On Channel Estimation for Massive MIMO With Pilot Contamination

Amin Khansefid, *Student Member, IEEE*, and Hlaing Minn, *Senior Member, IEEE*

***Abstract*—This letter proposes a simple channel estimator for a multi-cell massive MIMO system with pilot contamination. Under moderate or strong pilot contamination, the existing least square (LS) channel estimator suffers from substantial performance degradation while the existing minimum mean square error (MMSE) channel estimator offers much improved performance over the LS estimator. However, the existing MMSE estimator assumes perfect knowledge of cross-cell large-scale channel coefficients and such an assumption is practically unjustifiable. The proposed estimator overcomes the issues of both existing estimators and provides estimation performance close to that of the ideal MMSE estimator. An analytical mean square error (MSE) expression is also derived for the proposed estimator.**

***Index Terms*—Massive MIMO, channel estimator, pilot contamination, minimum mean square error.**

I. Introduction

Massive MIMO is a promising technology for next generation cellular systems [1] and channel knowledge is critical for its operation. In cellular massive MIMO, frequency reuse gives rise to interference in channel estimation, also known as pilot contamination. Due to pilot overhead cost, time division duplexing is usually considered in the literature [2]. Users transmit pilots in the uplink and then base station (BS) estimates channels based on least square (LS) [3] or minimum mean square error (MMSE) [4]–[6] methods. Inter- and intra-cell large-scale fading coefficients are assumed perfectly known in the existing MMSE method while their estimation is recently presented in [7] together with its asymptotic performance as the number of antennas M approaches ∞. But [7] uses orthogonal pilots and its pilot overhead is very large.

We propose a channel estimator that does not require knowledge of inter-cell large-scale fading coefficients, thus no overhead. The main idea is, instead of obtaining those individual coefficients, to use an effective parameter namely their sum plus a normalized noise variance. The proposed approach estimates that effective parameter and substitutes it back into the MMSE estimator. We also derive analytical MSE expression of the proposed estimator. The results show that the performance of the proposed channel estimator approaches that of the ideal MMSE channel estimator asymptotically ($M \to \infty$).

Notation: Vectors (matrices) are denoted by bold face small (big) letters. The superscripts T and H stand for transpose and conjugate transpose, respectively. $\boldsymbol{I}_K$ is the $K \times K$ identity matrix and $\mathbf{0}_L$ is the $L \times 1$ all zero vector. $\mathbb{E}\{\cdot\}$, $\mathbb{P}\{\cdot\}$, $\|.\|$, $\Re\{\cdot\}$, $\Gamma(\cdot)$, $B(\cdot,\cdot)$, and $\cos^{-1}(.)$ denote expectation, probability, Euclidean norm, real part, Gamma function, beta function, and arc-cosine, respectively. $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \mathbf{C})$ means $\mathbf{n}$ has a probability density function (pdf) of the zero-mean complex Gaussian vector with covariance matrix $\mathbf{C}$, $\mathcal{N}(0, \sigma^2)$ denotes a Gaussian pdf with zero mean and variance σ^2.

Manuscript received March 23, 2015; accepted July 1, 2015. Date of publication July 6, 2015; date of current version September 4, 2015. The associate editor coordinating the review of this paper and approving it for publication was K. J. Kim.

The authors are with the Department of Electrical Engineering, University of Texas at Dallas, Richardson, TX 75080 USA (e-mail: amin.khansefid@utdallas.edu; hlaing.minn@utdallas.edu).

Digital Object Identifier 10.1109/LCOMM.2015.2452912

II. Signal Model

We consider a multi-cell system with L cells where each cell has a BS with co-located M antenna elements and K randomly located single antenna users. We assume frequency-flat fading channels which are quasi-static within a frame and independent across users and antennas. Let g_{ilkm} represent complex gain of the channel from user k in cell l to the antenna m of the BS in cell i (or simply, of BS i). We can write $g_{ilkm} = \sqrt{\beta_{ilk}} h_{ilkm}$ where $\sqrt{\beta_{ilk}}$ (the same for all m) is the large-scale coefficient (for both path loss and log-normal shadowing) and h_{ilkm} is the small-scale coefficient (for multipath Rayleigh fading) with a $\mathcal{CN}(0, 1)$ distribution. The overall $M \times K$ channel matrix is denoted by $\mathbf{G}_{il}$, whose kth column $\mathbf{g}_{ilk} = [g_{ilk1}, \cdots, g_{ilkM}]^T$ represents the gains of the channels from user k in cell l to BS i. The same as in the literature, we treat $\{\beta_{ilk}\}$ as deterministic in the estimation.

A. Uplink Training

The users in uplink transmit τ pilot symbols and then each BS estimates the channels of its users. We assume users of different cells transmit the same set of pilots at the same time (a typical scenario in massive MIMO) and the pilot reuse factor is one. The pilot signals of K users are represented by a $K \times \tau$ matrix $\mathbf{\Phi}^H$ with orthogonality property $\mathbf{\Phi}^H \mathbf{\Phi} = \tau \mathbf{I}_K$, $(K \le \tau)$. The received pilot symbols at BS i are represented by an $M \times \tau$ matrix $\mathbf{Y}_i$ as

$$\mathbf{Y}_i = \sum_{l=1}^{L} \sqrt{q}\mathbf{G}_{il}\mathbf{\Phi}^H + \mathbf{N}_i \tag{1}$$

where q is the uplink pilot power or transmit signal to noise ratio (TX SNR) and $\mathbf{N}_i$ is an $M \times \tau$ noise matrix with independent and identically distributed elements of $\mathcal{CN}(0, 1)$. Let $\boldsymbol{\phi}_k$ denote the kth column of $\mathbf{\Phi}$. Then, for estimation of the channel $\mathbf{g}_{iik}$ at BS i, a sufficient statistic is

$$\mathbf{z}_{ik} = \frac{1}{\sqrt{q}\,\tau}\mathbf{Y}_i\boldsymbol{\phi}_k = \sum_{l=1}^{L} \mathbf{g}_{ilk} + \mathcal{CN}\left(\mathbf{0}, \frac{1}{\tau q}\mathbf{I}_M\right). \tag{2}$$

We have $\mathbf{z}_{ik} \sim \mathcal{CN}(\mathbf{0}_M, \zeta_{ik}\mathbf{I}_M)$ where

$$\zeta_{ik} \triangleq \sum_{l=1}^{L} \beta_{ilk} + \frac{1}{\tau q}. \tag{3}$$

For detection and precoding, BS i needs to know channels of the users in cell i, namely $\{\mathbf{g}_{iik} : \forall k\}$.

B. MMSE Channel Estimator

The Bayesian MMSE estimator requires knowledge of statistics of parameters to be estimated as well as those of noise and interference. Many massive MIMO works acquire channel knowledge based on MMSE estimation [4]–[6], and they assume perfect knowledge of all large-scale fading coefficients, i.e., $\{\beta_{ilk}, 1 \le i, l \le L, 1 \le k \le K\}$ which may not be justifiable in practice. With the perfect knowledge of $\{\beta_{ilk}\}$, the ideal MMSE estimator is given by [8]

$$\hat{\mathbf{g}}_{iik}^{\text{mmse}} = \frac{\beta_{iik}}{\zeta_{ik}} \mathbf{z}_{ik}. \tag{4}$$

We can write $\mathbf{g}_{iik} = \hat{\mathbf{g}}_{iik} + \epsilon_{iik}$, where $\hat{\mathbf{g}}_{iik}$ and estimation error ϵ_{iik} are independent due to the property of MMSE under the Gaussian model. The MSE per antenna of the ideal MMSE channel estimator is given by [8]

$$\eta_{ik}^{\text{mmse}} = \frac{1}{M}\mathbb{E}\left\{\left\|\hat{\mathbf{g}}_{iik}^{\text{mmse}} - \mathbf{g}_{iik}\right\|^2\right\} = \frac{\beta_{iik}(\zeta_{ik} - \beta_{iik})}{\zeta_{ik}}. \tag{5}$$

η_{ik}^{mmse} decreases with increasing q, decreasing β_{iik}, or decreasing β_{ilk} (smaller interference level).

Remark 1: Due to pilot contamination, as $q \to \infty$, $\eta_{ik}^{\text{mmse}} \to \beta_{iik}(1 - \frac{\beta_{iik}}{\sum_{l=1}^{L} \beta_{ilk}})$.

C. LS Channel Estimator

The least square estimator is given by [8]

$$\hat{\mathbf{g}}_{iik}^{\text{ls}} = \mathbf{z}_{ik}. \tag{6}$$

The MSE per antenna of the LS estimator is

$$\eta_{ik}^{\text{ls}} = \frac{1}{M}\mathbb{E}\left\{\left\|\hat{\mathbf{g}}_{iik}^{\text{ls}} - \mathbf{g}_{iik}\right\|^2\right\} = \zeta_{ik} - \beta_{iik}. \tag{7}$$

The LS estimator has a larger MSE than the MMSE but it does not need to know large-scale fading coefficients.

Remark 2: Due to pilot contamination, as $q \to \infty$, $\eta_{ik}^{\text{ls}} \to \sum_{l=1, l\neq i}^{L} \beta_{ilk}$.

III. Proposed Channel Estimator

In practice, acquiring knowledge of inter-cell large-scale fading coefficients may be unjustifiable due to excessive overhead; for example, for L cells with K users in each cell, each BS needs to estimate $(L-1)K$ inter-cell large-scale fading coefficients. A simple but effective solution to this issue is based on the observation that what we need is not individual $\{\beta_{ilk}\}$ as assumed in the existing works, but just ζ_{ik}. Thus, our approach estimates ζ_{ik} and replaces it in the MMSE estimator, yielding the following estimator.

Theorem 1: The proposed channel estimator defined by

$$\hat{\mathbf{g}}_{iik}^{\text{prop}} = M\beta_{iik}\frac{\mathbf{z}_{ik}}{\|\mathbf{z}_{ik}\|^2} \tag{8}$$

approaches the ideal MMSE estimator asymptotically with respect to M. Its MSE defined by $\eta_{ik}^{\text{prop}} \triangleq \frac{1}{M}\mathbb{E}\{\|\hat{\mathbf{g}}_{iik}^{\text{prop}} - \mathbf{g}_{iik}\|^2\}$ is given by

$$\eta_{ik}^{\text{prop}} = \frac{M}{M-1}\frac{\beta_{iik}^2}{\zeta_{ik}} + \beta_{iik} - 2\beta_{iik}\theta_{ik} \tag{9}$$

where $\theta_{ik} = \int_0^1 \int_{-1}^{1} \frac{\kappa_{ik}^2(1-t)+\kappa_{ik}w\sqrt{t(1-t)}}{\kappa_{ik}^2(1-t)+2\kappa_{ik}w\sqrt{t(1-t)}+t} f_T(t) f_W(w)\,dw\,dt$ with $\kappa_{ik} \triangleq \sqrt{\frac{\beta_{iik}}{\zeta_{ik}-\beta_{iik}}}$, and $f_T(t)$ and $f_W(w)$ are given by

$$f_T(t) = \frac{\Gamma(2M)}{(\Gamma(M))^2}\,(t(1-t))^{M-1},\ 0 < t < 1 \tag{10}$$

$$f_W(w) = \frac{M}{\pi}B\left(\frac{1}{2}, M\right)(1-w^2)^{M-\frac{1}{2}},\ |w| < 1. \tag{11}$$

Proof: The minimum variance unbiased estimator (MVUE) of ζ_{ik} given the observed signal $\mathbf{z}_{ik}$ reads as [8]

$$\hat{\zeta}_{ik} = \frac{\|\mathbf{z}_{ik}\|^2}{M}. \tag{12}$$

We can write $\zeta_{ik} = \hat{\zeta}_{ik} + e_{ik}$. Since $\mathbf{z}_{ik} \sim \mathcal{CN}(\mathbf{0}_M, \zeta_{ik}\mathbf{I}_M)$, $\hat{\zeta}_{ik}$ has a central Chi-square distribution with $2M$ degrees of freedom (DoF), $\mathbb{E}\{\hat{\zeta}_{ik}\} = \zeta_{ik}$, and $\text{var}\{\hat{\zeta}_{ik}\} = \zeta_{ik}^2/M$. Under the MVUE framework where the parameter to be estimated is treated to be non-random, estimation error e_{ik} has zero mean and its variance is the same as the variance of $\hat{\zeta}_{ik}$ given by

$$\text{var}\{e_{ik}\} = \frac{\zeta_{ik}^2}{M}. \tag{13}$$

Replacing ζ_{ik} in (4) with $\hat{\zeta}_{ik}$ yields the proposed estimator. As $\hat{\zeta}_{ik}$ approaches ζ_{ik} asymptotically, the proposed estimator approaches the ideal MMSE estimator asymptotically with respect to M. The proof of MSE is given in Appendix A. □

Remark 3: The average normalized squared Euclidean distance between $\hat{\mathbf{g}}_{iik}^{\text{prop}}$ and $\hat{\mathbf{g}}_{iik}^{\text{mmse}}$ is given by

$$\frac{1}{M}\mathbb{E}\left\{\|\hat{\mathbf{g}}_{iik}^{\text{prop}} - \hat{\mathbf{g}}_{iik}^{\text{mmse}}\|^2\right\} = \frac{1}{M-1}.\frac{\beta_{iik}^2}{\zeta_{ik}}. \tag{14}$$

The proof of (14) is given in Appendix B. From (3) and (14), we observe the average distance decreases with increasing M, decreasing q, increasing β_{ilk}, $i \neq l$, and decreasing β_{iik}.

IV. Numerical Results and Discussion

In this section we compare the performance of the proposed channel estimator with the ideal MMSE estimator and LS estimator. We use a typical multicell structure with $L = 7$ cells, $K = 10$ users in each cell, frequency reuse factor of 1, and $\tau = K$ pilot symbols. We consider both setups of fixed and random values for $\{\beta_{ilk}\}$. For the fixed case, we set $\beta_{llk} = 1$ and $\beta_{ilk} = a$, $\forall\, l \neq i$. For the random case, users in each cell are uniformly located in the disk around BS within radii $d_0 = 100$ and $d_1 = 1000$ m; $\{\beta_{ilk}\}$ are independently generated by $\beta_{ilk} = \psi/\left(\frac{d_{ilk}}{d_0}\right)^{\nu}$ where $\nu = 3.8$, $10\log_{10}(\psi) \sim \mathcal{N}(0, \sigma_{\text{shadow,dB}}^2)$ with $\sigma_{\text{shadow,dB}} = 8$, and d_{ilk} is the distance of user k in cell l to BS i.

Fig. 1 shows the channel estimation MSE versus TX SNR for $a = 0.05$ and $M = 30$. Analytical and simulation MSEs match for all the methods and for simplicity we include the simulation result of the proposed method only. With the increase of the pilot power, MSEs of all the methods decrease but the MSE gap between the ideal MMSE estimator and the proposed estimator increases (see Remark 3). There are MSE floors for all three estimators due to pilot contamination (see remarks 1 and 2). At low TX SNR, MSE of the proposed estimator is very close to

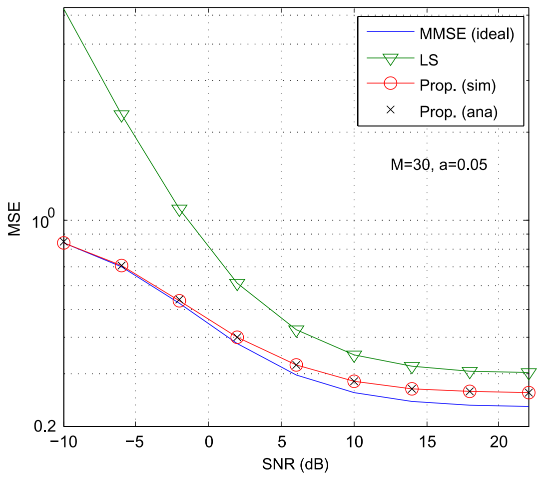

Fig. 1. Channel estimation MSE versus TX SNR.

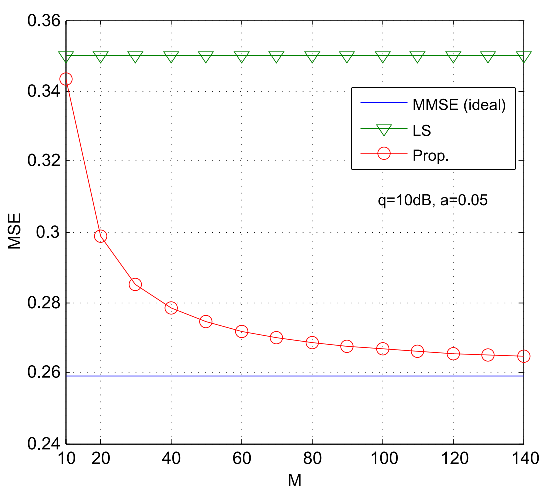

Fig. 2. Channel estimation MSE versus the number of antennas.

that of the ideal MMSE estimator while the MSE gap between LS and ideal MMSE are large.

In Fig. 2, we compare MSE versus the number of BS antennas M under the setting of $a = 0.05$ and TX SNR $q = 10$ dB. With increase of M, the MSE of the proposed estimator approaches that of the ideal MMSE, while the MSE of LS estimator does not change.

In Fig. 3, we show MSE performance with respect to the various levels of cross-cell interference (characterized by a) with $q = 10$ dB under two settings of $M = 30$ and $M = 90$. At a very low interference level, LS has slightly better MSE than the proposed method at a small M but not at a large M. As the interference level increases, the proposed method outperforms LS substantially and approaches the ideal MMSE performance (see Remark 3).

In Fig. 4, we evaluate MSE performance under random large-scale fading coefficients $\{\beta_{ilk}\}$ with $M = 30$. The results are obtained by averaging MSEs over 1000 realizations of $\{\beta_{ilk}\}$. Analytical MSEs match with the simulation MSEs. Relative performances of different estimators are similar to the fixed $\{\beta_{ilk}\}$ case. In addition, we investigate the sensitivity of the proposed estimator against inaccuracy of β_{iik} by using an estimate $\hat{\beta}_{iik} = \beta_{iik}(1 + \mathcal{N}(0, \sigma^2))$. The performance degradation

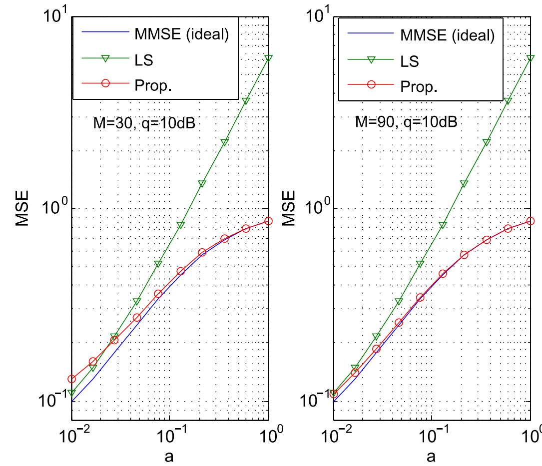

Fig. 3. Channel estimation MSE versus cross-cell interference level.

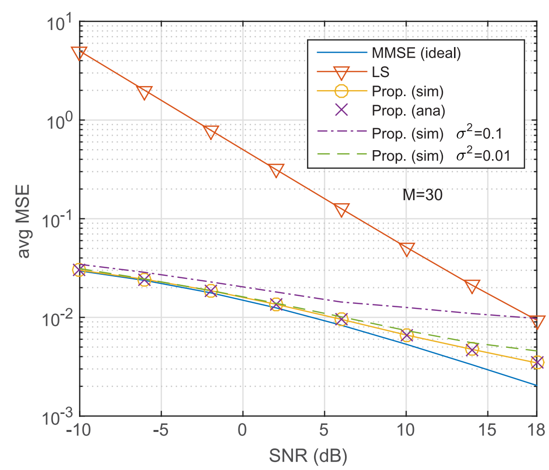

Fig. 4. Average channel estimation MSE under random $\{\beta_{ilk}\}$.

for $\sigma^2 = 0.1$ is noticeable at high SNR but that for $\sigma^2 = 0.01$ is insignificant. The proposed estimator still outperforms the LS estimator significantly.

V. Conclusion

We have developed a simple channel estimator for a multi-cell massive MIMO system with pilot contamination. The proposed estimator replaces the combined interference plus noise power term in the ideal MMSE estimator with its simple estimate, and yields MSE performance very close to that of the ideal MMSE estimator without requiring any additional overhead. We have also derived an analytical MSE expression of the proposed estimator which can be useful in system design and performance evaluation.

Appendix A

For the proof of MSE, we need the following Lemmas.

Lemma 1: If $\mathbf{x} \sim \mathcal{CN}(\mathbf{0}_M, \sigma_x^2 \mathbf{I}_M)$ and $\mathbf{y} \sim \mathcal{CN}(\mathbf{0}_M, \sigma_y^2 \mathbf{I}_M)$ are independent and $\frac{\mathbf{x}^H}{\|\mathbf{x}\|}\frac{\mathbf{y}}{\|\mathbf{y}\|} \triangleq Re^{j\theta}$, then θ is uniformly distributed in $[-\pi, \pi]$, and the pdf of R is given by

$$f_R(r) = 2Mr(1 - r^2)^{M-1}, \; 0 \le r \le 1. \tag{15}$$

Proof: Circular symmetry of $\mathbf{x}$ and $\mathbf{y}$ yields uniform distribution of θ. Next, $Z = \left|\frac{\mathbf{x}^H}{\|\mathbf{x}\|}\frac{\mathbf{y}}{\|\mathbf{y}\|}\right|^2$ has a beta pdf $f_Z(z) = M(1-z)^{(M-1)}$, $z \in [0,1]$ [9]. As $R = \sqrt{Z}$, transformation of random variable (RV) gives (15). □

Lemma 2: If R is with pdf given in (15), θ is uniformly distributed in $[-\pi, \pi]$ and they are independent, then the pdf of $W \triangleq R\cos(\theta)$ is given by (11).

Proof: From the uniform distribution of θ, we obtain

$$\mathbb{P}\left(\cos(\theta) \le \frac{w}{r}\right) = \begin{cases} 0 & \frac{w}{r} < -1 \\ 1 - \frac{1}{\pi}\cos^{-1}\left(\frac{w}{r}\right) & -1 \le \frac{w}{r} \le 1 \\ 1 & \frac{w}{r} > 1 \end{cases} \quad (16)$$

Next, the cumulative distribution function (cdf) of W is $F_W(w) = \mathbb{P}(R\cos(\theta) \le w) = \int_0^1 \mathbb{P}\left(\cos(\theta) \le \frac{w}{r}\right) f_R(r)dr$:

$$F_W(w) = \begin{cases} 0 & w < -1 \\ \mathcal{I}(w) & -1 \le w \le 0 \\ \mathcal{I}(w) + \int_0^w f_R(r)dr & 0 \le w \le 1 \\ 1 & w > 1 \end{cases} \quad (17)$$

where $\mathcal{I}(w) \triangleq \int_{|w|}^1 (1 - \frac{1}{\pi}\cos^{-1}(\frac{w}{r})) f_R(r)dr$. Then, we find $f_W(w) = \frac{dF_W(w)}{dw} = \int_{|w|}^1 \frac{2Mr(1-r^2)^{M-1}}{\pi\sqrt{r^2-w^2}}dr$, $|w| < 1$. Next, by change of variable as $z = r^2 - w^2$ and using [10, 3.191-1], we arrive at (11). □

Lemma 3: If $\mathbf{x} \sim \mathcal{CN}(\mathbf{0}_M, \mathbf{I}_M)$ and $\mathbf{y} \sim \mathcal{CN}(\mathbf{0}_M, \mathbf{I}_M)$ are independent, then $U = \frac{\|\mathbf{x}\|}{\|\mathbf{y}\|}$ has pdf given by

$$f_U(u) = \frac{2\Gamma(2M)}{(\Gamma(M))^2}\frac{u^{2M-1}}{(u^2+1)^{2M}}, u > 0. \quad (18)$$

Proof: We recall $\|\mathbf{x}\|^2$ and $\|\mathbf{y}\|^2$ are central Chi-square with the same pdf $f(v) = \frac{v^{M-1}}{\Gamma(M)}e^{-v}$. Then using the independence of $\|\mathbf{x}\|^2$ and $\|\mathbf{y}\|^2$, we find the cdf of U^2 and then by differentiation we find the pdf of U^2 as $f(z) = \frac{\Gamma(2M)}{(\Gamma(M))^2}\frac{z^{M-1}}{(z+1)^{2M}}$, $z > 0$. Finally by the square root transformation of U^2, we arrive at (18). □

Lemma 4: If $\mathbf{x} \sim \mathcal{CN}(\mathbf{0}_M, \sigma_x^2\mathbf{I}_M)$ and $\mathbf{y} \sim \mathcal{CN}(\mathbf{0}_M, \sigma_y^2\mathbf{I}_M)$ are independent, then

$$\mathbb{E}\left\{\frac{(\mathbf{x}+\mathbf{y})^H\mathbf{x}}{\|\mathbf{x}+\mathbf{y}\|^2}\right\} = \int_0^\infty \int_{-1}^1 \frac{(\kappa u + w)f_U(u)f_W(w)}{\kappa u + \frac{1}{\kappa u} + 2w}dwdu \quad (19)$$

where $\kappa \triangleq \sqrt{\frac{\sigma_x^2}{\sigma_y^2}}$, and $f_W(w)$ and $f_U(u)$ are given in (11) and (18), respectively.

Proof: By expanding and dividing numerator and denominator of the left hand side of (19) by $\|\mathbf{x}\|\|\mathbf{y}\|$, we can write

$$\mathbb{E}\left\{\frac{(\mathbf{x}+\mathbf{y})^H\mathbf{x}}{\|\mathbf{x}+\mathbf{y}\|^2}\right\} = \mathbb{E}\left\{\frac{\kappa U + Re^{-j\theta}}{\kappa U + \frac{1}{\kappa U} + 2W}\right\} \quad (20)$$

where $\frac{\|\mathbf{x}\|}{\|\mathbf{y}\|} \triangleq \kappa U$, $\frac{\mathbf{x}^H}{\|\mathbf{x}\|}\frac{\mathbf{y}}{\|\mathbf{y}\|} \triangleq Re^{j\theta}$, and $W \triangleq R\cos(\theta)$, and U, R and θ are independent. First, by conditioning on U and R, and averaging over θ, it can be seen the expected value of imaginary part of the left hand side of (20) is zero. So, (20) equals $\mathbb{E}\{\frac{\kappa U + W}{\kappa U + \frac{1}{\kappa U} + 2W}\}$. Next, from Lemmas 2 and 3, the pdfs of independent RVs U and W are given in (18) and (11) respectively. Then direct calculation of the expectation over those pdfs yields the result in (19). □

For proof of MSE, we expand $\eta_{ik}^{\text{prop}} = \frac{1}{M}\mathbb{E}\{\|\hat{\mathbf{g}}_{iik}^{\text{prop}}\|^2\} + \frac{1}{M}\mathbb{E}\{\|\mathbf{g}_{iik}\|^2\} - \frac{2}{M}\mathbb{E}\{\Re\{(\hat{\mathbf{g}}_{iik}^{\text{prop}})^H\mathbf{g}_{iik}\}\}$ and find these three expected values. From (8), $\|\hat{\mathbf{g}}_{iik}^{\text{prop}}\|^2 = M^2\beta_{iik}^2\frac{1}{\|\mathbf{z}_{ik}\|^2}$, and using the fact that $\frac{1}{\|\mathbf{z}_{ik}\|^2}$ has an inverse gamma distribution with mean $\mathbb{E}\left\{\frac{1}{\|\mathbf{z}_{ik}\|^2}\right\} = \frac{1}{(M-1)\zeta_{ik}}$, we have $\frac{1}{M}\mathbb{E}\{\|\hat{\mathbf{g}}_{iik}^{\text{prop}}\|^2\} = \frac{M}{M-1}\frac{\beta_{iik}^2}{\zeta_{ik}}$. Then we recall $\|\mathbf{g}_{iik}\|^2$ has a central Chi-square distribution and $\frac{1}{M}\mathbb{E}\{\|\mathbf{g}_{iik}\|^2\} = \beta_{iik}$. Finally, to find the expected value of the third term, we recall $\hat{\mathbf{g}}_{iik}^{\text{prop}} = M\beta_{iik}\frac{\mathbf{z}_{ik}}{\|\mathbf{z}_{ik}\|^2}$ and use Lemma 4 with $\mathbf{x} = \mathbf{g}_{iik}$, $\mathbf{y} = \mathcal{CN}(\mathbf{0}_M, (\zeta_{ik} - \beta_{iik})\mathbf{I}_M)$. After that to make the boundary of integration finite, we use change of variable in the integration as $t = \frac{1}{1+u^2}$ and we reformulate this integral as θ_{ik} as given in Theorem 1. After finding these three expectations, by substituting them back in the expansion of η_{ik}^{prop}, we complete the proof.

Appendix B

Here, we present proof for (14). We expand $\frac{1}{M}\mathbb{E}\{\|\hat{\mathbf{g}}_{iik}^{\text{prop}} - \hat{\mathbf{g}}_{iik}^{\text{mmse}}\|^2\} = \frac{1}{M}\mathbb{E}\{\|\hat{\mathbf{g}}_{iik}^{\text{mmse}}\|^2\} + \frac{1}{M}\mathbb{E}\{\|\hat{\mathbf{g}}_{iik}^{\text{prop}}\|^2\} - \frac{2}{M}\mathbb{E}\{\Re\{(\hat{\mathbf{g}}_{iik}^{\text{mmse}})^H\hat{\mathbf{g}}_{iik}^{\text{prop}}\}\}$. Then we compute these different expectations. First by recalling $\hat{\mathbf{g}}_{iik}^{\text{mmse}} \sim \mathcal{CN}(\mathbf{0}_M, \frac{\beta_{iik}^2}{\zeta_{ik}}\mathbf{I}_M)$, and using the fact that $\|\hat{\mathbf{g}}_{iik}^{\text{mmse}}\|^2$ has a central Chi-square distribution, we have $\frac{1}{M}\mathbb{E}\{\|\hat{\mathbf{g}}_{iik}^{\text{mmse}}\|^2\} = \frac{\beta_{iik}^2}{\zeta_{ik}}$. Next, from Appendix A, $\frac{1}{M}\mathbb{E}\{\|\hat{\mathbf{g}}_{iik}^{\text{prop}}\|^2\} = \frac{M}{M-1}\frac{\beta_{iik}^2}{\zeta_{iik}}$. For the last term, by substituting $\hat{\mathbf{g}}_{iik}^{\text{mmse}}$ and $\hat{\mathbf{g}}_{iik}^{\text{prop}}$ from (4) and (8) respectively, we have $(\hat{\mathbf{g}}_{iik}^{\text{mmse}})^H\hat{\mathbf{g}}_{iik}^{\text{prop}} = M\frac{\beta_{iik}^2}{\zeta_{ik}}$. By substituting these quantities back in the expansion, we arrive at (14).

References

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

[2] L. Lu, G. Li, A. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: Benefits and challenges," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 742–758, Oct. 2014.

[3] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.

[4] 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.

[5] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?" *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 160–171, Feb. 2013.

[6] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "The multicell multiuser MIMO uplink with very large antenna arrays and a finite-dimensional channel," *IEEE Trans. Commun.*, vol. 61, no. 6, pp. 2350–2361, Jun. 2013.

[7] A. Ashikhmin, T. L. Marzetta, and L. Li, "Interference reduction in multi-cell massive MIMO systems I: Large-scale fading precoding and decoding," to be published.

[8] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory*, New York, NY, USA: Pearson Education, 1993.

[9] N. Jindal, "MIMO broadcast channels with finite-rate feedback," *IEEE Trans. Inf. Theory*, vol. 52, no. 11, pp. 5045–5060, Nov. 2006.

[10] A. Jeffrey and D. Zwillinger, *Table of Integrals, Series, and Products*. New York, NY, USA: Academic, 2007.