

Joint UL and DL Spectral Efficiency Optimization of Superimposed Pilots in Massive MIMO

Daniel Verenzuela*, Emil Björnson*, Luca Sanguinetti^{†‡}.

*Department of Electrical Engineering (ISY), Linköping University, Linköping, Sweden.

[†]Dipartimento di Ingegneria dell'Informazione, University of Pisa, Pisa, Italy.

[‡]Large Networks and System Group (LANEAS), CentraleSupélec, Université Paris-Saclay, Gif-sur-Yvette, France.

Abstract—The reuse of pilot sequences in a Massive MIMO system leads to pilot contamination, which reduces the channel estimation quality and adds coherent interference in the data transmission. A standard method to reduce pilot contamination, known as regular pilots (RPs), is to increase the pilot overhead and reuse pilots more sparsely in the network. Another approach, denoted as superimposed pilots (SPs), is to send a superposition of pilot and data symbols which allows the system to reuse pilots far more sparsely. This work performs a comparative analysis of RPs and SPs in Massive MIMO considering the joint spectral efficiency (SE) of the uplink (UL) and downlink (DL) communications. A rigorous DL lower bound on the capacity with SPs is derived and multiobjective optimization theory is used to compare the UL and DL SE between RPs and SPs. Numerical results indicate that RPs and SPs give comparable SE when both methods are optimized.

I. INTRODUCTION

The increasing demand for data traffic is the driving force for the development of wireless communication technologies [1]. Massive MIMO makes use of a large number of antennas at base stations (BSs) to simultaneously serve several user equipments (UEs) by means of spatial multiplexing. This technology improves the spectral efficiency (SE) of the system by orders of magnitude compared to single antenna communications and is considered as a viable option for next generation wireless networks [2].

Obtaining accurate channel state information (CSI) in multiuser MIMO systems is highly important. The time division duplexing (TDD) transmission protocol has been proposed for Massive MIMO where the UEs send uplink (UL) pilot signals that are used by the BSs to estimate the channels. These estimates are then employed to detect UL data signals and to precode (by exploiting channel reciprocity) data signals to UEs in the downlink (DL) [3].

The number of available orthogonal pilot sequences is limited by the size of the coherence block. Therefore, pilots need to be reused among cells in practical networks. This causes interference in the channel estimation process that, in turn, adds coherent interference in the UL and DL data transmissions, giving rise to the so-called pilot contamination effect [4]–[9]. Different approaches have been proposed to mitigate the effect of pilot contamination, such as methods based on

semi-blind channel estimation [4]–[6] or the exploitation of the spatial channel correlations [7]–[9]. A standard method is to use more symbols for pilots than served UEs per cell, which allows the network to reuse pilots less frequently, thereby reducing the effect of pilot contamination at the expense of increasing the pilot overhead [10]–[12]. This method is known in the Massive MIMO literature as regular pilots (RPs).

Another approach to perform channel estimation is to send pilot and data symbols simultaneously. This is referred to as supersimposed pilots (SPs) and has been studied in estimation and detection theory for a couple of decades [13]–[19]. Recently, in [17]–[19] the use of SPs have been proposed as an alternative method to reduce pilot contamination in Massive MIMO. The main idea is for UEs to simultaneously send pilot and data signals in the UL allowing the use of longer pilot sequences that are thus reused less frequently in the network. In doing so, interference in the channel estimation process is reduced. However, this method adds interference from UL data signals into the channel estimates. This causes strong correlation with the received signal in the UL, which results into extra coherent and non-coherent interference that limits the performance of the UL detection [20]. In the DL, the effect of SPs are different since the UL channel estimates and DL data signals are independent. In this case, the use of longer pilot sequences may have potential benefits.

In this work, RPs and SPs are compared looking at the joint optimization of UL and DL spectral efficiencies. To this end, closed-form rate expressions for the DL are derived, from which the different sources of interference are identified and an analytical comparison between RPs and SPs is offered. Then, a multiobjective optimization problem is formulated considering as objectives the UL and DL sum rates and as optimization variables: the proportion between UL and DL samples; the length of pilot sequences (with RPs); and the proportion of power between pilot and data symbols (with SPs). The analysis shows that with SPs there is an extra correlation between the precoding vectors and the channels to all other UEs (not only to the ones using the same pilot). This creates coherent interference, which limits the SE of the system.

II. SYSTEM MODEL

Consider a multicell system where Φ is the set containing the indices of all BSs. Each BS is equipped with M antennas and serves K UEs. The analysis is focused on an arbitrary

This paper has received funding from ELLIIT and the Swedish Foundation for Strategic Research (SSF) and the Swedish Research Council. The work of L. Sanguinetti was supported in part by ERC Starting MORE under Grant 305123.

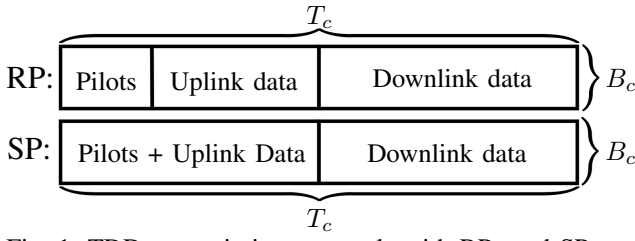


Fig. 1: TDD transmission protocols with RPs and SPs.

BS, denoted as BS_0 , and an arbitrary UE served by BS_0 , denoted as UE_{0k} . A TDD transmission protocol is considered, as illustrated in Fig. 1, where the UL and DL transmissions are divided in time and channel reciprocity is exploited to use the UL channel estimates for DL data precoding. A block fading channel model is considered where the communication channels are assumed to remain static during a time T_c and frequency-flat within a bandwidth B_c , in which both UL and DL transmissions take place. This time-frequency block is referred to as a coherence block and contains $\tau_c = T_c B_c$ complex samples of which, τ_{ul} are used for the UL and τ_{dl} for the DL, with $\tau_c = \tau_{ul} + \tau_{dl}$. The total system bandwidth is B_w and B_w/B_c is assumed to be an integer for simplicity. The channel between the M antennas of $BS_{l'}$ and UE_{li} is defined as $\mathbf{h}_{l'i} \sim \mathcal{CN}(\mathbf{0}, \beta_{l'i} \mathbf{I}_M)$, where $\beta_{l'i}$ accounts for the large-scale fading between $BS_{l'}$ and UE_{li} . The UL pilot and data symbols transmitted by UE_{li} are denoted by ϕ_{li} and \mathbf{s}_{li} , respectively, where the elements of \mathbf{s}_{li} are uncorrelated, have zero mean and unit variance. The transmission powers of UL pilot and data symbols are denoted as q_{li} and ρ_{li} , respectively.¹ Let $\mathbf{x}_{li} \in \mathbb{C}^{\tau_{ul}}$ denote the UL symbols transmitted by UE_{li} as

$$\mathbf{x}_{li} = \begin{cases} \begin{bmatrix} \sqrt{q_{li}} \phi_{li}^T, \sqrt{\rho_{li}} \mathbf{s}_{li}^T \end{bmatrix}^T & \text{with RPs} \\ \sqrt{q_{li}} \phi_{li} + \sqrt{\rho_{li}} \mathbf{s}_{li} & \text{with SPs} \end{cases} \quad (1)$$

where \cdot^T denotes transpose operation. The received signal at the M antennas of BS_0 over the τ_{ul} samples is

$$\mathbf{Z}_0 = \sum_{l \in \Phi} \sum_{i=1}^K \mathbf{h}_{0li} \mathbf{x}_{li}^T + \mathbf{N}_0 \quad (2)$$

where $\mathbf{N}_0 = [\mathbf{n}_{01}, \dots, \mathbf{n}_{0\tau_{ul}}]$ is the noise matrix with $\mathbf{n}_{0j} \sim \mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{I}_M) \forall j \in \{1, \dots, \tau_{ul}\}$ and noise variance σ^2 . The DL data symbol transmitted by BS_l towards UE_{li} in an arbitrary sample of the coherence block is denoted as $s_{li} \in \mathbb{C}$. The received signal at UE_{0k} is

$$y_{0k} = \sum_{l \in \Phi} \sum_{i=1}^K \mathbf{h}_{l0k}^H \mathbf{w}_{li} \sqrt{p_{li}} s_{li} + n_{0k} \quad (3)$$

where p_{li} is the transmission power that BS_l allocates to UE_{li} , which is assumed to be the same with RPs and SPs.¹ The receiver noise at UE_{0k} is $n_{0k} \sim \mathcal{CN}(0, \sigma^2)$. The precoding vector from BS_l to UE_{li} is denoted by \mathbf{w}_{li} and normalized

¹This assumption is made for ease of notation. In Section IV, the power is allocated differently with RPs and SPs.

such that $\mathbb{E}\{\|\mathbf{w}_{li}\|^2\} = 1$. This vector is chosen based on the channel estimates obtained from the UL pilot signals.

III. CHANNEL ESTIMATION

To estimate the channels with both RPs and SPs, linear minimum mean-squared error (LMMSE) estimation based on the UL pilot signals is implemented. Let τ_p out of τ_{ul} samples be used for transmitting pilot signals, then $\tau_p \in [K, \tau_{ul}]$ with RPs and $\tau_p = \tau_{ul}$ with SPs. To avoid cumbersome pilot allocation methods, assume that each BS assigns K orthogonal pilot sequences uniformly at random within its cell. The pilot sequence assigned to UE_{li} is defined as $\phi_{li} \in \mathbb{C}^{\tau_p}$ such that

$$\phi_{li}^H \phi_{l'i'} = \begin{cases} \tau_p & \text{for } i' = i \\ 0 & \text{for } i' \neq i \end{cases} \quad (4)$$

which means that, within each cell, the assigned pilots are mutually orthogonal. The pilot assignment is done independently by each BS, so that if $l' \neq l$ then

$$\phi_{li}^H \phi_{l'i'} = \begin{cases} \tau_p & \text{with probability } \frac{1}{\tau_p} \\ 0 & \text{with probability } 1 - \frac{1}{\tau_p}. \end{cases} \quad (5)$$

For ease of notation, we use $\chi_{l'i'}^{li(\tau_p)} = \frac{\phi_{li}^H \phi_{l'i'}}{\tau_p} \in \{0, 1\}$ to indicate when $UE_{l'i'}$ has been assigned the same pilot as UE_{li} .² To estimate the channel \mathbf{h}_{ll_i} , BS_l performs a despreading operation on the received UL pilot signal (2)³ by multiplying it with $\phi_{li}^*/\sqrt{\tau_p}$. This leads to [20]

$$\mathbf{z}_{li} = \sqrt{q_{li} \tau_p} \mathbf{h}_{ll_i} + \sum_{l' \in \Phi \setminus \{l\}} \sum_{i'=1}^K \chi_{l'i'}^{li(\tau_p)} \sqrt{q_{l'i'} \tau_p} \mathbf{h}_{l'l'i'} + \mathbf{a}_{li} \quad (6)$$

where

$$\mathbf{a}_{li} = \begin{cases} \bar{\mathbf{n}}_p & \text{with RPs} \\ \underbrace{\sum_{l' \in \Phi} \sum_{i'=1}^K \sqrt{\frac{\rho_{l'i'}}{\tau_{ul}}} \mathbf{h}_{l'l'i'} \mathbf{s}_{l'i'}^T \phi_{li}^*}_{\text{Interference due to SPs data symbols}} + \sum_{j=1}^{\tau_{ul}} \mathbf{n}_{0j} \frac{[\phi_{li}]_j^*}{\sqrt{\tau_{ul}}} & \text{with SPs.} \end{cases} \quad (7)$$

In the case of RPs, the term $\bar{\mathbf{n}}_p = \sum_{j=1}^{\tau_p} \mathbf{n}_{0j} \frac{[\phi_{li}]_j^*}{\sqrt{\tau_p}}$ represents the aggregated noise from τ_p (out of τ_{ul}) UL symbols and is distributed as $\bar{\mathbf{n}}_p \sim \mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{I}_M)$. The LMMSE estimate of \mathbf{h}_{ll_i} at BS_l with RPs and SPs is given in the next lemma [20].

Lemma 1. Consider a fixed realization of $\chi_{l'i'}^{li(\tau_p)} \forall l' \in \Phi \setminus \{l\}, i' \in \{1, \dots, K\}$. The LMMSE estimate of \mathbf{h}_{ll_i} with RPs and SPs is

$$\hat{\mathbf{h}}_{ll_i} = \frac{\bar{\gamma}_{li}}{\sqrt{q_{li} \tau_p}} \mathbf{z}_{li} \quad (8)$$

where

$$\bar{\gamma}_{li} = \frac{q_{li} \tau_p \beta_{ll_i}}{\sum_{l' \in \Phi} \sum_{i'=1}^K \chi_{l'i'}^{li(\tau_p)} q_{l'i'} \tau_p \beta_{l'l'i'} + b_l + \sigma^2} \quad (9)$$

²Further details on the random pilot allocation method are found in [20].

³In the case of RPs, the received signal from pilot symbols refers to the τ_p samples where ϕ_{li}^T is transmitted.

and

$$b_l = \begin{cases} 0 & \text{with RPs} \\ \sum_{l' \in \Phi} \sum_{i'=1}^K \rho_{l'i'} \beta_{ll'i'} & \text{with SPs.} \end{cases} \quad (10)$$

Proof: See [20, Sec. III]. ■

The parameter $\bar{\gamma}_{li} \in [0, 1]$ accounts for the quality of the channel estimates. Notice that on average, as τ_p increases, the quality of the estimates improves given that there are more observations available and the pilots are reused more sparsely [20]. However, the samples used for channel estimation are limited by the physical properties of the channel. In particular, $\tau_p \in [K, \tau_{ul}]$ with RPs and $\tau_p = \tau_{ul}$ with SPs. Thus, there is always an estimation error due to noise and pilot contamination (and data symbol interference with SPs).

Remark 1. Notice that if the performance of the UL data rates is neglected, then it is always possible to find a value of $\tau_p \leq \tau_{ul}$ for which the quality of the channel estimates with RPs is greater than with SPs. This is due to the presence of interference from data symbols in the UL with SPs (see (9)). Note that a similar observation is made in [18, Sec. III] with respect to the pilot reuse factor τ_p/K . That is, for larger pilot reuse factors the effect of pilot contamination is reduced at the expense of a larger overhead in the UL.

Remark 2. As it follows from (6), with RPs \mathbf{z}_{li} is a sum of independent circularly symmetric complex Gaussian random variables. This means that the LMMSE estimates coincide with the true MMSE estimates; the channel estimates and their errors are thus independent [21]. On the other hand, with SPs \mathbf{z}_{li} contains an extra interference term (see the first term in second line of (7)) that is not Gaussian due to the presence of data symbols. Thus, with SPs the LMMSE estimates are not the true MMSE estimates, and the channel estimates and its errors are only uncorrelated and not independent [21]. This latter fact will play a key role in deriving the achievable rates.

IV. DL ACHIEVABLE RATES

To evaluate the performance of the system, closed-form expressions for the achievable rates of UE_{0k} with RPs and SPs are derived using the capacity bound methodology from [3]. For a fixed bandwidth B_w , the SE is obtained by scaling the achievable rates with $1/B_w$. Since BS_0 and UE_{0k} are chosen arbitrarily, these rates are representative for any UE in the system with RPs or SPs. To obtain an achievable rate for UE_{0k} under the assumption that UEs do not have CSI, the term $|\mathbb{E}\{\mathbf{h}_{00k}^H \mathbf{w}_{0k}\}| \sqrt{p_{0k} \varsigma_{0k}}$ is added and subtracted from y_{0k} in (3). The result can be interpreted as an equivalent single-input single-output (SISO) system with deterministic gain $|\mathbb{E}\{\mathbf{h}_{00k}^H \mathbf{w}_{0k}\}|^2 p_{0k}$ and non-Gaussian effective noise with both RPs and SPs. A DL achievable rate, i.e., lower bound on the capacity, for UE_{0k} is thus given in the next lemma.

Lemma 2. An achievable rate for the DL between BS_0 and UE_{0k} with RPs and SPs is

$$R_{0k} = B_w \frac{\tau_{dl}}{\tau_c} \log_2(1 + \text{SINR}_{0k}) \quad (11)$$

where

$$\text{SINR}_{0k} = \frac{|\mathbb{E}\{\mathbf{h}_{00k}^H \mathbf{w}_{0k}\}|^2 p_{0k}}{\sum_{l \in \Phi} \sum_{i=1}^K \mathbb{E}\{|\mathbf{h}_{l0k}^H \mathbf{w}_{li}|^2\} p_{li} - |\mathbb{E}\{\mathbf{h}_{00k}^H \mathbf{w}_{0k}\}|^2 p_{0k} + \sigma^2} \quad (12)$$

is the effective signal-to-interference plus noise ratio (SINR) expression of UE_{0k} .

Proof: It follows from applying standard bounding techniques on the capacity of the deterministic effective channel $|\mathbb{E}\{\mathbf{h}_{00k}^H \mathbf{w}_{0k}\}|^2$ with uncorrelated non-Gaussian effective noise [3, Ch. 2]. The variance of this noise is given in the denominator of (12) which follows from having independent and zero-mean DL data symbols and noise. ■

To obtain further insights into the interplay of design parameters such as number of BS antennas and length of the coherence block, maximum ratio transmission (MRT) is used and a closed-form expression for the effective SINR in (12) is derived. In particular, it is assumed that the precoding vector used by BS_l to send data symbols towards UE_{li} is defined as

$$\mathbf{w}_{li} = \frac{v_{li} \hat{\mathbf{h}}_{lli}}{\sqrt{\mathbb{E}\left\{\|v_{li} \hat{\mathbf{h}}_{lli}\|^2\right\}}} = \sqrt{\frac{\gamma_{li}}{M \tau_p q_{li} \beta_{lli}}} \mathbf{z}_{li} \quad (13)$$

where $v_{li} = 1/(\bar{\gamma}_{li} \sqrt{M \beta_{lli}})$ such that the effective channel gain of UE_{0k} (i.e., $|\mathbb{E}\{\mathbf{h}_{00k}^H \mathbf{w}_{0k}\}|^2$) with both RPs and SPs is $M \beta_{00k} \gamma_{0k}$. The parameter γ_{li} reflects the quality of the channel estimate $\hat{\mathbf{h}}_{lli}$ and is defined as

$$\gamma_{li} = \mathbb{E}\left\{\frac{1}{\bar{\gamma}_{li}}\right\}^{-1} = \frac{q_{li} \tau_p \beta_{lli}}{q_{li} \tau_p \beta_{lli} + \sum_{l' \in \Phi \setminus \{l\}} \sum_{i'=1}^K q_{l'i'} \beta_{ll'i'} + b_l + \sigma^2}. \quad (14)$$

This follows from the definition of $\chi_{l'i'}^{li(\tau_p)}$ in Section III considering that $\sum_{i'=1}^K \chi_{l'i'}^{li(\tau_p)}$ with $l' \neq l$ is a Bernoulli distributed random variable with success probability K/τ_p and independent across $l' \in \Phi \setminus \{l\}$. Thus, the probability of a UE in each cell having the same pilot as UE_{0k} is K/τ_p . Assuming that this UE is selected uniformly at random within the cell, the following result holds

$$\mathbb{E}\left\{\sum_{l' \in \Phi \setminus \{l\}} \sum_{i'=1}^K \chi_{l'i'}^{li(\tau_p)} q_{l'i'} \beta_{ll'i'}\right\} = \sum_{l' \in \Phi \setminus \{l\}} \frac{K}{\tau_p} \frac{1}{K} \left(\sum_{i'=1}^K q_{l'i'} \beta_{ll'i'}\right). \quad (15)$$

The choice of MRT is further motivated by its low complexity and the fact that as the number of antennas goes to infinity the SE of MRT is close to optimal for i.i.d. Rayleigh fading channels, by virtue of asymptotic favorable propagation and channel hardening [3]. With RPs, the expectations in (12) are obtained by first exploiting the properties of circularly symmetric complex Gaussian random vectors (see Lemma 3 in the Appendix) and the fact that the channel estimates and errors are independent as in [3]. Second, the results in (14) and (15) are used to average over the random pilot assignment. To differentiate the results between RPs and SPs, the following

$$\text{SINR}_{0k}^{\text{SP}} = \frac{Mp_{0k}^{\text{SP}}\beta_{00k}\gamma_{0k}^{\text{SP}}}{\underbrace{\frac{M}{\tau_{ul}} \sum_{l \in \Phi \setminus \{0\}} \sum_{i=1}^K \frac{q_{0k}^{\text{SP}}\beta_{l0k}}{q_{li}^{\text{SP}}\beta_{lli}} p_{li}^{\text{SP}} \beta_{l0k} \gamma_{li}^{\text{SP}}}_{\text{Coherent interference caused by pilot contamination}} + \underbrace{\frac{M}{\tau_{ul}} \sum_{l \in \Phi} \sum_{i=1}^K \frac{\rho_{0k}^{\text{SP}}\beta_{l0k}}{q_{li}^{\text{SP}}\beta_{lli}} p_{li}^{\text{SP}} \beta_{l0k} \gamma_{li}^{\text{SP}}}_{\text{Additional coherent interference}} + \underbrace{\sum_{l \in \Phi} \beta_{l0k} \sum_{i=1}^K p_{li}^{\text{SP}} + \sigma^2}_{\text{Non-coherent interference and noise}}} \quad (18)$$

notation is introduced: the UL power for pilot symbols sent by UE_{li} and DL power for data symbols sent to UE_{li} are denoted as q_{li}^{RP} and p_{li}^{RP} , respectively. The terms γ_{0k}^{RP} and γ_{li}^{RP} are defined as in (14) for $\tau_p \in [K, \tau_{ul}]$ and $b_l = 0$. Finally, the closed-form SINR expression with RPs is

$$\text{SINR}_{0k}^{\text{RP}} = \frac{Mp_{0k}^{\text{RP}}\beta_{00k}\gamma_{0k}^{\text{RP}}}{\frac{M}{\tau_p} \sum_{l \in \Phi \setminus \{0\}} \sum_{i=1}^K \frac{q_{0k}^{\text{RP}}\beta_{l0k}}{q_{li}^{\text{RP}}\beta_{lli}} p_{li}^{\text{RP}} \beta_{l0k} \gamma_{li}^{\text{RP}} + \sum_{l \in \Phi} \beta_{l0k} \sum_{i=1}^K p_{li}^{\text{RP}} + \sigma^2} \quad (16)$$

In the case of SPs, the channel estimates and errors are not independent and therefore a different approach is needed to obtain the closed-form expression of the effective SINR. This result is summarized in the next theorem.

Theorem 1. *The closed-form expression of UE_{0k} effective SINR in (12) with SPs is given by (18) at the top of this page. The UL transmission powers of pilot and data symbols sent by UE_{li} are denoted as q_{li}^{SP} and ρ_{li}^{SP} , respectively. The DL transmission power of data symbols sent to UE_{li} is p_{li}^{SP} . The term γ_{li}^{SP} is defined as in (14) where $\tau_p = \tau_{ul}$ and b_l is given in (10).*

Proof: The proof can be found in the Appendix. ■

The term *coherent interference* is used to identify the interfering precoded signals that add constructively when they are sent through the channel so that the received power is proportional to M . *Non-coherent interference* refers instead to the interfering precoded signals that do not add constructively when sent through the channel and therefore the received power does not scale with M . In (18), it is shown that the pilot contamination term with SPs is reduced by a factor of $1/\tau_{ul}$ in comparison to $1/\tau_p$ with RPs (see the first term in the denominator of (16)) since $\tau_p \leq \tau_{ul}$. On the other hand, there is additional coherent interference with SPs due to the data symbol interference in the channel estimates. After the submission of this paper, a similar analysis is conducted in [19] arriving also to an expression for the DL achievable rate with SPs [19, Eq. (13)]. This expression is almost the same as (18) with a small difference in the noise term.

V. SE COMPARISON BETWEEN RPs AND SPs

Notice that in Lemma 2 the difference between the DL data rates with RPs and SPs lies only in the definition of the channel estimates and power allocation. Thus, with both RPs and SPs there is a strong relationship between the UL and DL data rates due to the influence of the channel estimation quality in the DL rates. In the case of RPs, the DL data rates increase monotonically with τ_p since the channel estimates improve (see (14)) and pilot contamination is reduced (see (16)). However, the UL data rates are zero when $\tau_p = \tau_{ul}$ [20].

On the other hand, with SPs the DL data rates increase when more power is allocated to the pilot sequences since the term b_l becomes less significant (see (10) and (14)) and the *additional coherence interference* is reduced (see (18)). However, if the entire power budget is allocated to pilot sequences, then the UL data rates are zero [20]. Moreover, with both RPs and SPs, the proportion between τ_{ul} and τ_{dl} affects the UL and DL data rates. To offer a comprehensive comparison between RPs and SPs, multiobjective optimization theory is used with the UL and DL rates as the two conflicting objectives. Multiobjective optimization is a mathematical framework to solve optimization problems with multiple objectives that are mutually conflicting; that is, improving one of the objectives would eventually deteriorate other ones [22]. In this class of optimization problems, the multiple objectives are not in any particular order and since they are mutually conflicting, there is no global optimum. To obtain insights into the relationship between objectives, an attainable objective set is defined as the set of all objective points that can be achieved by some configuration of the available resources. Moreover, the subset of the attainable objective set for which no individual objective can be improved without deteriorating some other objective is called the Pareto boundary and is the closest to an optimal solution that exists in multiobjective optimization. As an example, consider an M -antenna BS that communicates with a single-antenna UE using the TDD protocol defined with RPs in Section II (see Fig. 1) assuming imperfect CSI and maximum ratio processing. The objectives considered are the UL and DL rates defined as in [3] for an UL and DL signal-to-noise ratio (SNR) of 0 dB. The available resources are the samples in the coherence interval τ_c which can be used either for DL, UL, or pilots in any proportion as long as $\tau_p \geq 1$, $\tau_{ul} \geq 1$ and $\tau_c \geq 1$. The attainable objective set (or rate region in this case) is illustrated in Fig. 2 as a shaded area and the solid line depicts the Pareto boundary.

A. Multiobjective Optimization with RPs and SPs

Denote by $R_{0k}^{(\text{SC})}$ and $\text{SINR}_{0k}^{(\text{SC})}$ the data rate and effective SINR of UE_{0k}, respectively, where the notation SC stands for the scenario to be evaluated. The scenarios DL-RP and DL-SP refer to the DL with RPs and SPs, respectively, where the data rates and effective SINRs are defined as in Section IV. The scenarios UL-RP and UL-SP refer to the UL with RPs and SPs, respectively, for which the data rates are defined as

$$R_{0k}^{(\text{UL-RP})} = B_w \left(\frac{\tau_{ul} - \tau_p}{\tau_c} \right) \log_2 \left(1 + \text{SINR}_{0k}^{(\text{UL-RP})} \right) \quad (17)$$

$$R_{0k}^{(\text{UL-SP})} = B_w \frac{\tau_{ul}}{\tau_c} \log_2 \left(1 + \text{SINR}_{0k}^{(\text{UL-SP})} \right) \quad (19)$$

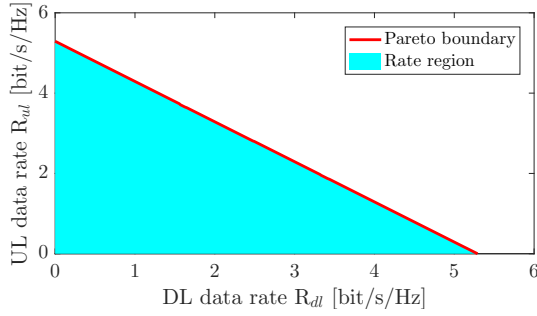


Fig. 2: UL and DL rate region (i.e., attainable objective set) and Pareto boundary for UL and DL SNR = 0 dB.

where $\text{SINR}_{0k}^{(\text{UL-RP})}$ and $\text{SINR}_{0k}^{(\text{UL-SP})}$ are respectively the UE $_{0k}$ effective UL SINRs with RPs and SPs, given in [20]. The optimization objectives considered here are the UL and DL total sum data rates defined as

$$f^{(\text{sc})} = \frac{\tau^{(\text{sc})}}{\tau_c} \sum_{l \in \Phi} \sum_{k=1}^K \log_2 \left(1 + \text{SINR}_{lk}^{(\text{sc})} \right) \quad (20)$$

in which $\tau^{(\text{sc})}$ accounts for the symbols used for data transmission, defined in the UL as $\tau^{(\text{UL-RP})} = \tau_{ul} - \tau_p$ with RPs and $\tau^{(\text{UL-SP})} = \tau_{ul}$ with SPs, whereas in the DL $\tau^{(\text{DL-RP})} = \tau^{(\text{DL-SP})} = \tau_{dl}$ with both RPs and SPs.

The power allocation has an important impact on the data rates. In the UL, power control is essential to avoid UEs with good channels to overshadow the received signals from UEs with poor channels. This is particularly important in Massive MIMO where the use of low-cost hardware (e.g., low resolution analog-to-digital converters) is key to reduce complexity and power consumption. The average transmission powers per data and pilot symbols with RPs are denoted as ρ_{ul-d}^{RP} and ρ_{ul-p}^{RP} , respectively; and with SPs as ρ_{ul-d}^{SP} and ρ_{ul-p}^{SP} , respectively. Following [11], statistical channel inversion power control is assumed in the UL where

$$\rho_{li}^{\text{RP}} = \frac{\rho_{ul-d}^{\text{RP}}}{\beta_{li}}, \quad q_{li}^{\text{RP}} = \frac{\rho_{ul-p}^{\text{RP}}}{\beta_{li}}, \quad \rho_{li}^{\text{SP}} = \frac{\rho_{ul-d}^{\text{SP}}}{\beta_{li}}, \quad q_{li}^{\text{SP}} = \frac{\rho_{ul-p}^{\text{SP}}}{\beta_{li}}. \quad (21)$$

Note that in the UL, the transmission power per symbol with SPs is shared between pilot and data symbols whereas with RPs this power can be fully utilized since pilot and data symbols are transmitted distjointly. The average transmission power per UL symbol is denoted as ρ_{ul} so that

$$\begin{aligned} \rho_{ul-d}^{\text{RP}} &= \rho_{ul-p}^{\text{RP}} = \rho_{ul}, \\ \rho_{ul-p}^{\text{SP}} &= \Delta \rho_{ul}, \quad \rho_{ul-d}^{\text{SP}} = (1 - \Delta) \rho_{ul}, \end{aligned} \quad (22)$$

where $\Delta \in [0, 1]$ is the proportion between the power allocated to pilot and data symbols with SPs.

In the DL, the power allocation is assumed to be performed by each cell individually to make the network scalable with the number of cells. That is, each BS allocates the power of its served UEs and the intercell interference is updated in a sequential manner until convergence is achieved.

Two power allocation strategies are considered: maximum-minimum (max-min) fairness and maximum-sum (max-sum) rate. The max-min fairness problem aims at maximizing the minimum SINR providing equal SINR to all UEs in each particular cell. This problem can be solved by formulating an equivalent linear feasibility problem for a specific target SINR and then conducting a bisection search to find the maximum SINR target. Note that a closed-form solution for this problem also exists [3, Ch. 5]. The max-sum rate problem aims at maximizing the aggregated SE of UEs within each cell. Since statistical channel inversion is employed in the UL along with a random pilot allocation scheme, it can be shown⁴ that the maximum DL sum rate problem in each individual cell is a convex optimization problem that can be solved through water-filling power allocation.

The available resources (or variables) for the multiobjective optimization problem are then: the number of UL symbols τ_{ul} (note that since τ_c is fixed the number of DL symbols is obtained directly as $\tau_{dl} = \tau_c - \tau_{ul}$), the number of pilot symbols τ_p with RPs and the proportion of UL power Δ between pilot and data symbols with SPs. The multiobjective optimization problems with RPs and SPs can be cast as

$$\underset{\tau_p \in [K, \tau_{ul}], \tau_{ul} \in [K, \tau_c]}{\text{maximize}} \quad \left\{ f^{(\text{UL-RP})}, f^{(\text{DL-RP})} \right\} \quad (23)$$

and

$$\underset{\Delta \in [0, 1], \tau_{ul} \in [K, \tau_c]}{\text{maximize}} \quad \left\{ f^{(\text{UL-SP})}, f^{(\text{DL-SP})} \right\}. \quad (24)$$

To obtain the Pareto boundary of (23) and (24), the ϵ -constraint method is used where one objective is maximized while the other is maintained above a certain target and vice-versa [23]. A point on the Pareto boundary is found when one of the objectives cannot be further increased without decreasing the other. An exhaustive search can be conducted over τ_{ul} and τ_p since these are integers. For $\Delta \in [0, 1]$, $f^{(\text{DL-SP})}$ increases monotonically since the quality of the channel estimates improves with Δ . In the case of $f^{(\text{UL-SP})}$, it can be proved⁵ that the denominator of $\text{SINR}_{lk}^{\text{UL-SP}} \forall l \in \Phi, k \in \{1, \dots, K\}$ is a convex function of Δ . This means that $f^{(\text{UL-SP})}$ is a quasi-concave function of Δ and the maximum point can be found through bisection search algorithms.

VI. NUMERICAL RESULTS

Numerical results are used to get insights into the sum rate performance with RPs and SPs. The BSs are dropped in a square area following an homogeneous PPP⁶ with a density of 100 [BSs/km²] where the area is properly scaled to simulate an average number of 30 BSs. The wraparound technique is used to avoid cell edge effects. K UEs per cell are distributed uniformly within the Voronoi region around each BS. The

⁴With both RPs and SPs, when combining (14) with (21) $\gamma_{li} = \gamma_{li'} \forall i, i' \in \{1, \dots, K\}$. Then, by fixing the total DL power for each BS, the denominator of (16) and (18) is the same for all UEs within the same cell. Thus, R_{0k}^{RP} and R_{0k}^{SP} are concave functions of p_{0k}^{RP} and p_{0k}^{SP} respectively.

⁵It follows from combining (21) and (22) with $\text{SINR}_{0k}^{(\text{UL-SP})}$ in [20] and evaluating the second derivative of the denominator with respect to Δ .

⁶Due to the translation invariance property, the average sum rates per cell is equal to the average data rate per UEs times K .

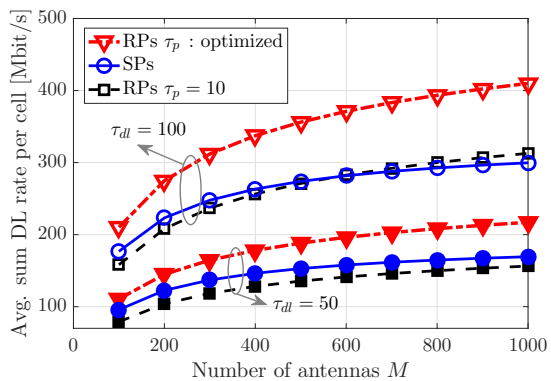


Fig. 3: Average sum DL rate per cell versus M for $B_w = 20$ [MHz], $K = 10$, $\tau_c = 200$ and $\frac{\rho_{ul}}{\sigma^2} = 0.25$. The DL power allocation is max-sum rate per cell. The maximum DL power is set so that the lowest cell edge SNR is equal to $\frac{\rho_{ul}}{\sigma^2}$.

large-scale fading is modeled as $\beta_{l_{ik}} = \omega^{-1} d_{l_{ik}}^{-\alpha}$ where $d_{l_{ik}}$ [km] is the distance between BS $_{l_i}$ and UE $_{l_{ik}}$, $\alpha = 3.76$ is the pathloss exponent and ω is the pathloss at a reference distance of 1 [km]. In all results with SPs, the proportion between UL pilot and data symbol power Δ is optimized to maximize the UL sum rate and the received pilot symbols are estimated and subtracted from the received UL signal for data detection [20].

Fig. 3 depicts the average DL sum rate per cell versus number of antennas with $K = 10$, $\tau_c = 200$ and $\tau_{dl} = \{50, 100\}$ to illustrate the effect of available UL symbols for channel estimation. With RPs, two cases are considered for selecting the number of channel estimation symbols: *i*) τ_p is optimized to maximize the UL sum rate and *ii*) τ_p is fixed and equal to K . The results show that for $\tau_p = K$ and $\tau_{dl} = 50$, the average DL sum rate per cell with RPs is smaller than with SPs. However, when τ_p is optimized to maximize the UL sum rate the resulting average DL sum rate per cell with RPs is greater than with SPs. These results are in line with [20] indicating that the reduction of pilot contamination offered by SPs is counteracted by the extra interference from data symbols in the channel estimation process and RPs have greater SE when τ_p is optimized. It is worth mentioning that for $\tau_{dl} = 50$ the relative SE of SPs with respect to RPs is better since there are more samples available in the UL for improving the channel estimates. Thus, when RPs and SPs have the same proportion of τ_{ul} and τ_{dl} the benefits of SPs are more noticeable for τ_{ul} much larger than K .

Fig. 4 and Fig. 5 depict the Pareto boundary for the UL and DL sum data rates of all BSs with max-min fairness and max-sum rate DL power allocation, respectively, with $K = 20$. In contrast to Fig. 3, the proportion of τ_{ul} and τ_{dl} is optimized independently with RPs and SPs. The results show that when τ_c is very close to K the rate region with RPs is contained within the rate region with SPs. This follows from the fact that in this extreme case RPs have almost no samples for the UL or DL data transmission whereas SPs can use many more UL samples for data. On the other hand, when τ_c is significantly

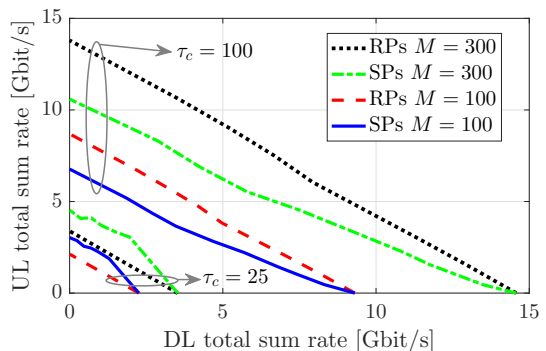


Fig. 4: Pareto boundary of UL and DL total sum rate for $B_w = 20$ [MHz], $K = 20$ and $\frac{\rho_{ul}}{\sigma^2} = 0.25$. The DL power allocation is max-min fairness. The maximum DL power is set so that the lowest cell edge SNR is equal to $\frac{\rho_{ul}}{\sigma^2}$.

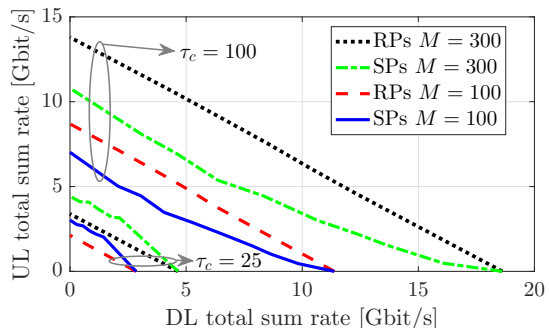


Fig. 5: Pareto boundary of UL and DL total sum rate for $B_w = 20$ [MHz], $K = 20$ and $\frac{\rho_{ul}}{\sigma^2} = 0.25$. The DL power allocation is max-sum rate per cell. The maximum DL power is set so that the lowest cell edge SNR is equal to $\frac{\rho_{ul}}{\sigma^2}$.

larger than K the rate region with SPs is contained within the rate region with RPs. When the number of BS antennas is increased, the corresponding Pareto boundaries with RPs and SPs are expanded, allowing higher sum rates in both UL and DL, while the behavior with respect to τ_c and K remains unchanged. Notice that in practice τ_c would be at least in the order of hundreds, even for high mobility scenarios [3]. Therefore, for systems designed to multiplex tens of UEs, RPs with optimized pilot overhead provide better SE.

VII. CONCLUSION

The use of SPs in the UL of Massive MIMO allows a reduction of the pilot contamination in the DL at the expense of adding interference in the channel estimation process that reduces the channel estimation quality and creates additional coherent interference in the DL. In comparison with SPs, RPs are able to offer higher average DL sum rate per cell when the length of the pilot sequences is optimized to maximize the UL sum rate. The result of multiobjective optimization for UL and DL sum rates indicates that SPs offer higher SE when the number of samples in the coherence block is very small and many UEs need to be multiplexed. However, in more practical cases where the coherence block is typically

several times larger than the number of multiplexed UEs than RPs with optimized pilot length offer better SE. Further improvements with SPs can be achieved through iterative decoding algorithms. However, it is not clear if the gains are large enough to justify the extra amount of signal processing needed. Moreover, these algorithms can potentially be used along with RPs.

APPENDIX PROOF OF THEOREM 1

Lemma 3. Consider two independent complex random vectors $\mathbf{x} \sim \mathcal{CN}(\mathbf{0}, \sigma_x^2 \mathbf{I}_M)$ and $\mathbf{y} \sim \mathcal{CN}(\mathbf{0}, \sigma_y^2 \mathbf{I}_M)$. The following results hold [3]

$$\mathbb{E} \left\{ \|\mathbf{x}\|^4 \right\} = M(M+1)\sigma_x^4 \quad (25)$$

$$\mathbb{E} \left\{ |\mathbf{x}^H \mathbf{y}|^2 \right\} = M\sigma_x^2 \sigma_y^2. \quad (26)$$

A closed-form expression for SINR_{0k} in (12) with SPs is obtained by using the results in Lemma 3 to first calculate

$$\begin{aligned} \left| \mathbb{E} \{ \mathbf{h}_{00k}^H \mathbf{w}_{0k} \} \right|^2 &= \frac{\gamma_{0k}^{\text{SP}}}{M\tau_{ul}q_{0k}^{\text{SP}}\beta_{00k}} \left| \mathbb{E} \{ \mathbf{h}_{00k}^H \mathbf{z}_{0k} \} \right|^2 \\ &= \frac{\gamma_{0k}^{\text{SP}} \left| \mathbb{E} \left\{ \|\mathbf{h}_{00k}\|^2 \left(\sqrt{q_{0k}^{\text{SP}}\tau_{ul}} + \sqrt{\frac{\rho_{0k}^{\text{SP}}}{\tau_{ul}}} \mathbf{s}_{0k}^T \phi_{0k}^* \right) \right\} \right|^2}{M\tau_{ul}q_{0k}^{\text{SP}}\beta_{00k}} \\ &= M\beta_{00k}\gamma_{0k}^{\text{SP}}. \end{aligned} \quad (27)$$

Second, let $a_{l'i'}^{li} = \chi_{l'i'}^{li(\tau_{ul})} \sqrt{q_{l'i'}^{\text{SP}}\tau_{ul}} + \sqrt{\frac{\rho_{l'i'}^{\text{SP}}}{\tau_{ul}}} \mathbf{s}_{l'i'}^T \phi_{l'i'}^*$ for $l, l' \in \Phi$, $i, i' \in \{1, \dots, K\}$ and compute

$$\begin{aligned} \sum_{i=1}^K \mathbb{E} \left\{ |\mathbf{h}_{l0k}^H \mathbf{w}_{li}|^2 \right\} p_{li}^{\text{SP}} &= \mathbb{E} \left\{ \sum_{i=1}^K \frac{p_{li}^{\text{SP}} \gamma_{li}^{\text{SP}} \mathbb{E} \left\{ |\mathbf{h}_{l0k}^H \mathbf{z}_{li}|^2 \right\} \chi_{l'i'}^{li(\tau_{ul})}}{M\tau_{ul}q_{li}^{\text{SP}}\beta_{li}} \right\} \\ &= \mathbb{E} \left\{ \sum_{i=1}^K \frac{p_{li}^{\text{SP}} \gamma_{li}^{\text{SP}}}{M\tau_{ul}q_{li}^{\text{SP}}\beta_{li}} \left(\mathbb{E} \left\{ \|\mathbf{h}_{l0k}\|^4 |a_{0k}^{li}|^2 \right\} \chi_{l'i'}^{li(\tau_{ul})} \right) \right. \\ &\quad \left. + \mathbb{E} \left\{ \|\mathbf{h}_{l0k}\|^2 \left| \chi_{l'i'}^{li(\tau_{ul})} \right| \sum_{\substack{l' \in \Phi \\ \{l', i'\} \neq \{0, k\}}} \sum_{i=1}^K \mathbb{E} \left\{ |a_{l'i'}^{li}|^2 \right\} \chi_{l'i'}^{li(\tau_{ul})} \right) \right\} \\ &= \frac{M}{\tau_{ul}} \mathbb{E} \left\{ \sum_{i=1}^K \beta_{l0k} p_{li}^{\text{SP}} \gamma_{li}^{\text{SP}} \left(\left(\chi_{0k}^{li(\tau_{ul})} q_{0k}^{\text{SP}} \tau_{ul} + \rho_{0k}^{\text{SP}} \right) \frac{\beta_{l0k}}{q_{li}^{\text{SP}} \beta_{li}} + \frac{1}{\gamma_{li}^{\text{SP}}} \right) \right\} \\ &= \begin{cases} M p_{0k}^{\text{SP}} \beta_{l0k} \gamma_{0k}^{\text{SP}} + \frac{M}{\tau_{ul}} \sum_{i=1}^K \frac{\rho_{0k}^{\text{SP}} \beta_{l0k} p_{0i}^{\text{SP}} \gamma_{0i}^{\text{SP}}}{q_{0i}^{\text{SP}} \beta_{00i}} + \beta_{00k} \sum_{i=1}^K p_{0i}^{\text{SP}} & \text{for } l = 0 \\ \frac{M}{\tau_{ul}} \sum_{i=1}^K \frac{(q_{0k}^{\text{SP}} + \rho_{0k}^{\text{SP}}) \beta_{l0k} p_{li}^{\text{SP}} \gamma_{li}^{\text{SP}}}{q_{li}^{\text{SP}} \beta_{li}} + \beta_{l0k} \sum_{i=1}^K p_{li}^{\text{SP}} & \text{for } l \neq 0 \end{cases} \quad (28) \end{aligned}$$

where the last equality follows from the definition of $\chi_{0k}^{li(\tau_{ul})}$ in Section III and the result in (15). Finally, plugging (27) and (28) into (12) concludes the proof.

REFERENCES

- [1] Cisco, "Visual networking index: Global mobile data traffic forecast update, 2014–2019," Tech. Rep., 2015.
- [2] E. G. 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, 2014.
- [3] T. L. Marzetta, E. G. Larsson, H. Yang, and H. Q. Ngo, *Fundamentals of Massive MIMO*. Cambridge Press, 2016.
- [4] J. Vinogradova, E. Björnson, and E. G. Larsson, "On the separability of signal and interference-plus-noise subspaces in blind pilot decontamination," in *Proc. IEEE ICASSP*, 2016, pp. 3421–3425.
- [5] H. Q. Ngo and E. G. Larsson, "EVD-based channel estimation in multicell multiuser MIMO systems with very large antenna arrays," in *Proc. IEEE ICASSP*, 2012, pp. 3249–3252.
- [6] R. R. Müller, L. Cottatellucci, and M. Vehkaperä, "Blind pilot decontamination," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 773–786, 2014.
- [7] H. Huh, G. Caire, H. C. Papadopoulos, and S. A. Ramprasad, "Achieving "massive MIMO" spectral efficiency with a not-so-large number of antennas," *IEEE Trans. Wireless Commun.*, vol. 11, no. 9, pp. 3226–3239, 2012.
- [8] 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, 2013.
- [9] E. Björnson, J. Hoydis, and L. Sanguinetti, "Massive MIMO has unlimited capacity," *Submitted to IEEE Trans. Wireless Commun.*, 2017. [Online]. Available: <http://arxiv.org/abs/1705.00538>
- [10] Y. Li, Y.-H. Nam, B. L. Ng, and J. Zhang, "A non-asymptotic throughput for massive MIMO cellular uplink with pilot reuse," in *Proc. IEEE GLOBECOM*, 2012, pp. 4500–4504.
- [11] E. Björnson, E. Larsson, and M. Debbah, "Massive MIMO for maximal spectral efficiency: How many users and pilots should be allocated?" *IEEE Trans. Wireless Commun.*, vol. 15, no. 2, pp. 1293–1308, 2016.
- [12] R. Mochaourab, E. Björnson, and M. Bengtsson, "Adaptive pilot clustering in heterogeneous massive MIMO networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 8, pp. 5555–5568, 2016.
- [13] P. Hoeher and F. Tufvesson, "Channel estimation with superimposed pilot sequence," in *IEEE Proc. GLOBECOM*, 1999, pp. 2162–2166.
- [14] G. T. Zhou, M. Viberg, and T. McKelvey, "A first-order statistical method for channel estimation," *IEEE Signal Process. Lett.*, vol. 10, no. 3, pp. 57–60, 2003.
- [15] A. T. Asyhari and S. ten Brink, "Orthogonal or superimposed pilots? a rate-efficient channel estimation strategy for stationary MIMO fading channels," *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 2776–2789, 2017.
- [16] F. Li, H. Wang, M. Ying, W. Zhang, and J. Lu, "Channel estimations based on superimposed pilots for massive MIMO uplink systems," in *IEEE Proc. WCSP*, 2016, pp. 1–5.
- [17] H. Zhang, S. Gao, D. Li, H. Chen, and L. Yang, "On superimposed pilot for channel estimation in multicell multiuser MIMO uplink: Large system analysis," *IEEE Trans. Veh. Technol.*, vol. 65, no. 3, pp. 1492–1505, 2016.
- [18] K. Upadhyaya, S. A. Vorobyov, and M. Vehkaperä, "Superimposed pilots are superior for mitigating pilot contamination in massive MIMO," *IEEE Trans. Signal Process.*, vol. 65, no. 11, pp. 2917–2932, 2017.
- [19] K. Upadhyaya, S. A. Vorobyov, and M. Vehkaperä, "Downlink performance of superimposed pilots in massive MIMO systems," *Submitted to IEEE Trans. Wireless Commun.*, 2017. [Online]. Available: <https://arxiv.org/abs/1606.04476v2>
- [20] D. Verenzuela, E. Björnson, and L. Sanguinetti, "Spectral and energy efficiency of superimposed pilots in uplink massive MIMO," *Submitted to IEEE Trans. Wireless Commun.*, 2017. [Online]. Available: <https://arxiv.org/abs/1709.07722>
- [21] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice Hall, 1993.
- [22] E. Björnson, E. Jorswieck, M. Debbah, and B. Ottersten, "Multiobjective signal processing optimization: The way to balance conflicting metrics in 5g systems," *IEEE Signal Process. Mag.*, vol. 31, no. 6, pp. 14–23, 2014.
- [23] E. Björnson and E. Jorswieck, "Optimal resource allocation in coordinated multi-cell systems," *Foundations and Trends® in Communications and Information Theory*, vol. 9, no. 2–3, pp. 113–381, 2013. [Online]. Available: <http://dx.doi.org/10.1561/01000000069>