Massive MIMO Systems with Hardware-Constrained Base Stations


Document Version:
Peer reviewed version

Queen's University Belfast - Research Portal:
Link to publication record in Queen's University Belfast Research Portal

General rights
Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.

Download date:03. Jan. 2019
MASSIVE MIMO SYSTEMS WITH HARDWARE-CONSTRAINED BASE STATIONS

Emil Björnson†‡ Michail Matthaiou† Mérouane Debbah*

*Alcatel-Lucent Chair on Flexible Radio, SUPELEC, Gif-sur-Yvette, France
†ACCESS Centre, Dept. of Signal Processing, KTH Royal Institute of Technology, Stockholm, Sweden
‡ECIT Institute, Queen’s University Belfast, U.K. and S2, Chalmers University of Technology, Sweden

ABSTRACT

Massive multiple-input multiple-output (MIMO) systems are cellular networks where the base stations (BSs) are equipped with unconventionally many antennas. Such large antenna arrays offer huge spatial degrees-of-freedom for transmission optimization; in particular, great signal gains, resilience to imperfect channel knowledge, and small inter-user interference are all achievable without extensive inter-cell coordination. The key to cost-efficient deployment of large arrays is the use of hardware-constrained base stations with low-cost antenna elements, as compared to today’s expensive and power-hungry BSs. Low-cost transceivers are prone to hardware imperfections, but it has been conjectured that the excessive degrees-of-freedom of massive MIMO would bring robustness to such imperfections. We herein prove this claim for an uplink channel with multiplicative phase-drift, additive distortion noise, and noise amplification. Specifically, we derive a closed-form scaling law that shows how fast the imperfections increase with the number of antennas.

Index Terms— Achievable uplink rates, channel estimation, massive MIMO, scaling laws, transceiver hardware imperfections.

1. INTRODUCTION

Massive densification, in terms of more antennas per unit area, is a key enabler to higher area throughput in future wireless networks [1]. This is achieved using multi-user MIMO techniques, by adding more antennas to the macro BSs and/or distributing the antennas by ultra-dense deployment of small BSs. These approaches are non-conflicting, since the former operates in current frequency bands and ultra-dense deployment of small BSs. Towards this end, [1] showed that the BSs are equipped with large antenna arrays with hundreds (or even thousands) of antennas, which are used to serve tens (or even hundreds) of users. In other words, the number of antennas, \( N \), and number of users per BS, \( K \), are unconventionally large, but can differ by an order of magnitude. For this reason, massive MIMO brings unprecedented spatial degrees-of-freedom, which enable precoding with strong signal gains, give near-orthogonal user channels, and resilience to imperfect channel knowledge [6]. Apart from achieving high area throughput, recent works have investigated additional ways to capitalize on the huge degrees-of-freedom offered by massive MIMO. Towards this end, [1] showed that massive MIMO enables implicit coordination between systems that operate in the same band. Moreover, it was shown in [4] and [7] that the transmit powers can be reduced as \( 1/\sqrt{N} \) with only a minor loss in throughput. This offers major reductions in the emitted power, but is actually bad from an energy efficiency (EE) perspective—the EE is maximized by increasing the transmit power with \( N \) to compensate for the increasing circuit power [8].

The work presented here explores whether the huge degrees-of-freedom offered by massive MIMO provide robustness to transceiver hardware imperfections; for example, phase drifts, quantization errors, and noise amplification. Robustness to hardware imperfections has been conjectured in overview articles, such as [5], and is notably important since the deployment cost of massive MIMO scales linearly with \( N \) unless we resort to using cheaper hardware with larger imperfections. Constant envelope precoding was analyzed in [2] to facilitate the use of power-efficient amplifiers, while the impact of phase drifts was analyzed and simulated in [10,11]. A preliminary proof of the conjecture was provided in [12], but the authors considered only additive distortions and, thus, ignored other characteristics of hardware imperfections. It was shown that the distortion variance can increase as \( \sqrt{N} \) with only minor throughput losses.

In this work, we consider an uplink massive MIMO system with hardware imperfections that cause phase drifts, additive distortions, and noise amplification; this model is more general compared to [10,12] which investigated merely one of these effects. We derive a new linear minimum mean square error (LMMSE) channel estimator and closed-form achievable user rates. Based on the analytical results, we prove the conjecture by obtaining intuitive scaling laws that show how fast the hardware imperfections can be increased with \( N \). The results are validated numerically in a realistic simulation setup, while the impact on circuit design is considered in [13].

2. SYSTEM MODEL

This paper considers the uplink of a cellular network with \( L \geq 1 \) cells. Each cell consists of \( K \) single-antenna user equipments (UEs) that communicate simultaneously with a BS which is equipped with an array of \( N \) antennas. Our analysis holds for any \( N \) and \( K \), but we are primarily interested in massive MIMO topologies where \( N \gg K \gg 1 \). The channel from UE \( k \) in cell \( l \) to BS \( j \) is denoted as \( h_{jlk} = [h_{jlk}^{(1)} \ldots h_{jlk}^{(N)}]^T \in \mathbb{C}^N \) and is modeled as Rayleigh block fading; thus, it takes a static realization for a coherence block of \( T \) channel uses and independent realizations between blocks. Each channel is circularly symmetric complex Gaussian distributed with zero mean and covariance matrix \( \lambda_{jlk} I_N \); \( h_{jlk} \sim \mathcal{CN}(0, \lambda_{jlk} I_N) \).

\[ \text{†} \text{The assumption of independent fading implies that the array dimensions grow with } N \text{ to keep the inter-antenna distance sufficiently large. However, the analysis in this paper can be easily extended to spatially correlated channels as in [4] and [7], but at the cost of complicating the notation and results.} \]
The average channel attenuation $\lambda_{jk}\kappa > 0$ is different for each combination of BS and UE and depends, for example, on the distance.

The received signal $y_j(t) \in \mathbb{C}^N$ at BS $j$ at a given channel use $t \in \{1, \ldots, T\}$ in the coherence block is conventionally modeled as

$$y_j(t) = \sum_{l=1}^L H_{jl}x_l(t) + n_j(t)$$

where the transmit signal in cell $l$ is $x_l(t) = [x_{l1}(t) \ldots x_{lk}(t)]^T \in \mathbb{C}^K$ and $H_{jl} = [h_{jl1} \ldots h_{jlk}] \in \mathbb{C}^{N \times K}$. The signal $x_{lk}(t)$ sent by UE $k$ in cell $l$ at channel use $t$ is either a deterministic pilot symbol (used for channel estimation) or an information symbol from a Gaussian codebook; in any case, the expectation of the transmit power is bounded as $E[|x_{lk}(t)|^2] \leq P_{lk}$.

The thermal noise vector $n_j(t)$ is modeled as $\mathcal{C}\mathcal{N}(0, \sigma_j^2 I_N)$ independent in time and has variance $\sigma_j^2$. If we block-fade the channels, we will bring insights into the fundamental impact of the parameters $\delta$, $\kappa$, and $\xi$, which characterize the BS hardware imperfections.

### 3. PERFORMANCE ANALYSIS

In this section, we derive achievable user rates for the uplink system in (2) and scaling laws for how quickly the BS hardware imperfections can be increased with $N$ and still achieve non-zero rates.

#### 3.1. Channel Estimation

The achievable rates are computed under the assumption that the first $B \geq K$ channel uses of each coherence block are dedicated for pilot-based channel estimation. UE $k$ in cell $j$ transmits a pre-defined pilot sequence $x_{jk} = [x_{jk}(1) \ldots x_{jk}(B)]^T \in \mathbb{C}^B$.

The pilot sequences are selected arbitrarily under the above mentioned power constraints and our analysis supports any choice. However, it is reasonable to make the sequences $x_{jk}$, $\ldots$, $x_{jk}$ in cell $j$ linearly independent to avoid unnecessary intra-cell interference. Due to the limited coherence block length $T$, inter-cell interference is often unavoidable but the pilot sequences can also be designed and allocated to also reduce inter-cell interference [19].

For any given set of pilot sequences, we now derive estimators of the effective channels $\hat{h}_{jk}(t) = D_{\delta}(\psi_j) h_{jk}$ at any channel use $t > B$ for all $j, l, k$. The conventional multi-antenna channel estimators from [20][22] cannot be applied in this work since the generalized system model in (2) has two non-standard properties: The pilot transmission is corrupted by random phase-drifts and the distortion noise is statistically dependent on the channels. Therefore, we derive a new LMMSE estimator for the system model at hand.

Theorem 1. Let $\psi_j = [y_j(1) \ldots y_j(B)]^T \in \mathbb{CN}B$ denote the received signal at BS $j$ from the pilot transmission. The LMMSE estimate of $h_{jk}(t)$ at any channel use $t \geq B$ for any $l$ and $k$ is

$$\hat{h}_{jk}(t) = (\lambda_{jk} \hat{x}_{jk}(t) D_{\delta}(\psi_j^{-1} \otimes I_N) \psi_j$$

where $D_{\delta}(\Psi) \triangleq \text{diag}(e^{-\delta(t-1)}, e^{-\delta(t-2)}, \ldots, e^{-\delta(t-B)})$.

$$\Psi_j \triangleq \sum_{l=1}^L \sum_{m=1}^K \lambda_{lkm} x_{lm} + \sigma_j^2 \xi_{jk}$$

$\otimes$ is the Kronecker product, and element $(i_1, i_2)$ of $X_{\ell m} \in \mathbb{C}^{B \times B}$ is

$$[X_{\ell m}]_{i_1, i_2} = \begin{cases} |x_{\ell m}(i_1)|^2 (1+c_\ell^2), & i_1 = i_2, \\ x_{\ell m}(i_1) x_{\ell m}(i_2) e^{j \frac{\pi}{2} |i_1-i_2|}, & i_1 \neq i_2. \end{cases}$$

The corresponding error covariance matrix is

$$C_{\ell m} = \mathbb{E} \{ (h_{jk}(t) - \hat{h}_{jk}(t)) (h_{jk}(t) - \hat{h}_{jk}(t))^H \}$$

and the mean-squared error (MSE) becomes $\text{MSE}_{\ell m} = \text{tr}(C_{\ell m})$.

Proof. The general expression for an LMMSE estimator is $\hat{h}_{jk}(t) = \mathbb{E} \{ (h_{jk}(t) \psi_j^H) (\Psi_j^H \psi_j)^{-1} \psi_j \}$ (Chapter 12). The theorem follows from algebraic computation of the two expectations.

Although the channels are block fading, the phase-drifts caused by hardware imperfections make the effective channels $h_{jk}(t)$ change between channel uses. The new LMMSE estimator in Theorem 1 predicts the effective channel for each $t \in \{B + 1, \ldots, T\}$ during the data transmission. Next, we use these predictors to design receive filters and derive the corresponding achievable user rates.
This corollary shows that the distortion noise and receiver noise vanish as $N \to \infty$, while the phase-drifts only has a minor asymptotic impact since the numerator and denominator of the SINR in (13) are scaled symmetrically by $D_{\theta}(t)$. The terms that remain in the denominator depend on the pilot sequences $\tilde{x}_{im}$; hence, these terms are due to pilot contamination (PC) [3], that is, inter-user interference in the estimation phase. Intra-cell PC is typically removed by making the pilot sequences orthogonal in space (i.e., $\tilde{x}_{jk} x_{im} = 0$ for $k \neq m$), which can be achieved by using the columns of a DFT matrix as pilot sequences [25]. Unfortunately, the phase-drifts caused by hardware imperfections break any spatial orthogonality. Therefore, the only way to remove the intra-cell PC is to assign temporally orthogonal sequences within each cell (e.g., $x_{jk}(k) = \sqrt{p_j} \tilde{x}_{jk}$ and $x_{jk}(t) = 0$ for $t \neq k$). Since temporal orthogonality reduces the total pilot power per user, $\| \tilde{x}_{jk} \|^2$, by $1/K$, the simulations in Section 3 reveal that it is only beneficial for extremely large arrays. Inter-cell PC can generally not be removed because there are only $B$ orthogonal sequences in the whole network.

3.3. Scaling Laws on Hardware Imperfections

The asymptotic results in Corollary 1 reveal that the detrimental impact of hardware imperfections vanishes almost completely as $N$ grows large. This conclusion holds for any fixed values of the parameters $\delta$, $\kappa$, and $\xi$. The next corollary shows that it also holds if the parameters are increased with $N$ in a certain way.

Corollary 2. Suppose the hardware imperfection parameters are replaced as $\kappa' = \kappa_0 N^{\tau_1}$, $\xi = \xi_0 N^{\tau_2}$, and $\delta = \delta_0 (1 + \log_2(N^{\tau_3}))$, for some scaling parameters $\tau_1$, $\tau_2$, $\tau_3 \geq 0$ and some initial values $\kappa_0, \xi_0, \delta_0 \geq 0$. If

$$\max(\tau_1, \tau_2, \tau_3) + \frac{\delta_0(t - B)}{2} \leq 1$$

then $\text{SINR}_{jk}(t)$ with MRC converges to a non-zero limit as $N \to \infty$.

Proof. This is achieved by substituting the new parameters into the SINR in (7), multiplying all terms by $1/N^{1 - \tau_3 \delta_0(t - B)}$, and showing that the signal part is non-zero and the denominator is bounded.

The corollary proves that one can increase the hardware imperfections with the number of antennas. This is a very important result for practical deployments, because it indicates that one can make the cost scale with $N$ at a slower pace than linear by using cheaper hardware. This property has been conjectured in overview articles, such as [5], and was proved in [22] using a system model with only additive distortion noise. Corollary 2 shows explicitly that the conjecture holds also for multiplicative phase-drifts and noise amplifications.

Since Corollary 2 is derived for MRC, (14) provides a sufficient scaling condition for any other receive filter that performs better than MRC. The scaling law consists of two terms: $\max(\tau_1, \tau_2)$ and $\frac{\delta_0(t - B)}{2} \tau_3$. The first term $\max(\tau_1, \tau_2)$ shows that the additive distortion noise and noise amplification can be increased simultaneously and independently, while the sum of the two terms manifests a tradeoff between increasing hardware imperfections that cause additive and multiplicative distortions. The system is particularly vulnerable to phase-drifts due to its accumulation, as seen from the second term which increases with $t$ and from that $\delta$ can scale only logarithmically with $N$. We can accept larger variances if the coherence block $T$ is small, which is in line with the results in [10][11].
Fig. 1. The simulation scenario considers 16 square cells with wrap-around to avoid edge effects. Each cell is \(250 \times 250 \text{ m}\) and consists of 8 UEs uniformly distributed in different parts of the cell.

![SIMULATION SCENARIO](image)

Fig. 2. Sum rate for different numbers of antennas, different hardware imperfections, and spatially or temporally orthogonal pilots.

![SUM RATE GRAPH](image)

4. NUMERICAL RESULTS

The analytic results are evaluated in a simulation scenario with 16 cells and wrap-around to avoid edge effects; see Fig. 1. Each square cell is \(250 \times 250 \text{ m}\) and is divided into 8 virtual sectors, where each sector contains one uniformly distributed UE (with minimum distance 35 m). Each sector has an orthogonal pilot sequence, but the same pilot is reused in the corresponding sector of other cells.

The channel attenuations are based on the 3GPP propagation model in [26]: \(\lambda_{ljk} = 10^{d_{ljk}/-3.53} e^{j\Phi_{ljk}}\) where \(d_{ljk}\) is the distance in meters between BS \(j\) and UE \(k\) in cell \(l\) and \(s_{jk} \sim \mathcal{N}(0, 0.25)\) is a realization of the shadow-fading. The transmit powers are \(p_{jk} = -47 \text{ dBm}/\text{Hz}\), the thermal noise power is \(\sigma^2 = -174 \text{ dBm}/\text{Hz}\), \(B = 8\) is the pilot sequence length, and the coherence block is \(T = 500\).

We start by validating the asymptotic behaviors for fixed imperfections. Fig. 2 shows the sum rate of all users (in the 1 km\(^2\) area) as a function of the number of antennas \(N\). The performance is given for ideal hardware and two types of hardware imperfections that are specified in the figure. The simulation shows that the convergence to the upper limit in Corollary 1 is very slow—we used logarithmic scale on the horizontal axis because a million antennas is required for convergence. We observe that the sum rate reduces with hardware imperfections, but the loss is small and vanishes asymptotically.

Two types of pilot sequences are considered in Fig. 2: spatially orthogonal pilots selected from a DFT matrix [25] and temporally orthogonal pilots. As discussed in Section 3.2, spatially orthogonal pilots is a better choice at practical \(N\), although the limit in Corollary 1 might be slightly larger for temporally orthogonal pilots.

5. CONCLUSION

A prerequisite for practical deployment of massive MIMO systems is that each antenna element in the large BS arrays is manufactured using low-cost components, which unfortunately are prone to hardware imperfections. In this work, we derived a scaling law that proves that massive MIMO systems are robust to hardware imperfections. This is a property that has been conjectured in prior works but only proved for simplified channel models with only additive distortion noise. We considered a more accurate uplink model with multiplicative phase-drifts, additive distortion noise, and noise amplifications. We derived an LMMSE channel estimator and the achievable user rates under MRC. Based on this model, our closed-form scaling law manifests how fast the hardware imperfections can increase with \(N\), if non-zero user rates should be achieved. The simulation validates that the rate losses are small as compared to having ideal hardware. The scaling law reveals that the variance of the distortion noise and receiver noise can increase simultaneously as \(\sqrt{N}\), but the scaling should be slower if also the phase-drift variance increases with \(N\) (it can only increase logarithmically). Interestingly, the scaling results hold for other receive filters, such as the approximate MMSE filter.
6. REFERENCES


