\sigma_{\mathrm{UE}}^2}$.

III. COVARIANCE SHAPING AT THE UE-SIDE

Although MMSE precoding/combining can asymptotically remove any interference as M grows large [11], this no longer holds in the presence of a large number of UEs and finite BS antennas. In particular, the number of BS antennas needed to effectively separate interfering UEs with MMSE precoding/combining increases with their spatial correlation, i.e., as the UEs get closer to each other, and may become impractical in crowded scenarios. In this setting, the BS can spatially separate signals corresponding to different UEs and mitigate or eliminate pilot contamination if their channel covariance matrices lie on orthogonal supports, i.e., if $\mathbf{\Sigma}_k\mathbf{\Sigma}_j = \mathbf{0}$ for a given pair of UEs k and j (see, e.g., [11]). However, the degree of statistical orthogonality among the UEs is determined by their locations combined with the surrounding scattering environments and both these factors are generally beyond the control of network designers. Hence, signal subspace separation among the UEs rarely occurs in practice.

In this context, we propose a novel method relying uniquely on statistical CSI and referred to as *MIMO covariance shaping* (or simply *covariance shaping*), which is applied at the UE-side to enforce the aforementioned signal subspace separation in scenarios where the spatial selectivity of the BS is not sufficient to separate the UEs. While covariance shaping is conceived in such a way that it is agnostic to the operating frequency, we envision its application especially in sub-6 GHz frequency bands, where the availability of BS antennas may be limited and further exacerbates the lack of sufficient spatial selectivity. According to covariance shaping, the UEs preemptively apply a statistical beamforming vector, different for each UE, that aims at spatially separating their transmissions, thus drastically reducing both pilot contamination and interference. Here, the original MIMO channel of each UE is transformed into an effective multiple-input single-output (MISO) channel by combining the transmit/receive signal with the corresponding statistical beamforming vector. To this end, the knowledge of the channel covariance matrices of the UEs is required. Although the time scale at which such information must be acquired and used to compute the covariance shaping vectors increases with the mobility of the UEs (or, more generally, as the channel coherence time reduces), this still happens much less frequently than the instantaneous uplink pilot-aided channel estimation. Remarkably, the proposed method exploits the realistic non-Kronecker structure of massive MIMO channels that allows to suitably alter the channel statistics perceived at the BS by acting at the UE-side, thus turning a generally inconvenient model into a benefit.

Let $\mathbf{v}_k \in \mathbb{C}^{N\times 1}$ denote the statistical beamforming vector preemptively applied at UE k: in the rest of the paper, we refer to $\mathbf{v}_k$ as *covariance shaping vector*.[3] Hence, the original MIMO channel $\mathbf{H}_k$ between the BS and each UE k is transformed into the effective MISO channel $\bar{\mathbf{g}}_k \triangleq \mathbf{v}_k^{\mathrm{H}}\mathbf{H}_k \in \mathbb{C}^{1\times M}$. In this setting, it follows that $\bar{\mathbf{g}}_k \sim \mathcal{CN}(\mathbf{0}, \bar{\mathbf{\Phi}}_k)$, where $\bar{\mathbf{\Phi}}_k \in \mathbb{C}^{M\times M}$ is the effective channel covariance matrix defined as

$$\bar{\mathbf{\Phi}}_k \triangleq \mathbb{E}[\bar{\mathbf{g}}_k^{\mathrm{T}}\bar{\mathbf{g}}_k^*] \tag{6}$$

$$= \left((\mathbf{I}_M \otimes \mathbf{v}_k^{\mathrm{H}})\mathbf{\Sigma}_k(\mathbf{I}_M \otimes \mathbf{v}_k)\right)^{\mathrm{T}} \tag{7}$$

with $\mathbf{\Sigma}_k$ introduced in (1) and where $\mathbb{E}[\|\bar{\mathbf{g}}_k\|^2] = \mathrm{tr}(\bar{\mathbf{\Phi}}_k)$. In the rest of this section, we describe how the two phases of uplink pilot-aided channel estimation and downlink data transmission are modified under covariance shaping and we provide a tractable expression of the resulting ergodic achievable sum rate. The optimization of the covariance shaping vectors is discussed in Section IV. For ease of exposition, we focus on the case where covariance shaping is applied to all the UEs in the systems. Nonetheless, the proposed method may be applied to one or more subsets of closely spaced UEs, e.g., those who cannot be effectively separated with MMSE precoding/combining. In this context, UEs adopting covariance shaping and others employing the multiple antennas for spatial multiplexing can be served simultaneously by the BS in a transparent manner.

A. Uplink Pilot-Aided Channel Estimation

To estimate the effective channels $\{\bar{\mathbf{g}}_k\}_{k=1}^K$ resulting from covariance shaping, the BS assigns the same pilot vector $\mathbf{p}_p \in \mathbb{C}^{1\times\tau}$ to all UEs $k \in \mathcal{S}_p$, with $\{\|\mathbf{p}_p\|^2 = \tau\}_{p=1}^P$ and $\{\mathbf{p}_p\mathbf{p}_q^{\mathrm{H}} = 0\}_{p\neq q}$. The receive signal at the BS during the uplink pilot-aided channel estimation phase, which we denote by $\bar{\mathbf{Y}} \in \mathbb{C}^{M\times\tau}$, is given by (cf. (2))

$$\bar{\mathbf{Y}} \triangleq \sum_{p=1}^P \sum_{k\in\mathcal{S}_p} \sqrt{\rho_{\mathrm{UE}}}\bar{\mathbf{g}}_k^{\mathrm{H}}\mathbf{p}_p + \mathbf{Z}. \tag{8}$$

Then, the MMSE estimate of $\bar{\mathbf{g}}_k$, with $k \in \mathcal{S}_p$, is given by (cf. (3))

$$\hat{\bar{\mathbf{g}}}_k^{\mathrm{H}} \triangleq \frac{1}{\tau\sqrt{\rho_{\mathrm{UE}}}}\bar{\mathbf{\Phi}}_k\mathbf{Q}_k^{-1}\bar{\mathbf{Y}}\mathbf{p}_p^{\mathrm{H}} \in \mathbb{C}^{M\times 1} \tag{9}$$

where $\mathbf{Q}_k \triangleq \bar{\mathbf{\Phi}}_k + \sum_{j\in\mathcal{S}_p\backslash\{k\}}\bar{\mathbf{\Phi}}_j + \frac{1}{\tau\varrho_{\mathrm{UE}}}\mathbf{I}_M \in \mathbb{C}^{M\times M}$ is the normalized covariance matrix of the receive signal after

---

[3]Note that the UEs can be configured for transmission with covariance shaping by the BS in the same way as for codebook-based or non-codebook-based transmission in the current 5G NR implementations (see, e.g., [33, Ch. 11.3] for more details).

$$\gamma_k = \frac{\mathrm{tr}(\bar{\boldsymbol{\Phi}}_k \mathbf{Q}_k^{-1} \bar{\boldsymbol{\Phi}}_k)^2}{\sum_{j=1}^K \mathrm{tr}(\bar{\boldsymbol{\Phi}}_k \bar{\boldsymbol{\Phi}}_j \mathbf{Q}_j^{-1} \bar{\boldsymbol{\Phi}}_j) + \sum_{j \in \mathcal{S}_p \setminus \{k\}} \mathrm{tr}(\bar{\boldsymbol{\Phi}}_k \mathbf{Q}_j^{-1} \bar{\boldsymbol{\Phi}}_j)^2 + \frac{1}{\varrho_{\mathrm{BS}}} \|\mathbf{v}_k\|^2 \sum_{j=1}^K \mathrm{tr}(\bar{\boldsymbol{\Phi}}_j)} \tag{16}$$

correlation with $\mathbf{p}_p^{\mathrm{H}}$. Note that the estimation of the effective channels only requires one pilot vector per UE. Hence, the application of covariance shaping can potentially reduce the pilot length with respect to the estimation of the channel matrices $\{\mathbf{H}_k\}_{k=1}^K$ described in Section II-B, which requires N pilot vectors per UE.

B. Downlink Data Transmission

Focusing on the downlink data transmission, the BS now transmits only one symbol $s_k \in \mathbb{C}$ to each UE k, i.e., $\{L_k = 1\}_{k=1}^K$. Hence, we have $\mathbf{s} = [s_1, \ldots, s_K]^{\mathrm{T}} \in \mathbb{C}^{K \times 1}$ and the multi-UE precoding matrix becomes $\mathbf{W} = [\mathbf{w}_1, \ldots, \mathbf{w}_K] \in \mathbb{C}^{M \times K}$, where $\mathbf{w}_k \in \mathbb{C}^{M \times 1}$ is the precoding vector corresponding to s_k. The receive signal at UE k is thus given by (cf. (4))

$$\bar{y}_k \triangleq \sqrt{\rho_{\mathrm{BS}}} \bar{\mathbf{g}}_k \mathbf{w}_k s_k + \sqrt{\rho_{\mathrm{BS}}} \sum_{j \neq k} \bar{\mathbf{g}}_k \mathbf{w}_j s_j + \mathbf{v}_k^{\mathrm{H}} \mathbf{z}_k \in \mathbb{C} \tag{10}$$

and, assuming perfect CSI at the UEs, the resulting sum rate is given by (cf. (5))

$$\bar{R} \triangleq \sum_{k=1}^K \log_2 \left(1 + \frac{|\bar{\mathbf{g}}_k \mathbf{w}_k|^2}{\sum_{j \neq k} |\bar{\mathbf{g}}_k \mathbf{w}_j|^2 + \frac{1}{\varrho_{\mathrm{BS}}} \|\mathbf{v}_k\|^2} \right). \tag{11}$$

C. Ergodic Achievable Sum Rate

We now analyze the sum rate as a function of the covariance shaping vectors with the objective of characterizing the impact of the proposed framework on the system performance. In particular, we derive a tractable expression of the ergodic achievable sum rate by assuming that each UE k does not know the effective scalar channel $\bar{\mathbf{g}}_k \mathbf{w}_k = \mathbf{v}_k^{\mathrm{H}} \mathbf{H}_k \mathbf{w}_k$ instantaneously but only its expected value. Building on [32, Thm. 4.6], we can express the receive signal in (10) as

$$\bar{y}_k = \sqrt{\rho_{\mathrm{BS}}} \mathbb{E}[\bar{\mathbf{g}}_k \mathbf{w}_k] s_k + z_k' \tag{12}$$

where the effective noise term $z_k' \triangleq \sqrt{\rho_{\mathrm{BS}}} (\bar{\mathbf{g}}_k \mathbf{w}_k - \mathbb{E}[\bar{\mathbf{g}}_k \mathbf{w}_k]) s_k + \sqrt{\rho_{\mathrm{BS}}} \sum_{j \neq k} \bar{\mathbf{g}}_k \mathbf{w}_j s_j + \mathbf{v}_k^{\mathrm{H}} \mathbf{z}_k$ accounts for the lack of instantaneous CSI in addition to the interference and the noise at UE k. A lower bound on the signal-to-interference-plus-noise ratio (SINR) achievable by UE k can be obtained by treating z_k' as Gaussian noise, which yields the effective SINR

$$\gamma_k \triangleq \frac{\left| \mathbb{E}[\bar{\mathbf{g}}_k \mathbf{w}_k] \right|^2}{\mathbb{V}[\bar{\mathbf{g}}_k \mathbf{w}_k] + \sum_{j \neq k} \mathbb{E}[|\bar{\mathbf{g}}_k \mathbf{w}_j|^2] + \frac{1}{\varrho_{\mathrm{BS}}} \|\mathbf{v}_k\|^2}. \tag{13}$$

The expression in (13) can be further simplified by considering the case where the BS adopts maximum ratio transmission (MRT) precoding. In this setting, the multi-UE precoding matrix can be written as

$$\mathbf{W} = \frac{\hat{\bar{\mathbf{H}}}^{\mathrm{H}}}{\sqrt{\mathbb{E}[\|\bar{\mathbf{H}}\|_{\mathrm{F}}^2]}} \tag{14}$$

where $\bar{\mathbf{H}} \triangleq [\bar{\mathbf{g}}_1^{\mathrm{T}}, \ldots, \bar{\mathbf{g}}_K^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{C}^{K \times M}$ and $\hat{\bar{\mathbf{H}}} \triangleq [\hat{\bar{\mathbf{g}}}_1^{\mathrm{T}}, \ldots, \hat{\bar{\mathbf{g}}}_K^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{C}^{K \times M}$ contain the effective channels and their MMSE estimates defined in (9), respectively. Hence, when MRT precoding is adopted at the BS, the effective SINR of UE k in (13), with $k \in \mathcal{S}_p$, reduces to

$$\gamma_k = \frac{\left| \mathbb{E}[\bar{\mathbf{g}}_k \hat{\bar{\mathbf{g}}}_k^{\mathrm{H}}] \right|^2}{\mathbb{V}[\bar{\mathbf{g}}_k \hat{\bar{\mathbf{g}}}_k^{\mathrm{H}}] + \sum_{j \neq k} \mathbb{E}[|\bar{\mathbf{g}}_k \hat{\bar{\mathbf{g}}}_j^{\mathrm{H}}|^2] + \frac{1}{\varrho_{\mathrm{BS}}} \|\mathbf{v}_k\|^2 \mathbb{E}[\|\bar{\mathbf{H}}\|_{\mathrm{F}}^2]}. \tag{15}$$

This can be rewritten in closed form as in (16), shown at the top of this page, by substituting the following terms: $\mathbb{E}[\bar{\mathbf{g}}_k \hat{\bar{\mathbf{g}}}_k^{\mathrm{H}}] = \mathrm{tr}(\bar{\boldsymbol{\Phi}}_k \mathbf{Q}_k^{-1} \bar{\boldsymbol{\Phi}}_k)$, $\mathbb{V}[\bar{\mathbf{g}}_k \hat{\bar{\mathbf{g}}}_k^{\mathrm{H}}] = \mathrm{tr}(\bar{\boldsymbol{\Phi}}_k^2 \mathbf{Q}_k^{-1} \bar{\boldsymbol{\Phi}}_k)$, $\mathbb{E}[|\bar{\mathbf{g}}_k \hat{\bar{\mathbf{g}}}_j^{\mathrm{H}}|^2] = \mathrm{tr}(\bar{\boldsymbol{\Phi}}_k \bar{\boldsymbol{\Phi}}_j \mathbf{Q}_j^{-1} \bar{\boldsymbol{\Phi}}_j) + \mathbb{1}_{j \in \mathcal{S}_p} \mathrm{tr}(\bar{\boldsymbol{\Phi}}_k \mathbf{Q}_j^{-1} \bar{\boldsymbol{\Phi}}_j)^2$, and $\mathbb{E}[\|\bar{\mathbf{H}}\|_{\mathrm{F}}^2] = \sum_{k=1}^K \mathrm{tr}(\bar{\boldsymbol{\Phi}}_k)$. We refer to Appendix A for the detailed derivations. Note that, in the case of perfect channel estimation, i.e., when $\varrho_{\mathrm{UE}} \to \infty$ and all the UEs have orthogonal pilots, we have $\{\bar{\boldsymbol{\Phi}}_k \mathbf{Q}_k^{-1} \bar{\boldsymbol{\Phi}}_k = \bar{\boldsymbol{\Phi}}_k\}_{k=1}^K$ and $\{\mathbb{1}_{j \in \mathcal{S}_p} = 0\}_{j \neq k}$. In this context, the effective SINR in (16) simplifies as

$$\gamma_k = \frac{\mathrm{tr}(\bar{\boldsymbol{\Phi}}_k)^2}{\sum_{j=1}^K \mathrm{tr}(\bar{\boldsymbol{\Phi}}_k \bar{\boldsymbol{\Phi}}_j) + \frac{1}{\varrho_{\mathrm{BS}}} \|\mathbf{v}_k\|^2 \sum_{j=1}^K \mathrm{tr}(\bar{\boldsymbol{\Phi}}_j)}. \tag{17}$$

Finally, Proposition 1 presents an ergodic achievable sum rate with MRT in the two cases of perfect and imperfect channel estimation. The proof follows from [32, Thm. 4.6] and is thus omitted.

Proposition 1. *Assume that the BS adopts MRT precoding. Then, an ergodic achievable sum rate is given by*

$$\bar{R}^{\mathrm{lb}} = \sum_{k=1}^K \log_2(1 + \gamma_k) \tag{18}$$

with γ_k defined in (16) and in (17) for imperfect and perfect channel estimation, respectively. Furthermore, the following limit holds as $M \to \infty$:

$$\lim_{M \to \infty} \bar{R}^{\mathrm{lb}} = \mathbb{E}[\bar{R}] \tag{19}$$

with $\bar{R}$ defined in (11).

IV. COVARIANCE SHAPING OPTIMIZATION

In this section, we address the sum rate maximization through a proper design of the covariance shaping vectors at the UEs. To this end, we consider the variance of the inter-UE interference as a metric to measure the spatial correlation or, in other words, the degree of statistical orthogonality, between two interfering UEs. In the following, we first consider the simple case of $K = 2$ and then extend the resulting analysis to the general case of $K \geq 2$.

Let us define the inter-UE interference between UEs k and j after applying covariance shaping as

$$\Omega(\mathbf{v}_k, \mathbf{v}_j) \triangleq \frac{\bar{\mathbf{g}}_k(\mathbf{v}_k)\bar{\mathbf{g}}_j^{\mathrm{H}}(\mathbf{v}_j)}{\sqrt{\mathbb{E}\big[\|\bar{\mathbf{g}}_k(\mathbf{v}_k)\|^2\big]\mathbb{E}\big[\|\bar{\mathbf{g}}_j(\mathbf{v}_j)\|^2\big]}} \qquad (20)$$

where the notation $\bar{\mathbf{g}}_k(\mathbf{v}_k)$ makes explicit the dependence of the effective channel $\bar{\mathbf{g}}_k$ on the corresponding covariance shaping vector $\mathbf{v}_k$. The effective channels of UEs k and j yield asymptotic *favorable propagation* if they satisfy [32, Ch. 2.5.2]

$$\lim_{M\to\infty} \Omega(\mathbf{v}_k, \mathbf{v}_j) = 0. \qquad (21)$$

For a practical number of BS antennas M, a meaningful performance metric is the variance of $\Omega(\mathbf{v}_k, \mathbf{v}_j)$, expressed as

$$\delta(\mathbf{v}_k, \mathbf{v}_j) \triangleq \mathbb{V}\big[\Omega(\mathbf{v}_k, \mathbf{v}_j)\big] \qquad (22)$$

$$= \frac{\mathrm{tr}\big(\bar{\mathbf{\Phi}}_k(\mathbf{v}_k)\bar{\mathbf{\Phi}}_j(\mathbf{v}_j)\big)}{\mathrm{tr}\big(\bar{\mathbf{\Phi}}_k(\mathbf{v}_k)\big)\mathrm{tr}\big(\bar{\mathbf{\Phi}}_j(\mathbf{v}_j)\big)} \qquad (23)$$

where the notation $\bar{\mathbf{\Phi}}_k(\mathbf{v}_k)$ makes explicit the dependence of the effective channel covariance matrix $\bar{\mathbf{\Phi}}_k$ on the corresponding covariance shaping vector $\mathbf{v}_k$. Note that the terms in the numerator and denominator of (23) can be written as

$$\mathrm{tr}\big(\bar{\mathbf{\Phi}}_k(\mathbf{v}_k)\bar{\mathbf{\Phi}}_j(\mathbf{v}_j)\big) = \sum_{m,n=1}^{M} \mathbf{v}_k^{\mathrm{H}}\mathbf{\Sigma}_{k,mn}\mathbf{v}_k\mathbf{v}_j^{\mathrm{H}}\mathbf{\Sigma}_{j,nm}\mathbf{v}_j \quad (24)$$

$$= \sum_{m,n=1}^{M} \mathbf{v}_k^{\mathrm{H}}\mathbf{\Sigma}_{k,mn}\mathbf{v}_k\mathbf{v}_j^{\mathrm{H}}\mathbf{\Sigma}_{j,mn}^{\mathrm{H}}\mathbf{v}_j, \quad (25)$$

$$\mathrm{tr}\big(\bar{\mathbf{\Phi}}_k(\mathbf{v}_k)\big) = \mathbf{v}_k^{\mathrm{H}}\bigg(\sum_{m=1}^{M}\mathbf{\Sigma}_{k,mm}\bigg)\mathbf{v}_k \qquad (26)$$

where we recall that $\mathbf{\Sigma}_{k,mn}$ represents the cross-covariance matrix between the mth and nth columns of the original channel $\mathbf{H}_k$, i.e., before applying covariance shaping, as defined in (1). Observe that $\delta(\mathbf{v}_k, \mathbf{v}_j) = 0$ implies $\bar{\mathbf{\Phi}}_k\bar{\mathbf{\Phi}}_j = \mathbf{0}$, i.e., $\bar{\mathbf{\Phi}}_k$ and $\bar{\mathbf{\Phi}}_j$ lie on orthogonal supports [11]. Indeed, considering the eigenvalue decomposition of the effective channel covariance matrices, which may be written as $\{\bar{\mathbf{\Phi}}_i = \mathbf{U}_i\mathbf{\Lambda}_i\mathbf{U}_i^{\mathrm{H}}\}_{i\in\{k,j\}}$, the condition of statistical orthogonality is satisfied if $\mathbf{U}_k = \mathbf{U}_j$ and $\mathrm{tr}(\mathbf{\Lambda}_k\mathbf{\Lambda}_j) = 0$, which implies the rank deficiency of both $\bar{\mathbf{\Phi}}_k$ and $\bar{\mathbf{\Phi}}_j$. Clearly, in the general case, this imposes M^2 conditions whereas only $2N$ variables can be adjusted: this means that the resulting system of equations can be solved when $N \geq \frac{M^2}{2}$, which is generally not verified in practice since $M \gg KN$ in massive MIMO scenarios. Hence, since full signal subspace separation can hardly be achieved by simple transceiver design at the UE-side, it is of interest to minimize the spatial correlation between each pair of interfering UEs.

A. Two-UE Case

In the two-UE case, i.e., for $K = 2$, the covariance shaping vectors of UEs k and j are computed by solving the optimization problem

$$\begin{aligned} \underset{\mathbf{v}_k, \mathbf{v}_j}{\text{minimize}} \quad & \delta(\mathbf{v}_k, \mathbf{v}_j) \\ \text{subject to} \quad & \|\mathbf{v}_k\|^2 = \|\mathbf{v}_j\|^2 = 1 \end{aligned} \qquad \text{(P1)}$$

Algorithm 1 (Covariance shaping: Alternating optimization algorithm)

Data: $\mathbf{v}_k^{(0)}, \mathbf{v}_j^{(0)}, \mathbf{\Sigma}_k, \mathbf{\Sigma}_j$, and the accuracy ϵ. Fix $n = 1$.

`While` $\big|\delta(\mathbf{v}_k^{(n)}, \mathbf{v}_j^{(n)}) - \delta(\mathbf{v}_k^{(n-1)}, \mathbf{v}_j^{(n-1)})\big|/\delta(\mathbf{v}_k^{(n)}, \mathbf{v}_j^{(n)}) > \epsilon$

   `S.1:` Given $\mathbf{v}_j^{(n-1)}$, compute $\{\eta_{j,mn}(\mathbf{v}_j^{(n-1)})\}_{m,n=1}^M$ as in (27).

   `S.2:` Compute $\mathbf{v}_k^{(n)}$ as in (29).

   `S.3:` Given $\mathbf{v}_k^{(n)}$, compute $\{\eta_{k,mn}(\mathbf{v}_k^{(n)})\}_{m,n=1}^M$ as in (27).

   `S.4:` Compute $\mathbf{v}_j^{(n)}$ as in (29).

`End`

`S.5:` Fix $\mathbf{v}_k = \mathbf{v}_k^{(n)}$ and $\mathbf{v}_j = \mathbf{v}_j^{(n)}$.

---

with $\delta(\mathbf{v}_k, \mathbf{v}_j)$ defined in (23). Although problem (P1) is not convex in either $\mathbf{v}_k$ or $\mathbf{v}_j$, a suboptimal solution can be efficiently obtained via alternating optimization, as suggested in [1]. Let us introduce the definition

$$\eta_{j,mn}(\mathbf{v}_j) \triangleq \frac{\mathbf{v}_j^{\mathrm{H}}\mathbf{\Sigma}_{j,mn}^{\mathrm{H}}\mathbf{v}_j}{\mathbf{v}_j^{\mathrm{H}}\big(\sum_{m=1}^{M}\mathbf{\Sigma}_{j,mm}\big)\mathbf{v}_j}. \qquad (27)$$

The optimal covariance shaping vector of UE k for a given $\mathbf{v}_j$, denoted by $\mathbf{v}_k^\star$, is obtained as

$$\mathbf{v}_k^\star = \underset{\mathbf{v}_k : \|\mathbf{v}_k\|^2 = 1}{\operatorname{argmin}} \frac{\sum_{m,n=1}^{M}\mathbf{v}_k^{\mathrm{H}}\big(\eta_{j,mn}(\mathbf{v}_j)\mathbf{\Sigma}_{k,mn}\big)\mathbf{v}_k}{\mathbf{v}_k^{\mathrm{H}}\big(\sum_{m=1}^{M}\mathbf{\Sigma}_{k,mm}\big)\mathbf{v}_k}. \quad (28)$$

Since (28) is in the form of generalized Rayleigh quotient, it admits the solution

$$\mathbf{v}_k^\star = \mathbf{u}_{\min}\Bigg[\bigg(\sum_{m=1}^{M}\mathbf{\Sigma}_{k,mm}\bigg)^{-1}\bigg(\sum_{m,n=1}^{M}\eta_{j,mn}(\mathbf{v}_j)\mathbf{\Sigma}_{k,mn}\bigg)\Bigg] \qquad (29)$$

and the optimal covariance shaping vector of UE j for a given $\mathbf{v}_k$ is obtained in a similar way. Hence, problem (P1) is solved by alternating the optimization between $\mathbf{v}_k$ and $\mathbf{v}_j$ until a predetermined convergence criterion is satisfied, e.g., until the difference between the values of the objective in consecutive iterations is sufficiently small. This scheme is formalized in Algorithm 1, whose convergence properties are characterized in Proposition 2.

Proposition 2. *The alternating optimization algorithm described in Algorithm 1 converges to a stationary point of problem* (P1).

   Proof: Observe that the objective function in (28) does not depend on the scaling of $\mathbf{v}_k$ and, if we relax the nonconvex constraint $\|\mathbf{v}_k\|^2 = 1$ as $\|\mathbf{v}_k\|^2 \leq 1$, (29) remains a solution of (28). Then, according to [34, Corollary 2], every limit point of the sequence generated by the alternating optimization algorithm applied to problem (P1) with the relaxed constraints is a stationary point of the original problem. ∎

Algorithm 1 can be implemented in a centralized manner at the BS and the covariance shaping vectors are fed back to the corresponding UEs. Alternatively, it can be implemented
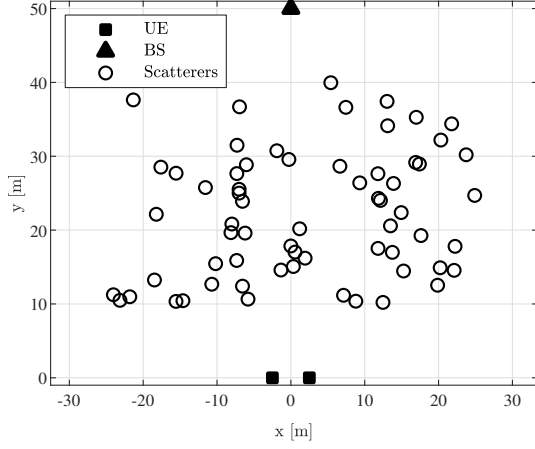
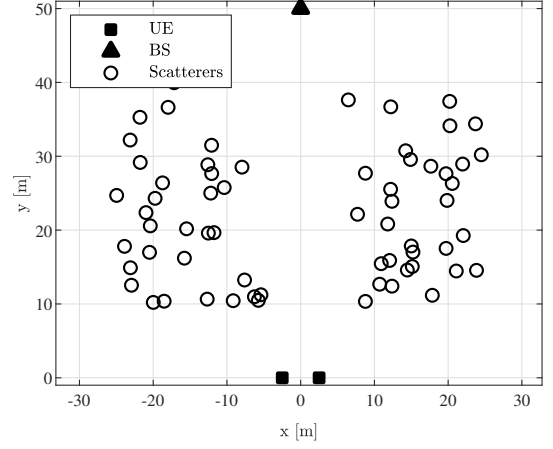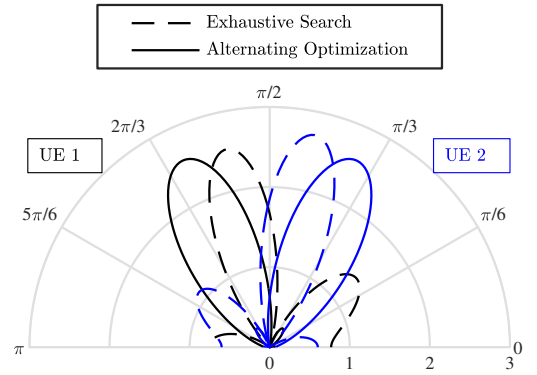(a) NLoS scenario with $K = 2$.
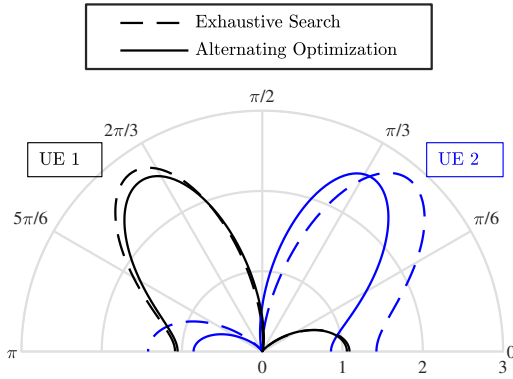
(b) Corresponding covariance shaping vectors.

Fig. 1. Two-UE case: (a) 2D map of the considered NLoS scenario with inter-UE distance $d = 4$ m; (b) corresponding covariance shaping vectors with $N = 2$ obtained with exhaustive search and with Algorithm 1.

(a) LoS scenario with $K = 2$.

(b) Corresponding covariance shaping vectors.

Fig. 2. Two-UE case: (a) 2D map of the considered LoS scenario with inter-UE distance $d = 4$ m; (b) corresponding covariance shaping vectors with $N = 2$ obtained with exhaustive search and with Algorithm 1.

in a distributed fashion at the UEs. In this case, each UE can compute its covariance shaping vector without any information exchange with the other UE provided that the channel statistics of the latter are known, the order of update is fixed, and the same initial points are used.[4] In fact, under these conditions, the entire alternating optimization procedure can be carried out locally and independently at each UE. Note that the same considerations hold for Algorithm 2 presented in the next section for the multi-UE case, where we also discuss the computational complexity.

The result of the covariance shaping optimization heavily depends on the physical scattering environment. Consider the non-line-of-sight (NLoS) scenario in Fig. 1(a), where there is no line-of-sight (LoS) path between the BS and the UEs, the UEs are equipped with uniform linear arrays (ULAs), and the inter-UE distance is $d = 4$ m. In this case, as shown in Fig. 1(b), the covariance shaping vectors tend to focus their power along reflected paths that interfere with each other as little as possible while also carrying sufficient channel power, which results in a nearly interference-free

---

[4]Note that $\mathbf{v}_k^{(0)}$ and $\mathbf{v}_j^{(0)}$ can be any predefined pair of normalized vectors provided by the BS.

transmission/reception. On the other hand, when a LoS path exists, as in the LoS scenario in Fig. 2(a), this generally carries more channel power than any other path. In this case, as shown in Fig. 2(b), the propagation directions selected by the covariance shaping vectors tend to partially capture the LoS path while also focusing some of their power along separated reflected paths in order to achieve some degree of statistical orthogonality between the UEs. Remarkably, in both Fig. 1(b) and Fig. 2(b), the result of Algorithm 1 is very close to the optimal solution of problem (P1) obtained with exhaustive search within a set of 10^3 uniformly distributed vectors over the N-dimensional unit sphere for each UE, where the former is characterized by negligible complexity with respect to the latter.

Kronecker Channel Model. Let us consider the particular case where each channel $\mathbf{H}_k$ follows the Kronecker channel model [28]. In this setting, we have $\mathbf{H}_k = \mathbf{R}_k^{\frac{1}{2}} \mathbf{H}_k^{(\mathrm{w})} \mathbf{T}_k^{\frac{1}{2}}$, with $\mathbf{R}_k$ and $\mathbf{T}_k$ defined in Section II and $\mathrm{vec}(\mathbf{H}_k^{(\mathrm{w})}) \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{NM})$. Accordingly, the channel covariance matrix in (1) can be expressed as $\mathbf{\Sigma}_k = \mathbf{T}_k^{\mathrm{T}} \otimes \mathbf{R}_k$ with block elements given by $\mathbf{\Sigma}_{k,mn} = T_{k,mn}^* \mathbf{R}_k$, where $T_{k,mn}$ denotes the (m, n)th element of $\mathbf{T}_k$. Hence, from (23), we have

$$\delta(\mathbf{v}_k, \mathbf{v}_j)$$

$$= \frac{\mathrm{tr}\big((\mathbf{v}_k^{\mathrm{H}} \mathbf{\Sigma}_{k,mn} \mathbf{v}_k)_{m,n=1}^M (\mathbf{v}_j^{\mathrm{H}} \mathbf{\Sigma}_{j,mn}^{\mathrm{H}} \mathbf{v}_j)_{m,n=1}^M\big)}{\mathrm{tr}\big((\mathbf{v}_k^{\mathrm{H}} \mathbf{\Sigma}_{k,mn} \mathbf{v}_k)_{m,n=1}^M\big) \mathrm{tr}\big((\mathbf{v}_j^{\mathrm{H}} \mathbf{\Sigma}_{j,mn}^{\mathrm{H}} \mathbf{v}_j)_{m,n=1}^M\big)} \quad (30)$$

$$= \frac{\mathrm{tr}\big((T_{k,mn} \mathbf{v}_k^{\mathrm{H}} \mathbf{R}_k \mathbf{v}_k)_{m,n=1}^M (T_{j,mn} \mathbf{v}_j^{\mathrm{H}} \mathbf{R}_j \mathbf{v}_j)_{m,n=1}^M\big)}{\mathrm{tr}\big((T_{k,mn} \mathbf{v}_k^{\mathrm{H}} \mathbf{R}_k \mathbf{v}_k)_{m,n=1}^M\big) \mathrm{tr}\big((T_{j,mn} \mathbf{v}_j^{\mathrm{H}} \mathbf{R}_j \mathbf{v}_j)_{m,n=1}^M\big)}$$
(31)

$$= \frac{\mathrm{tr}(\mathbf{T}_k \mathbf{T}_j)}{\mathrm{tr}(\mathbf{T}_k) \mathrm{tr}(\mathbf{T}_j)}. \quad (32)$$

It is straightforward to observe that, in this case, $\delta(\mathbf{v}_k, \mathbf{v}_j)$ is independent of $\mathbf{v}_k$ and $\mathbf{v}_j$. Hence, under the Kronecker channel model, it is not possible to alter the channel statistics perceived at one end of the communication link by designing the transceiver at the other end: as a consequence, no meaningful effective channel separation can be performed. This is in accordance with the properties of the Kronecker channel model, whereby the transmit and receive covariance matrices are independent. In this context, the signal subspace separation is exclusively determined by the physical scattering environment and can only be achieved when $\mathrm{tr}(\mathbf{T}_k \mathbf{T}_j) = 0$, which is rarely satisfied in practice [27], [28].

B. Multi-UE Case

In the general case, i.e., for $K \geq 2$, the covariance shaping vectors for each UE are computed by solving the optimization problem

$$
\begin{aligned}
& \underset{\{\mathbf{v}_k\}_{k=1}^K}{\text{minimize}} && \sum_{k \neq j} \delta(\mathbf{v}_k, \mathbf{v}_j) \\
& \text{subject to} && \{\|\mathbf{v}_k\|^2 = 1\}_{k=1}^K
\end{aligned}
\quad \text{(P2)}
$$

with $\delta(\mathbf{v}_k, \mathbf{v}_j)$ defined in (23). Although problem (P2) is not convex in any of the optimization variables $\{\mathbf{v}_k\}_{k=1}^K$, a suboptimal solution can be efficiently obtained via block coordinate descent, which can be interpreted as an extension of the alternating optimization approach presented in the previous section for the two-UE case. The optimal covariance shaping vector of UE k for given $\{\mathbf{v}_j\}_{j \neq k}$ is obtained as

$$\mathbf{v}_k^\star = \underset{\mathbf{v}_k : \|\mathbf{v}_k\|^2 = 1}{\mathrm{argmin}} \frac{\sum_{m,n=1}^M \mathbf{v}_k^{\mathrm{H}} \big(\sum_{j \neq k} \eta_{j,mn}(\mathbf{v}_j) \mathbf{\Sigma}_{k,mn} \big) \mathbf{v}_k}{\mathbf{v}_k^{\mathrm{H}} \big(\sum_{m=1}^M \mathbf{\Sigma}_{k,mm} \big) \mathbf{v}_k}$$
(33)

with $\eta_{j,mn}(\mathbf{v}_j)$ defined in (27). Like (28), (33) is in the form of generalized Rayleigh quotient and thus admits the solution

$$\mathbf{v}_k^\star = \mathbf{u}_{\min}\left[\left(\sum_{m=1}^M \mathbf{\Sigma}_{k,mm} \right)^{-1} \left(\sum_{m,n=1}^M \sum_{j \neq k} \eta_{j,mn}(\mathbf{v}_j) \mathbf{\Sigma}_{k,mn} \right) \right]$$
(34)

and the optimal covariance shaping vectors of the other UEs j for given $\{\mathbf{v}_k\}_{k \neq j}$ are obtained in a similar way. Hence, problem (P2) is solved by optimizing the strategy of each UE given the strategies of the other UEs until a predetermined convergence criterion is satisfied. Furthermore, at each iteration i, the update $\mathbf{v}_k^{(i)} = \alpha \mathbf{v}_k^\star + (1 - \alpha) \mathbf{v}_k^{(i-1)}$ can be used to

---

Algorithm 2 (Covariance shaping: Block coordinate descent algorithm)

Data: $\{\mathbf{v}_k^{(0)}\}_{k=1}^K$, $\{\mathbf{\Sigma}_k\}_{k=1}^K$, $\alpha \in (0,1]$, and the accuracy ϵ.
   Fix $n = 1$.
`While`
$$\sum_{k \neq j} \frac{\left| \delta(\mathbf{v}_k^{(n)}, \mathbf{v}_j^{(n)}) - \delta(\mathbf{v}_k^{(n-1)}, \mathbf{v}_j^{(n-1)}) \right|}{\delta(\mathbf{v}_k^{(n)}, \mathbf{v}_j^{(n)})} > \epsilon$$

`For` $k = 1, \ldots, K$
   `S.1:` Given $\{\mathbf{v}_j^{(n-1)}\}_{j \neq k}$, compute
      $\{\eta_{j,mn}(\mathbf{v}_j^{(n)})\}_{m,n=1}^M$ as in (27), $\forall j \neq k$.
   `S.2:` Compute $\mathbf{v}_k^\star$ as in (34).
   `S.3:` Update $\mathbf{v}_k^{(n)} = \alpha \mathbf{v}_k^\star + (1 - \alpha) \mathbf{v}_k^{(n-1)}$.
`End`
`End`
`S.4:` Fix $\{\mathbf{v}_k = \mathbf{v}_k^{(n)}\}_{k=1}^K$.

---

limit the variation of the covariance shaping vectors between consecutive iterations, where the step size $\alpha \in (0,1]$ is chosen to strike the proper balance between convergence speed and accuracy (see, e.g., [10], [35]). The proposed scheme is formalized in Algorithm 2, whose convergence properties are characterized in Proposition 3.

Proposition 3. *The block coordinate descent algorithm described in Algorithm 2 converges to a limit point of problem* (P2).

   Proof: At each iteration i of Algorithm 1, the covariance shaping vector of UE k results from solving (28), which admits the optimal solution in (29). Hence, the sequence $\{\delta(\mathbf{v}_k^{(i)}, \mathbf{v}_j^{(i)})\}_i$ is non-increasing since

$$\delta(\mathbf{v}_k^{(i)}, \mathbf{v}_j^{(i-1)}) \leq \delta(\mathbf{v}_k^{(i-1)}, \mathbf{v}_j^{(i-1)}). \quad (35)$$

Moreover, since $\delta(\mathbf{v}_k, \mathbf{v}_j) \geq 0$, the sequence $\{\delta(\mathbf{v}_k^{(i)}, \mathbf{v}_j^{(i)})\}_i$ converges to a finite non-negative value. ∎

The computational complexity of Algorithm 2 is essentially dictated by step `S.2`, which corresponds to the computation of the optimal covariance shaping vector of UE k for given $\{\mathbf{v}_j\}_{j \neq k}$. This is expressed in (34) and has a complexity $\mathcal{O}(N^{2.37})$ due to the eigenvalue decomposition and inverse matrix operations [25]. Hence, the overall computational complexity of Algorithm 2 is $\mathcal{O}(IKN^{2.37})$, where I is the number of iterations required for convergence. Note that both N and I are fairly modest since the UEs are equipped with a small-to-moderate number of antennas and the algorithm converges in very few iterations (i.e., less than 10 in the simulation scenarios considered in Section V). We can thus conclude that Algorithms 1 and 2 are remarkably efficient in terms of computation complexity.

V. NUMERICAL RESULTS AND DISCUSSION

In this section, we present numerical results to evaluate the gains achieved by the proposed covariance shaping framework.

To this end, we compare the following alternative transmission/reception schemes, where the first is based on covariance shaping and the second is considered as a reference.[5]

(A) **Covariance shaping.** The multiple antennas at the UE are employed to implement covariance shaping and one data stream per UE is transmitted using MMSE precoding. Specifically, each UE k applies its covariance shaping vector $\mathbf{v}_k$, obtained with Algorithm 1 or Algorithm 2 (depending on the value of K), during both the uplink pilot-aided channel estimation phase and the downlink data transmission phase, as discussed in Section IV. The BS obtains the MMSE estimates of the resulting effective channels $\{\bar{\mathbf{g}}_k\}_{k=1}^K$ as in (9) based on the transmission of $P < K$ orthogonal pilot vectors by the UEs. These estimates are used to compute the $(M \times K)$-dimensional MMSE precoding matrix at the BS as in [11].

(B) **Spatial multiplexing.** The multiple antennas at the UE are employed for spatial multiplexing and N data streams per UE are transmitted using MMSE precoding. Specifically, the BS obtains the MMSE estimates of the channel matrices $\{\mathbf{H}_k\}_{k=1}^K$ as in (3) based on the transmission of $P < K$ orthogonal pilot matrices by the UEs. These estimates are used to compute the $(M \times KN)$-dimensional MMSE precoding matrix at the BS as in [11].

Some comments are in order.

- The reference scheme in (B) requires antenna-specific pilots for the estimation of the channel matrices $\{\mathbf{H}_k\}_{k=1}^K$. In particular, N orthogonal pilot vectors are assigned to each UE to avoid any pilot contamination among the antennas of the same UE, which implies $\tau \geq PN$. On the other hand, the estimation of the effective channels $\{\bar{\mathbf{g}}_k\}_{k=1}^K$ resulting from covariance shaping requires $\tau \geq P$ and, therefore, the pilot length can potentially be reduced.

- The reference scheme in (B) allows the transmission of up to N data streams per UE, while only one data stream per UE is transmitted when covariance shaping is used. In the following, we demonstrate how the proposed covariance shaping method can effectively outperform spatial multiplexing in scenarios where the UEs exhibit highly overlapping channel covariance matrices and the spatial selectivity of the BS is not sufficient to separate such UEs.

We evaluate different scenarios in which the BS serves K closely spaced UEs. We assume that the UEs are equipped with ULAs, whereas both ULAs and uniform planar arrays (UPAs) are considered at the BS. In this context, we define the ULA response at each UE for the angle of impingement θ as

$$\mathbf{a}(\theta) \triangleq [1, e^{-2\pi\delta\cos(\theta)}, \ldots, e^{-2\pi\delta(N-1)\cos(\theta)}]^{\mathrm{T}} \in \mathbb{C}^{N\times 1} \tag{36}$$

where $\delta = 0.5$ represents the ratio between the antenna spacing and the signal wavelength. Likewise, we define the

[5]The relevant spatial multiplexing methods in [36], [37] were also tested as reference schemes. Nonetheless, they are not included in this section since they are always outperformed by the reference scheme in (B) in the considered simulation scenarios.

Parameter	Symbol	Value
Number of BS antennas	M	128
Number of UEs	K	$\{2, 4, 8\}$
Number of UE antennas	N	2
Noise variance at the BS	σ_{BS}^2	-80 dBm
Noise variance at the UEs	σ_{UE}^2	-80 dBm
Transmit power at the BS	ρ_{BS}	30 dBm
Transmit power at the UEs	ρ_{UE}	25 dBm
Number of orthogonal pilots	P	$\lceil \frac{K}{2} \rceil$
Pilot length	τ	$\{1, 2, 4\}$
Ricean factor	κ	2.5 (LoS) 0 (NLoS)
Pathloss exponent	β	2
Inter-UE distance	d	$\{2, 4\}$ m

TABLE I. Simulation parameters (unless otherwise stated).

UPA response at the BS for the angles of impingement ϕ along the azimuth direction and ψ along the elevation direction as

$$\mathbf{b}(\phi, \psi)$$
$$\triangleq [1, e^{-2\pi\delta\cos(\phi)}, \ldots, e^{-2\pi\delta(M_x-1)\cos(\phi)}]^{\mathrm{T}}$$
$$\otimes [1, e^{-2\pi\delta\sin(\psi)}, \ldots, e^{-2\pi\delta(M_y-1)\sin(\psi)}]^{\mathrm{T}} \in \mathbb{C}^{M\times 1} \tag{37}$$

where M_x and M_y represent the number of BS antennas along the azimuth and elevation directions, respectively, with $M = M_x M_y$. The ULA response at the BS for the angle of impingement ϕ can be recovered from (37) by setting $M_x = M$, $M_y = 1$, and $\psi = 0$. Then, the instantaneous channel between the BS and each UE k follows the discrete physical channel model in [38] and is given by

$$\mathbf{H}_k = \sqrt{\frac{\kappa}{1+\kappa}} d_k^{-\frac{\beta}{2}} \mathbf{a}(\theta_k) \mathbf{b}(\phi_k, \psi_k)^{\mathrm{H}}$$
$$+ \sqrt{\frac{1}{1+\kappa}} \sum_{u=1}^{U} d_{k,u}^{-\frac{\beta}{2}} \alpha_{k,u} \mathbf{a}(\theta_{k,u}) \mathbf{b}(\phi_{k,u}, \psi_{k,u})^{\mathrm{H}} \tag{38}$$

where κ is the Ricean factor, U is the number of reflected paths, d_k and $d_{k,u}$ are the distances of the LoS path and of the uth reflected path, respectively, θ_k and $\theta_{k,u}$ are the angles of impingement at the UE of the LoS path and of the uth reflected path, respectively, ϕ_k (resp. ψ_k) and $\phi_{k,u}$ (resp. $\psi_{k,u}$) are the angles of impingement at the BS of the LoS path and of the uth reflected path along the azimuth (resp. elevation) direction, respectively, $\alpha_{k,u}$ is the random phase delay of the uth reflected path, and β is the pathloss exponent. Unless otherwise stated, we use the simulation parameters listed in Table I. The following results are obtained by means of Monte Carlo simulations with 5×10^3 independent channel realizations. Lastly, we point out that, in the considered simulation scenarios, Algorithms 1 and 2 are always observed to converge in very few iterations, i.e., less than 10.

A. Two-UE case

Following the discussion in Section IV-A, we first study the simple case of $K = 2$ UEs in order to highlight the key features of the proposed covariance shaping framework. In this setting, we assume $P = 1$, i.e., the UEs are assigned the
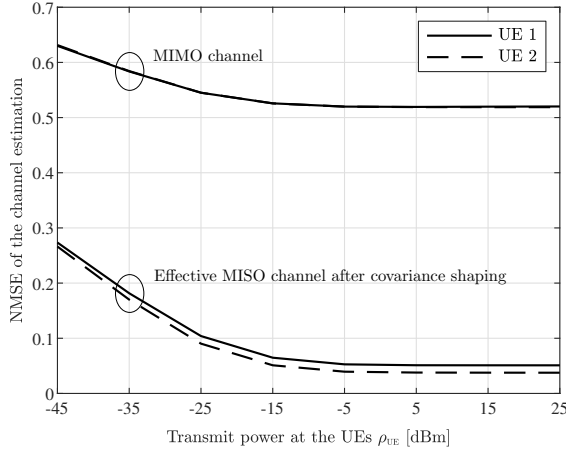
Fig. 3. Two-UE case: NMSE of the channel estimation versus the transmit power at the UEs ρ_{UE} for the NLoS scenario depicted in Fig. 1(a), with ULA at the BS.

Fig. 4. Two-UE case: average sum rate versus the transmit power at the BS ρ_{BS} for the NLoS scenario depicted in Fig. 1(a), with $d = 4$ m, ULA at the BS, and for different values of M.

Fig. 5. Two-UE case: average sum rate versus the transmit power at the BS ρ_{BS} for the LoS scenario depicted in Fig. 2(a), with $M = 128$, ULA at the BS, and for different values of d.

same pilot, and we remark that there is no pilot contamination among the antennas of the same UE for the reference scheme in (B). Assuming that the BS is equipped with a ULA, we consider the LoS and NLoS scenarios illustrated in Fig. 1(a) and in Fig. 2(a), respectively, where the corresponding covariance shaping vectors are shown in Fig. 1(b) and in Fig. 2(b), respectively. Observe that, in both cases, the inter-UE distance is $d = 4$ m: this is quite small compared with the distance of about 50 m between the UEs and the BS, and results in highly overlapping channel covariance matrices.

Let us define the normalized MSE (NMSE) of the channel estimation for UE k as

$$\overline{\text{NMSE}}_k \triangleq \mathbb{E}\left[\frac{\|\hat{\bar{\mathbf{g}}}_k - \bar{\mathbf{g}}_k\|^2}{\|\bar{\mathbf{g}}_k\|^2}\right] \tag{39}$$

for the effective channels $\{\hat{\bar{\mathbf{g}}}_k\}_{k=1}^K$ resulting from covariance shaping and as

$$\text{NMSE}_k \triangleq \frac{1}{N}\sum_{n=1}^N \mathbb{E}\left[\frac{\|\hat{\mathbf{g}}_{k,n} - \mathbf{g}_{k,n}\|^2}{\|\mathbf{g}_{k,n}\|^2}\right] \tag{40}$$

for the channel matrices $\{\hat{\mathbf{H}}_k\}_{k=1}^K$. Considering the NLoS scenario in Fig. 1(a), Fig. 3 plots the NMSE of the channel estimation versus the transmit power at the UEs.[6] Thanks to the improved statistical orthogonality between the UEs, covariance shaping considerably increases the channel estimation accuracy with respect to the case where the original MIMO channels are estimated in the presence of pilot contamination among the antennas of different UEs.

Considering again the NLoS scenario in Fig. 1(a), Fig. 4 plots the average sum rate versus the transmit power at the BS ρ_{BS} for different numbers of BS antennas M. For $M \leq 128$, it is straightforward to see that the signal subspace separation enforced by covariance shaping during both the uplink pilot-aided channel estimation phase and the downlink data transmission phase has a highly beneficial effect on the system
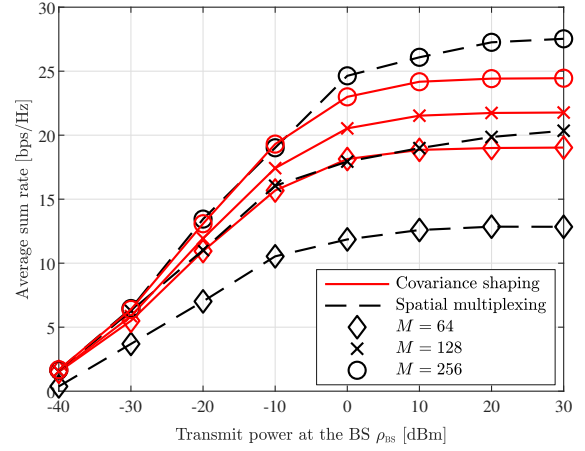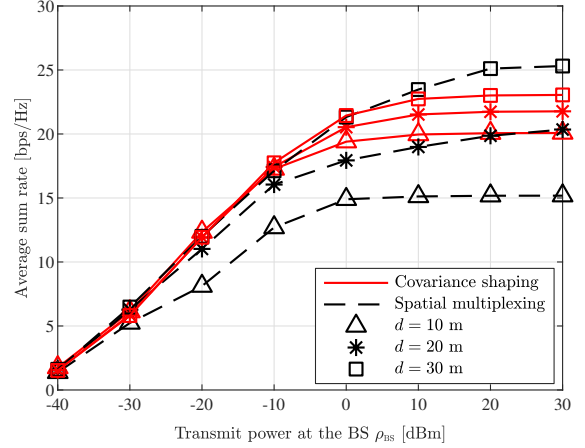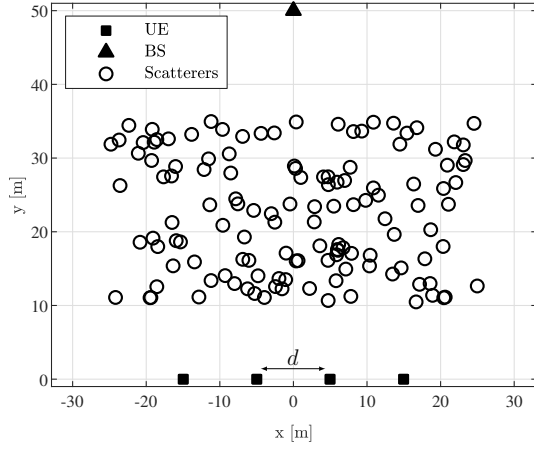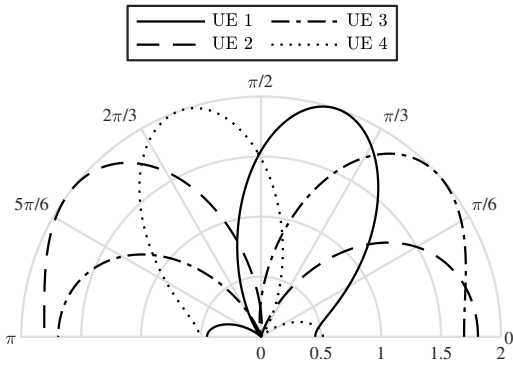
[6]Note that the transmit signal-to-noise ratio (SNR) is given by simply dividing the transmit power at the BS by the noise variance at the BS, which is given in Table I.

performance. However, covariance shaping is outperformed by the reference scheme in (B) for $M = 256$. In this setting, the enhanced spatial selectivity of the BS allows to effectively separate the UEs and promotes the transmission of multiple data streams per UE. Focusing now on the LoS scenario in Fig. 2(a), Fig. 5 plots the average sum rate versus the transmit power at the BS ρ_{BS}, with $M = 128$ and for different values of the inter-UE distance d. Evidently, the reference scheme in (B) is markedly limited by the overlapping channel covariance matrices of the UEs and is able to effectively separate the UEs only for very large inter-UE distances (i.e., $d = 30$ m). On the other hand, covariance shaping is able to enforce statistical orthogonality even when the UEs are closely spaced.

B. Multi-UE Case

Following the discussion in Section IV-B, we now examine the general case of $K \geq 2$. We begin by considering the NLoS scenario in Fig. 6(a) with $K = 4$. Here, the BS is equipped with a ULA and we vary the inter-UE distance d in order to quantify its impact on the degree of statistical orthogonality among the UEs. In this setting, $P = 2$ orthogonal pilots

(a) NLoS scenario with $K = 4$.

(b) Corresponding covariance shaping vectors.

Fig. 6. Multi-UE case: (a) 2D map of the considered NLoS scenario with $K = 4$ and variable inter-UE distance d; (b) corresponding covariance shaping vectors with $N = 2$ obtained with Algorithm 2.

are assigned such that two adjacent UEs always utilize orthogonal pilots. The covariance shaping vectors obtained with Algorithm 2 are shown in Fig. 6(b): these tend to focus their power along reflected paths that are as orthogonal as possible to each other while also carrying sufficient channel power (cf. Fig. 1(b)). Fig. 7 plots the average sum rate versus the inter-UE distance d, with $M = 128$ and $\rho_{\mathrm{BS}} = 30$ dBm. Note that, under the considered pilot assignment, pilot contamination occurs only between UEs at distance $2d$. In this context, we also study the case where the UEs are scheduled into two separate groups of non-adjacent UEs without pilot contamination (scheduling), as opposed to the case where all the UEs are the served simultaneously with pilot contamination (no scheduling). As in Fig. 5, the reference scheme in (B) outperforms covariance shaping only for very large inter-UE distances (i.e., $d \geq 13$ m). In this regard, the scheduling approach is shown to further deteriorate the performance of spatial multiplexing due to the pre-log factor of $1/2$ in the sum rate despite avoiding the pilot contamination.

We conclude this section by considering the NLoS scenario in Fig. 8 with $K = 8$. Here, the BS is equipped with a UPA and is placed at a height of 20 m with respect to the UEs and the scatterers. Moreover, the UEs are separated into two
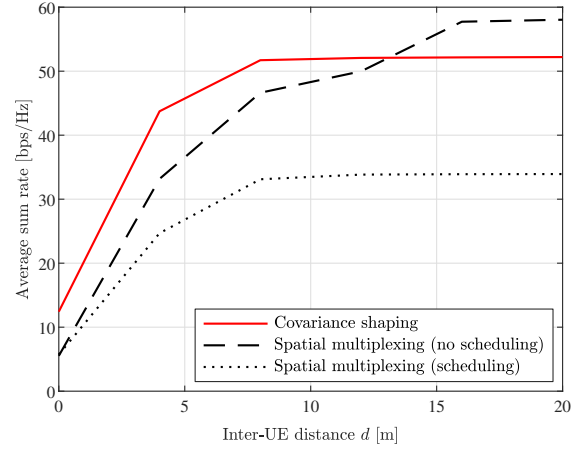

Fig. 7. Multi-UE case: average sum rate versus the inter-UE distance d for the NLoS scenario depicted in Fig. 6(a), with $M = 128$, ULA at the BS, and $\rho_{\mathrm{BS}} = 30$ dBm.

groups with inter-group distance $D = 20$ m and the inter-UE distance is $d = 2$ m, which results in highly overlapping channel covariance matrices within the same group. In this setting, $P = 4$ orthogonal pilots are assigned such that two adjacent UEs always utilize orthogonal pilots and covariance shaping is applied independently to each group. Fig. 9 plots the average sum rate versus the transmit power at the BS ρ_{BS} for different numbers of BS antennas M. Note that $M = 64$ corresponds to $M_x = M_y = 8$, whereas $M = 128$ corresponds to $M_x = 16$ and $M_y = 8$. As in Fig. 7, we also study the case where the UEs in each group are scheduled into two separate subgroups of non-adjacent UEs without pilot contamination among groups (scheduling). The signal subspace separation enforced by covariance shaping during both the uplink pilot-aided channel estimation phase and the downlink data transmission phase becomes even more beneficial when the BS is equipped with a UPA with limited spatial selectivity in the azimuth direction. Again, the scheduling approach is shown to further deteriorate the performance of spatial multiplexing due to the pre-log factor of $1/2$ in the sum rate despite avoiding the pilot contamination within the same group.

VI. CONCLUSIONS

In this paper, we introduce the novel concept of MIMO covariance shaping as a means to achieve statistical orthogonality among interfering UEs. It consists in preemptively applying a statistical beamforming at each UE during both the uplink pilot-aided channel estimation phase and the downlink data transmission phase, aiming at enforcing a separation of the signal subspaces of the UEs that would be otherwise highly overlapping. The proposed MIMO covariance shaping framework exploits the realistic non-Kronecker structure of massive MIMO channels, which allows to suitably alter the channel statistics perceived at the BS by designing the transceiver at the UE-side. To compute the covariance shaping strategies, we present a low-complexity block coordinate descent algorithm that minimizes the inter-UE interference (as a metric to measure the spatial correlation), which is proved to converge to a limit point of the original nonconvex problem or to a
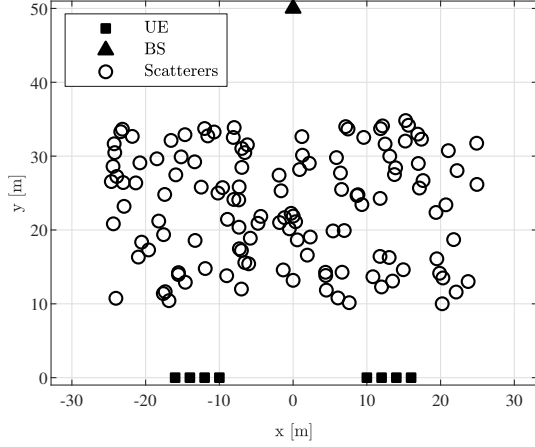
Fig. 8. Multi-UE case: (a) 2D map of the considered NLoS scenario with $K = 8$, inter-UE distance $d = 2$ m, and inter-group distance $D = 20$ m.
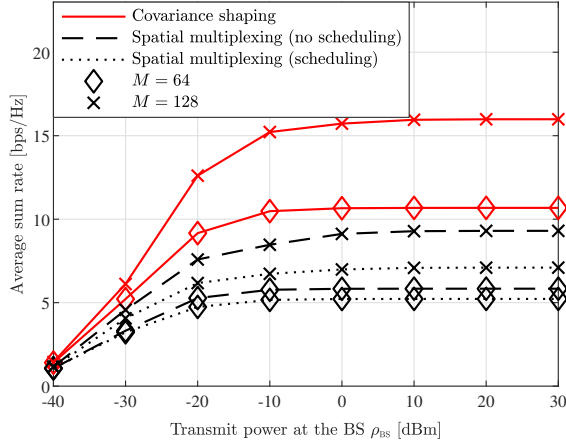

Fig. 9. Multi-UE case: average sum rate versus the transmit power at the BS ρ_{BS} for the NLoS scenario depicted in Fig. 8, with UPA at the BS and for different values of M.

stationary point in the two-UE case. We provide numerical results characterizing several scenarios where MIMO covariance shaping outperforms a reference schemes employing the multiple antennas at the UE for spatial multiplexing. Specifically, this occurs when the spatial selectivity of the BS is not sufficient to separate UEs exhibiting high spatial correlation, (e.g., when the BS is equipped with a UPA with limited spatial selectivity in the azimuth direction) and the channel estimation is limited by strong pilot contamination. Future work will consider MIMO covariance shaping in combination with multi-stream transmission and will analyze the trade-off between statistical orthogonality and spatial multiplexing.

APPENDIX A
DERIVATIONS OF THE EFFECTIVE SINR IN (16)

In this Appendix, we derive the expectation terms in (15) to obtain the expression of the effective SINR in (16). First of all, given the multi-UE effective channel matrix $\bar{\mathbf{H}} = [\bar{\mathbf{g}}_1^{\text{T}}, \dots, \bar{\mathbf{g}}_K^{\text{T}}]^{\text{T}}$, where the effective channels $\{\bar{\mathbf{g}}_k\}_{k=1}^K$ are

independent and $\bar{\mathbf{g}}_k \sim \mathcal{CN}(\mathbf{0}, \bar{\bar{\mathbf{\Phi}}}_k)$, it is easy to verify that

$$\mathbb{E}[\|\bar{\mathbf{H}}\|_{\text{F}}^2] = \sum_{k=1}^{K} \text{tr}(\bar{\bar{\mathbf{\Phi}}}_k). \tag{41}$$

Furthermore, recalling the definition of $\hat{\bar{\mathbf{g}}}_k$ in (9), we have

$$\mathbb{E}[\bar{\mathbf{g}}_k \hat{\bar{\mathbf{g}}}_k^{\text{H}}] = \text{tr}(\bar{\bar{\mathbf{\Phi}}}_k \mathbf{Q}_k^{-1} \bar{\bar{\mathbf{\Phi}}}_k). \tag{42}$$

We now focus on deriving $\mathbb{V}[\bar{\mathbf{g}}_k \hat{\bar{\mathbf{g}}}_k^{\text{H}}]$, with $k \in \mathcal{S}_p$. We have

$$\mathbb{V}[\bar{\mathbf{g}}_k \hat{\bar{\mathbf{g}}}_k^{\text{H}}]$$
$$= \mathbb{E}\left[|\bar{\mathbf{g}}_k \hat{\bar{\mathbf{g}}}_k^{\text{H}}|^2\right] - \left|\mathbb{E}[\bar{\mathbf{g}}_k \hat{\bar{\mathbf{g}}}_k^{\text{H}}]\right|^2 \tag{43}$$
$$= \mathbb{E}\left[\bar{\mathbf{g}}_k \bar{\bar{\mathbf{\Phi}}}_k \mathbf{Q}_k^{-1}\left(\bar{\mathbf{g}}_k^{\text{H}} + \sum_{j \in \mathcal{S}_p \setminus \{k\}} \bar{\mathbf{g}}_j^{\text{H}} + \frac{1}{\tau \sqrt{\rho_{\text{UE}}}} \mathbf{Z} \mathbf{p}_p^{\text{H}}\right)\right.$$
$$\left. \times \left(\bar{\mathbf{g}}_k + \sum_{j \in \mathcal{S}_p \setminus \{k\}} \bar{\mathbf{g}}_j + \frac{1}{\tau \sqrt{\rho_{\text{UE}}}} \mathbf{p}_p \mathbf{Z}^{\text{H}}\right) \mathbf{Q}_k^{-1} \bar{\bar{\mathbf{\Phi}}}_k \bar{\mathbf{g}}_k^{\text{H}}\right]$$
$$- \text{tr}(\bar{\bar{\mathbf{\Phi}}}_k \mathbf{Q}_k^{-1} \bar{\bar{\mathbf{\Phi}}}_k)^2 \tag{44}$$
$$= \mathbb{E}[\bar{\mathbf{g}}_k \bar{\bar{\mathbf{\Phi}}}_k \mathbf{Q}_k^{-1} \bar{\mathbf{g}}_k^{\text{H}} \bar{\mathbf{g}}_k \mathbf{Q}_k^{-1} \bar{\bar{\mathbf{\Phi}}}_k \bar{\mathbf{g}}_k^{\text{H}}]$$
$$+ \sum_{j \in \mathcal{S}_p \setminus \{k\}} \mathbb{E}[\bar{\mathbf{g}}_k \bar{\bar{\mathbf{\Phi}}}_k \mathbf{Q}_k^{-1} \bar{\mathbf{g}}_j^{\text{H}} \bar{\mathbf{g}}_j \mathbf{Q}_k^{-1} \bar{\bar{\mathbf{\Phi}}}_k \bar{\mathbf{g}}_k^{\text{H}}]$$
$$+ \frac{1}{\tau^2 \rho_{\text{UE}}} \mathbb{E}[\bar{\mathbf{g}}_k \bar{\bar{\mathbf{\Phi}}}_k \mathbf{Q}_k^{-1} \mathbf{Z} \mathbf{p}_p^{\text{H}} \mathbf{p}_p \mathbf{Z}^{\text{H}} \mathbf{Q}_k^{-1} \bar{\bar{\mathbf{\Phi}}}_k \bar{\mathbf{g}}_k^{\text{H}}]$$
$$- \text{tr}(\bar{\bar{\mathbf{\Phi}}}_k \mathbf{Q}_k^{-1} \bar{\bar{\mathbf{\Phi}}}_k)^2 \tag{45}$$

where in (44) we have used the expression in (42) and where (45) follows from the independence between $\bar{\mathbf{g}}_k$ and $\bar{\mathbf{g}}_j$, $\forall k \neq j$, and between $\bar{\mathbf{g}}_k$ and $\mathbf{Z}$. Then, we write $\bar{\mathbf{g}}_k^{\text{H}} = \bar{\bar{\mathbf{\Phi}}}_k^{\frac{1}{2}} \mathbf{x}_k$, with $\mathbf{x}_k \sim \mathcal{CN}(0, \mathbf{I}_M)$, and obtain

$$\mathbb{V}[\bar{\mathbf{g}}_k \hat{\bar{\mathbf{g}}}_k^{\text{H}}]$$
$$= \mathbb{E}\left[\mathbf{x}_k^{\text{H}} \bar{\bar{\mathbf{\Phi}}}_k^{\frac{1}{2}} \bar{\bar{\mathbf{\Phi}}}_k \mathbf{Q}_k^{-1} \bar{\bar{\mathbf{\Phi}}}_k^{\frac{1}{2}} \mathbf{x}_k \mathbf{x}_k^{\text{H}} \bar{\bar{\mathbf{\Phi}}}_k^{\frac{1}{2}} \mathbf{Q}_k^{-1} \bar{\bar{\mathbf{\Phi}}}_k \bar{\bar{\mathbf{\Phi}}}_k^{\frac{1}{2}} \mathbf{x}\right]$$
$$+ \sum_{j \in \mathcal{S}_p \setminus \{k\}} \text{tr}(\bar{\bar{\mathbf{\Phi}}}_k^2 \mathbf{Q}_k^{-1} \bar{\bar{\mathbf{\Phi}}}_j \mathbf{Q}_k^{-1} \bar{\bar{\mathbf{\Phi}}}_k)$$
$$+ \frac{1}{\tau \varrho_{\text{UE}}} \text{tr}(\bar{\bar{\mathbf{\Phi}}}_k^2 \mathbf{Q}_k^{-2} \bar{\bar{\mathbf{\Phi}}}_k) - \text{tr}(\bar{\bar{\mathbf{\Phi}}}_k \mathbf{Q}_k^{-1} \bar{\bar{\mathbf{\Phi}}}_k)^2 \tag{46}$$
$$= \text{tr}(\bar{\bar{\mathbf{\Phi}}}_k^2 \mathbf{Q}_k^{-1} \bar{\bar{\mathbf{\Phi}}}_k \mathbf{Q}_k^{-1} \bar{\bar{\mathbf{\Phi}}}_k) + \text{tr}(\bar{\bar{\mathbf{\Phi}}}_k \mathbf{Q}_k^{-1} \bar{\bar{\mathbf{\Phi}}}_k)^2$$
$$+ \sum_{j \in \mathcal{S}_p \setminus \{k\}} \text{tr}(\bar{\bar{\mathbf{\Phi}}}_k^2 \mathbf{Q}_k^{-1} \bar{\bar{\mathbf{\Phi}}}_j \mathbf{Q}_k^{-1} \bar{\bar{\mathbf{\Phi}}}_k)$$
$$+ \frac{1}{\tau \varrho_{\text{UE}}} \text{tr}(\bar{\bar{\mathbf{\Phi}}}_k^2 \mathbf{Q}_k^{-2} \bar{\bar{\mathbf{\Phi}}}_k) - \text{tr}(\bar{\bar{\mathbf{\Phi}}}_k \mathbf{Q}_k^{-1} \bar{\bar{\mathbf{\Phi}}}_k)^2 \tag{47}$$
$$= \text{tr}\left(\bar{\bar{\mathbf{\Phi}}}_k^2 \mathbf{Q}_k^{-1}\left(\bar{\bar{\mathbf{\Phi}}}_k + \sum_{j \in \mathcal{S}_p \setminus \{k\}} \bar{\bar{\mathbf{\Phi}}}_j + \frac{1}{\tau \varrho_{\text{UE}}} \mathbf{I}_M\right) \mathbf{Q}_k^{-1} \bar{\bar{\mathbf{\Phi}}}_k\right)$$
$$\tag{48}$$
$$= \text{tr}(\bar{\bar{\mathbf{\Phi}}}_k^2 \mathbf{Q}_k^{-1} \bar{\bar{\mathbf{\Phi}}}_k) \tag{49}$$

where (46) is obtained by substituting $\mathbb{E}[\mathbf{Z} \mathbf{p}_p^{\text{H}} \mathbf{p}_p \mathbf{Z}^{\text{H}}] = \tau \sigma_{\text{BS}}^2 \mathbf{I}_M$, in (47) we have exploited the fact that $\mathbb{E}[\mathbf{x}_k \mathbf{x}_k^{\text{H}} \mathbf{A} \mathbf{x}_k \mathbf{x}_k^{\text{H}}] = \mathbf{A} + \text{tr}(\mathbf{A}) \mathbf{I}_M$ for a given Hermitian matrix $\mathbf{A} \in \mathbb{C}^{M \times M}$ [13, App. A.2], and in (49) we have used the definition of $\mathbf{Q}_k$ provided in Section III-A. Finally, we focus on deriving $\mathbb{E}\left[|\bar{\mathbf{g}}_k \hat{\bar{\mathbf{g}}}_j^{\text{H}}|^2\right]$, with $k \in \mathcal{S}_p$ and with $j \in \mathcal{S}_q$.

Following similar steps as above, we have

$$\mathbb{E}\big[|\bar{\mathbf{g}}_k \hat{\bar{\mathbf{g}}}_j^{\mathrm{H}}|^2\big]$$

$$= \mathbb{E}\bigg[\bar{\mathbf{g}}_k \bar{\boldsymbol{\Phi}}_j \mathbf{Q}_j^{-1}\bigg(\bar{\mathbf{g}}_j^{\mathrm{H}} + \sum_{l \in \mathcal{S}_q \setminus \{j\}} \bar{\mathbf{g}}_l^{\mathrm{H}} + \frac{1}{\tau\sqrt{\rho_{\mathrm{UE}}}}\mathbf{Z}\mathbf{p}_q^{\mathrm{H}}\bigg)$$

$$\times \bigg(\bar{\mathbf{g}}_j + \sum_{l \in \mathcal{S}_q \setminus \{j\}} \bar{\mathbf{g}}_l + \frac{1}{\tau\sqrt{\rho_{\mathrm{UE}}}}\mathbf{p}_q\mathbf{Z}^{\mathrm{H}}\bigg)\mathbf{Q}_j^{-1}\bar{\boldsymbol{\Phi}}_j\bar{\mathbf{g}}_k^{\mathrm{H}}\bigg] \quad (50)$$

$$= \mathbb{E}[\bar{\mathbf{g}}_k \bar{\boldsymbol{\Phi}}_j \mathbf{Q}_j^{-1}\bar{\mathbf{g}}_j^{\mathrm{H}}\bar{\mathbf{g}}_j\mathbf{Q}_j^{-1}\bar{\boldsymbol{\Phi}}_j\bar{\mathbf{g}}_k^{\mathrm{H}}]$$

$$+ \sum_{l \in \mathcal{S}_q \setminus \{j\}} \mathbb{E}[\bar{\mathbf{g}}_k \bar{\boldsymbol{\Phi}}_j \mathbf{Q}_j^{-1}\bar{\mathbf{g}}_l^{\mathrm{H}}\bar{\mathbf{g}}_l\mathbf{Q}_j^{-1}\bar{\boldsymbol{\Phi}}_j\bar{\mathbf{g}}_k^{\mathrm{H}}]$$

$$+ \frac{1}{\tau^2\rho_{\mathrm{UE}}}\mathbb{E}[\bar{\mathbf{g}}_k \bar{\boldsymbol{\Phi}}_k \mathbf{Q}_k^{-1}\mathbf{Z}\mathbf{p}_q^{\mathrm{H}}\mathbf{p}_q\mathbf{Z}^{\mathrm{H}}\mathbf{Q}_k^{-1}\bar{\boldsymbol{\Phi}}_k\bar{\mathbf{g}}_k^{\mathrm{H}}] \quad (51)$$

$$= \mathrm{tr}(\bar{\boldsymbol{\Phi}}_k \bar{\boldsymbol{\Phi}}_j \mathbf{Q}_j^{-1}\bar{\boldsymbol{\Phi}}_j\mathbf{Q}_j^{-1}\bar{\boldsymbol{\Phi}}_j)$$

$$+ \sum_{l \in \mathcal{S}_q \setminus \{j\}} \mathrm{tr}(\bar{\boldsymbol{\Phi}}_k \bar{\boldsymbol{\Phi}}_j \mathbf{Q}_j^{-1}\bar{\boldsymbol{\Phi}}_l\mathbf{Q}_j^{-1}\bar{\boldsymbol{\Phi}}_j)$$

$$+ \mathbb{1}_{q=p}\mathrm{tr}(\bar{\boldsymbol{\Phi}}_k \mathbf{Q}_j^{-1}\bar{\boldsymbol{\Phi}}_j)^2 + \frac{1}{\tau\varrho_{\mathrm{UE}}}\mathrm{tr}(\bar{\boldsymbol{\Phi}}_k \bar{\boldsymbol{\Phi}}_j\mathbf{Q}_j^{-2}\bar{\boldsymbol{\Phi}}_j) \quad (52)$$

$$= \mathrm{tr}\bigg(\bar{\boldsymbol{\Phi}}_k \bar{\boldsymbol{\Phi}}_j \mathbf{Q}_j^{-1}\bigg(\bar{\boldsymbol{\Phi}}_j + \sum_{l \in \mathcal{S}_q \setminus \{j\}} \bar{\boldsymbol{\Phi}}_l + \frac{1}{\tau\varrho_{\mathrm{UE}}}\mathbf{I}_M\bigg)\mathbf{Q}_j^{-1}\bar{\boldsymbol{\Phi}}_j\bigg)$$

$$+ \mathbb{1}_{q=p}\mathrm{tr}(\bar{\boldsymbol{\Phi}}_k \mathbf{Q}_j^{-1}\bar{\boldsymbol{\Phi}}_j)^2 \quad (53)$$

$$= \mathrm{tr}(\bar{\boldsymbol{\Phi}}_k \bar{\boldsymbol{\Phi}}_j\mathbf{Q}_j^{-1}\bar{\boldsymbol{\Phi}}_j) + \mathbb{1}_{q=p}\mathrm{tr}(\bar{\boldsymbol{\Phi}}_k \mathbf{Q}_j^{-1}\bar{\boldsymbol{\Phi}}_j)^2. \quad (54)$$

Note that the second term in (54) is active only if UEs k and j are assigned the same pilot.

REFERENCES

[1] P. Mursia, I. Atzeni, D. Gesbert, and L. Cottatellucci, "Covariance shaping for massive MIMO systems," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Abu Dhabi, UAE, Dec. 2018.

[2] ——, "On the performance of covariance shaping in massive MIMO systems," in *Proc. IEEE Int. Workshop Computer-Aided Modeling and Design of Commun. Links and Netw. (CAMAD)*, Barcelona, Spain, Sep. 2018.

[3] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 74–80, Feb. 2014.

[4] E. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.

[5] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, Jun. 2014.

[6] N. Rajatheva, I. Atzeni, E. Björnson *et al.*, "White paper on broadband connectivity in 6G," Jun. 2020. [Online]. Available: http://jultika.oulu.fi/files/isbn9789526226798.pdf

[7] J. Zhang, E. Björnson, M. Matthaiou, D. Wing, K. Ng, H. Yang, and D. J. Love, "Prospective multiple antenna technologies for beyond 5G," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 8, pp. 1637–1660, Aug. 2020.

[8] X. Chen, D. W. K. Ng, W. Yu, E. G. Larsson, N. Al-Dhahir, and R. Schober, "Massive access for 5G and beyond," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 3, pp. 615–637, Mar. 2021.

[9] 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, Mar. 2017.

[10] I. Atzeni, B. Gouda, and A. Tölli, "Distributed precoding design via over-the-air signaling for cell-free massive MIMO," *IEEE Trans. Wireless Commun.*, vol. 20, no. 2, pp. 1201–1216, Feb. 2021.

[11] E. Björnson, J. Hoydis, and L. Sanguinetti, "Massive MIMO has unlimited capacity," *IEEE Trans. Wireless Commun.*, vol. 17, no. 1, pp. 574–590, Jan. 2018.

[12] J. Jose, A. Ashikhmin, T. L. Marzetta, and S. Vishwanath, "Pilot contamination and precoding in multi-cell TDD systems," *IEEE Trans. Wireless Commun.*, vol. 10, no. 8, pp. 2640–2651, Aug. 2011.

[13] T. L. Marzetta, E. G. Larsson, and H. Yang, *Fundamentals of Massive MIMO*. Cambridge University Press, 2016.

[14] J. Nam, J. Y. Ahn, A. Adhikary, and G. Caire, "Joint spatial division and multiplexing: Realizing massive MIMO gains with limited channel state information," in *Proc. Conf. on Inf. Sci. and Sys. (CISS)*, Princeton, USA, Mar. 2012.

[15] A. Adhikary, J. Nam, J. Y. Ahn, and G. Caire, "Joint spatial division and multiplexing—The large-scale array regime," *IEEE Trans. Inf. Theory*, vol. 59, no. 10, pp. 6441–6463, Jun. 2013.

[16] J. Nam, A. Adhikary, J. Y. Ahn, and G. Caire, "Joint spatial division and multiplexing: Opportunistic beamforming, user grouping and simplified downlink scheduling," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 876–890, Oct. 2014.

[17] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, "A coordinated approach to channel estimation in large-scale multiple-antenna systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 264–273, Feb. 2013.

[18] H. Yin, D. Gesbert, and L. Cottatellucci, "Dealing with interference in distributed large-scale MIMO systems: A statistical approach," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 942–953, Oct. 2014.

[19] X. Zhang, D. P. Palomar, and B. Ottersten, "Statistically robust design of linear MIMO transceivers," *IEEE Trans. Signal Process.*, vol. 56, no. 8, pp. 3678–3689, Aug. 2008.

[20] M. Dai and B. Clerckx, "Transmit beamforming for MISO broadcast channels with statistical and delayed CSIT," *IEEE Trans. Commun.*, vol. 63, no. 4, pp. 1202–1215, Jan. 2015.

[21] A. Liu and V. Lau, "Hierarchical interference mitigation for massive MIMO cellular networks," *IEEE Trans. Signal Process.*, vol. 62, no. 18, pp. 4786–4797, Sep. 2014.

[22] A. Padmanabhan, A. Tölli, and I. Atzeni, "Distributed two-stage multi-cell precoding," in *Proc. IEEE Int. Workshop Signal Process. Adv. in Wireless Commun. (SPAWC)*, Cannes, France, Jul. 2019.

[23] L. You, X. Gao, X. Xia, N. Ma, and Y. Peng, "Pilot reuse for massive MIMO transmission over spatially correlated Rayleigh fading channels," *IEEE Trans. Wireless Commun.*, vol. 14, no. 6, pp. 3352–3366, Jun. 2015.

[24] N. N. Moghadam, H. Shokri-Ghadikolaei, G. Fodor, M. Bengtsson, and C. Fischione, "Pilot precoding and combining in multiuser MIMO networks," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 7, pp. 1632–1648, Jul. 2017.

[25] H. Yin, L. Cottatellucci, D. Gesbert, R. R. Müller, and G. He, "Robust pilot decontamination based on joint angle and power domain discrimination," *IEEE Trans. Signal Process.*, vol. 64, no. 11, pp. 2990–3003, Feb. 2016.

[26] R. Müller, L. Cottatellucci, and M. Vehkaperä, "Blind pilot decontamination," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 773–786, Oct. 2014.

[27] V. Raghavan, J. H. Kotecha, and A. M. Sayeed, "Why does the Kronecker model result in misleading capacity estimates?" *IEEE Trans. Inf. Theory*, vol. 56, no. 10, pp. 4843–4864, Oct. 2010.

[28] W. Weichselberger, M. Herdin, H. Özcelik, and E. Bonek, "A stochastic MIMO channel model with joint correlation of both link ends," *IEEE Trans. Wireless Commun.*, vol. 5, no. 1, pp. 90–100, Jan. 2006.

[29] Y. C. Eldar and A. V. Oppenheim, "Covariance shaping least-squares estimation," *IEEE Trans. Signal Process.*, vol. 51, no. 3, pp. 686–697, Mar. 2003.

[30] A. K. Hassan, M. Moinuddin, U. M. Al-Saggaf, O. Aldayel, T. N. Davidson, and T. Y. Al-Naffouri, "Performance analysis and joint statistical beamformer design for multi-user MIMO systems," *IEEE Commun. Lett.*, vol. 24, no. 10, pp. 2152–2156, Oct. 2020.

[31] A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications*. Cambridge University Press, 2006.

[32] E. Björnson, J. Hoydis, and L. Sanguinetti, "Massive MIMO networks: Spectral, energy, and hardware efficiency," *Foundations and Trends® Signal Process.*, vol. 11, no. 3–4, pp. 154–655, Nov. 2017.

[33] E. Dahlman, S. Parkvall, and J. Skold, *5G NR: The Next Generation Wireless Access Technology*. Academic Press, 2018.

[34] L. Grippo and M. Sciandrone, "On the convergence of the block nonlinear Gauss-Seidel method under convex constraints," *Operation Res. Lett.*, vol. 26, no. 3, pp. 127–136, Apr. 2000.

[35] G. Scutari, F. Facchinei, P. Song, D. P. Palomar, and J.-S. Pang, "Decomposition by partial linearization: Parallel optimization of multi-agent systems," *IEEE Trans. Signal Process.*, vol. 62, no. 3, pp. 641–656, Feb. 2014.

[36] Q. H. Spencer, A. L. Swindlehurst, and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels," *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 461–471, Feb. 2004.

[37] S. Christensen, R. Agarwal, E. Carvalho, and J. Cioffi, "Weighted sum-rate maximization using weighted MMSE for MIMO-BC beamforming design," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 4792–4799, Dec. 2008.

[38] A. M. Sayeed, "Deconstructing multiantenna fading channels," *IEEE Trans. Signal Process.*, vol. 50, no. 10, pp. 2563–2579, Oct. 2002.

Laura Cottatellucci (Member, IEEE) received the Master degree from La Sapienza University, Rome, Italy, the Ph.D. degree from the Technical University of Vienna, Austria, in 2006, and the Habilitation degree from the University of Nice-Sophia Antipolis, France. Since December 2017, she is Professor for Digital Communications at the Institute of Digital Communications of Friedrich Alexander Universität (FAU) of Erlangen-Nürnberg (Germany) and Adjunct Professor in EURECOM, France, since September 2021. She worked in Telecom Italia (1995-2000) as responsible of industrial projects and as a Senior Research in Forschungszentrum Telekommunikation Wien Austria (Apr. 2000-Sep. 2005). She was a Research Fellow in INRIA, France, (Oct.-Dec. 2005) and at the University of South Australia (2006). She was also Assistant Professor (Dec. 2006-Nov. 2017) and Adjunct Professor (Mar. 2018-Aug. 2019) in EURECOM, France. Her research interests lie in the field of communications theory and signal processing for wireless communications, satellite and complex networks. She served as associate editor of the IEEE Transactions on Communications and the IEEE Transactions on Signal Processing (Feb 2016-2020). She is an elected member of the IEEE Technical Committee on Signal Processing for Communications and Networking since 2017.

Placido Mursia (Member, IEEE) received his B.Sc. degree in "Telecommunication Engineering" from Politecnico of Turin in 2015. He received the M.Sc. degree with honors in 2018 in "Communications and Computer Networks Engineering" and "Mobile Communications" in the context of a Double Degree program between Politecnico of Turin and EURECOM (Telecom ParisTech), respectively. He obtained his Ph.D from Sorbonne Université of Paris, at the Communication systems department of EURECOM in 2021. He is currently a research scientist in the 5GN group at NEC Laboratories Europe GmbH. His research interests lie in convex optimization, signal processing and wireless communication. He received the Best Paper Award in the Wireless Communications Symposium at IEEE ICC 2019.

David Gesbert (Fellow, IEEE) is currently serving as Director of EURECOM, Sophia Antipolis, France (www.eurecom.fr). He received the Ph.D. degree from TelecomParis, France, in 1997. From 1997 to 1999, he was with the Information Systems Laboratory, Stanford University. He was the Founding Engineer of Iospan Wireless Inc., a Stanford spin off pioneering MIMO-OFDM (currently Intel). Before joining EURECOM in 2004, he was with the Department of Informatics, University of Oslo, as an Adjunct Professor. He has published about 350 articles and 25 patents, 7 of them winning IEEE Best paper awards. He has been the Technical Program Co-Chair for ICC2017 and has been named a Thomson-Reuters Highly Cited Researchers in computer science. He is a Board Member for the OpenAirInterface (OAI) Software Alliance. In 2015, he has been awarded an ERC Advanced Grant. In 2020, he was awarded funding by the French Interdisciplinary Institute on Artificial Intelligence for a Chair in the area of AI for the future IoT. In 2021, he received the Grand Prix in Research from IMT-French Academy of Sciences.

Italo Atzeni (Member, IEEE) received the MSc degree (Hons.) in telecommunications engineering from the University of Cagliari, Italy, in 2010 and the PhD degree (Hons.) in signal theory and communications from the Polytechnic University of Catalonia–BarcelonaTech, Spain, in 2014. Since 2019, he is with the Centre for Wireless Communications, University of Oulu, Finland, where he is currently an Assistant Professor. 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 previously held a research appointment at The Hong Kong University of Science and Technology, Hong Kong, in 2013. 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 Editor for the IEEE Communications Letters.