

# Throughput Scaling of MIMO Channels with Imperfect CSIT in the Low SNR Regime

Nadia Jamal, *Student Member, IEEE*, Patrick Mitran, *Senior Member, IEEE*

**Abstract**—We study the effect of channel estimation error on the performance of water-filling in point-to-point MIMO channels at low signal to noise ratios (SNR). In this regard, we derive the water-filling throughput of a MIMO channel in the presence of imperfect channel state information at the transmitter (CSIT). The asymptotic growth rate for the throughput,  $R$ , is then found and is compared with the asymptotic growth rate for the capacity with perfect CSIT,  $C^P$ , as a function of the signal to estimation error ratio (SER). We show that at low SNR, for moderate values of the SER, water-filling based on *erroneous channel estimates* can still achieve significant throughputs asymptotically.

**Index Terms**—Capacity scaling, Channel estimation error, multiple-input multiple-output (MIMO), Massive MIMO.

## I. INTRODUCTION

Multiple-input multiple-output (MIMO) technology delivers promising improvement to the data rate of wireless systems as it enables simultaneous transmission of multiple streams of data over independent paths [1]-[4]. MIMO channels with a large number of antennas (also known as Massive MIMO) and their capacity scaling results have been under extensive investigation recently [5]-[7].

In [1, Section IV.A.], having  $n$  transmit and receive antennas in a rich scattering environment, the capacity of a point-to-point MIMO channel with perfect channel state information at the transmitter (CSIT) and the mutual information with no CSIT at all were found using water-filling power allocation and equal power allocation respectively and were shown to scale linearly with  $n$ . The constant multiplier associated with linear scaling (the asymptotic growth rate) were derived by using the limiting distribution of the eigenvalues of the channel gain matrix as  $n$  grows large. It was shown in [1, Section IV.A.] that at low signal to noise ratios (SNR), water-filling with perfect CSIT provides significant throughput improvement asymptotically over equal power allocation.

In practice, perfect CSIT is not generally available and in addition, the transmitter may have no a priori reliable knowledge of the quality of channel estimation either. Thus, it is interesting to evaluate the performance loss in point-to-point MIMO channels if the transmitter performs water-filling based on erroneous channel estimates only and does not take the quality of channel estimation into account. Hence, we derive such water-filling throughput, show that it scales linearly with  $n$  asymptotically at low SNR, and find its asymptotic growth rate,  $R$ . Using the asymptotic growth rate for the capacity with perfect CSIT,  $C^P$ , as a baseline for comparison, we then compare  $R$  with  $C^P$  as a function of the signal to estimation error ratio (SER).

Our results indicate that at low SNR, for moderate values of the SER, water-filling based on *erroneous channel estimates* can still achieve significant throughputs asymptotically. In particular, for SER values such as 5 dB, 0 dB, and  $-5$  dB,  $R$  is found to be 86%, 70%, and 52% of  $C^P$  respectively.

Note that, in the high SNR regime, as shown in [1], equal power allocation (which requires no CSIT) provides the same asymptotic throughput as water-filling with perfect CSIT. Hence, in that regime, equal power allocation can be employed and thus studying the effect of channel estimation error on the performance of water-filling at high SNR is of no interest.

The rest of this letter is organized as follows: Section II introduces our channel model and Section III provides capacity analysis of no CSIT and perfect CSIT scenarios. The water-filling throughput with imperfect CSIT and its asymptotic growth rate,  $R$ , are derived in sections IV and V respectively. Section VI provides numerical results and compares  $R$  with the asymptotic growth rate for the capacity with perfect CSIT. Finally, Section VII concludes the letter.

## II. CHANNEL MODEL

Consider a point-to-point MIMO channel between a transmitter and a receiver each equipped with  $n$  antennas. This channel has a discrete-time model  $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w}$ , where  $\mathbf{x}$  and  $\mathbf{y}$  are  $n \times 1$  transmit and receive vectors respectively, and  $\mathbf{w}$  is the  $n \times 1$  additive white Gaussian noise vector. We assume that the entries of  $\mathbf{w}$  are i.i.d. and zero-mean circularly symmetric complex Gaussian (ZMCSCG) with variance  $N$  and for simplicity, we take  $N = 1$ . Furthermore,  $\mathbf{H}$  is the channel gain matrix whose entries are i.i.d. and ZMCSCG with variance  $2\sigma^2$ , i.e.,  $\sigma^2$  per real dimension. Throughout the letter, we assume that the channel is quasi-static fading, i.e., as far as the transmitter is concerned,  $\mathbf{H}$  is fixed for the duration of the transmission. In addition, denoting the total transmit power over  $n$  antennas as  $P$ , we measure the SNR as  $\rho = P/N = P$ .

## III. ANALYSIS OF NO CSIT AND PERFECT CSIT CASES

### A. No CSIT

In this scenario, since the transmitter has no knowledge of  $\mathbf{H}$ , we assume that it distributes the total transmit power  $P$  equally over  $n$  transmit antennas. Therefore, assuming  $\{\lambda_1, \lambda_2, \dots, \lambda_n\}$  as the eigenvalues of  $\mathbf{H}\mathbf{H}^\dagger$  (where  $(\cdot)^\dagger$  refers to conjugate transpose), the mutual information with equal power allocation is <sup>1</sup> [1, Section III.B.]

$$I_n^N = \sum_{i=1}^n \log(1 + P\lambda_i/n). \quad (1)$$

<sup>1</sup>Throughout this letter, logarithms are with respect to base 2.

The authors are with the Department of Electrical and Computer Engineering at the University of Waterloo, Ontario, Canada, (e-mail: n2jamal, pmitran@uwaterloo.ca).

### B. Perfect CSIT (best-case scenario)

In this scenario, since  $\mathbf{H}$  is perfectly known at the transmitter, water-filling power allocation is optimal [4], [1, Section III.A.]. Thus, the capacity of the MIMO channel with perfect CSIT is

$$C_n^P = \sum_{i=1}^n (\log(\lambda_i \mu))^+, \quad (2)$$

where  $(x)^+ = \max(0, x)$  and the water-filling level  $\mu$  satisfies

$$\sum_{i=1}^n (\mu - 1/\lambda_i)^+ = P. \quad (3)$$

### IV. WATER-FILLING THROUGHPUT WITH IMPERFECT CSIT

A more realistic scenario is when the channel gains are not perfectly known at the transmitter. This happens due to channel estimation errors which are nonzero. With Maximum Likelihood Estimation (MLE), the estimated channel gain matrix  $\hat{\mathbf{H}}$  is expressed as

$$\hat{\mathbf{H}} = \mathbf{H} + \mathbf{E}, \quad (4)$$

where  $\mathbf{E}$  is the channel estimation error matrix that has i.i.d. and ZMCSCG entries with variance  $2\sigma_e^2$  (Gaussian-distributed model for the estimated channel gains is reasonable in estimation methods such as MLE [8]). Furthermore,  $\mathbf{E}$  and  $\mathbf{H}$  are independent. Throughout the letter, we refer to the ratio  $\zeta = \sigma^2/\sigma_e^2$  as the signal to estimation error ratio (SER).

From (4), the entries of  $\hat{\mathbf{H}}$  are i.i.d. and ZMCSCG with variance  $2\sigma^2 + 2\sigma_e^2$ . Thus, the entries of  $\hat{\mathbf{H}}$  and  $\mathbf{H}$  are joint complex Gaussian distributed and the channel gain matrix  $\mathbf{H}$  can be written as [9]

$$\mathbf{H} = \eta \hat{\mathbf{H}} + \mathbf{X}, \quad (5)$$

where  $\eta = \zeta/(1+\zeta)$ ,  $\hat{\mathbf{H}}$  and  $\mathbf{X}$  are independent, and the entries of  $\mathbf{X}$  are i.i.d. and ZMCSCG with variance  $2\sigma^2/(1+\zeta)$ .

When only the channel estimate  $\hat{\mathbf{H}}$  is known at the transmitter, water-filling power allocation provides the throughput

$$R_n = \log \det [\mathbf{I}_n + \mathbf{H}\mathbf{Q}\mathbf{H}^\dagger], \quad (6)$$

in which  $\mathbf{I}_n$  denotes the  $n \times n$  identity matrix,  $\mathbf{Q} = \mathbf{V}\tilde{\mathbf{Q}}\mathbf{V}^\dagger$  is the transmit covariance matrix where  $\mathbf{V}$  is obtained from the singular value decomposition (SVD)  $\hat{\mathbf{H}} = \mathbf{U}\hat{\Sigma}^{1/2}\mathbf{V}^\dagger$  with  $\hat{\Sigma} = \text{diag}(\hat{\lambda}_1, \hat{\lambda}_2, \dots, \hat{\lambda}_n)$ , and  $\tilde{\mathbf{Q}}$  is an  $n \times n$  diagonal matrix with the  $i$ th diagonal entry

$$\tilde{Q}_i = (\hat{\mu} - 1/\hat{\lambda}_i)^+, \quad (7)$$

as the power allocated to the  $i$ th stream [4]. Furthermore, the water-filling level  $\hat{\mu}$  in (7) satisfies

$$\sum_{i=1}^n (\hat{\mu} - 1/\hat{\lambda}_i)^+ = P. \quad (8)$$

Using (5) and replacing  $\mathbf{Q} = \mathbf{V}\tilde{\mathbf{Q}}\mathbf{V}^\dagger$  and  $\hat{\mathbf{H}} = \mathbf{U}\hat{\Sigma}^{1/2}\mathbf{V}^\dagger$  in (6), we find

$$\begin{aligned} R_n &= \log \det \left[ \mathbf{I}_n + (\eta \hat{\mathbf{H}} + \mathbf{X}) \mathbf{Q} (\eta \hat{\mathbf{H}} + \mathbf{X})^\dagger \right] \\ &= \log \det \left[ \mathbf{I}_n + \eta^2 \hat{\Sigma} \tilde{\mathbf{Q}} + \eta \mathbf{K} \tilde{\mathbf{Q}} \hat{\Sigma}^{1/2} \right. \\ &\quad \left. + \eta \hat{\Sigma}^{1/2} \tilde{\mathbf{Q}} \mathbf{K}^\dagger + \mathbf{K} \tilde{\mathbf{Q}} \mathbf{K}^\dagger \right], \end{aligned} \quad (9)$$

where  $\mathbf{K} = \mathbf{U}^\dagger \mathbf{X} \mathbf{V}$ . Equation (9) can be written as

$$R_n = \log \det \mathbf{Y} + \log \det \left[ \mathbf{I}_n + \mathbf{Y}^{-1/2} (\mathbf{L} + \mathbf{T}) \mathbf{Y}^{-1/2} \right], \quad (10)$$

where

$$\mathbf{Y} = \mathbf{I}_n + \eta^2 \hat{\Sigma} \tilde{\mathbf{Q}}, \quad (11)$$

$$\mathbf{L} = \eta \mathbf{K} \tilde{\mathbf{Q}} \hat{\Sigma}^{1/2} + \eta \hat{\Sigma}^{1/2} \tilde{\mathbf{Q}} \mathbf{K}^\dagger, \quad (12)$$

$$\mathbf{T} = \mathbf{K} \tilde{\mathbf{Q}} \mathbf{K}^\dagger. \quad (13)$$

### V. ASYMPTOTIC GROWTH RATES

The throughputs  $I_n^N$ ,  $C_n^P$ , and  $R_n$  presented in sections III and IV, are random variables. In [1, Section IV.A], the almost-sure convergence of scaled throughputs  $I_n^N/n$  and  $C_n^P/n$  is shown as  $n \rightarrow \infty$ . Furthermore, the asymptotic growth rates  $I^N$  and  $C^P$ , for no CSIT and perfect CSIT respectively, were found in [1] for the special case that the entries of  $\mathbf{H}$  have unit variance ( $2\sigma^2 = 1$ ).

In this section, we first state the result of Theorem IV.1 in [1] for arbitrary  $\sigma$  and  $\sigma_e$ . Subsequently, we show the almost-sure convergence of the scaled water-filling throughput with imperfect CSIT,  $R_n/n$ , as  $n \rightarrow \infty$  at low SNR and derive the asymptotic growth rate for the throughput,  $R$ .

Let  $\mathbf{G}$  be an  $n \times n$  matrix that has i.i.d. and ZMCSCG entries with unit variance. According to Theorem IV.1 in [1], the scaled (by  $1/n$ ) empirical eigenvalue distribution of  $\mathbf{G}\mathbf{G}^\dagger$ , converges almost surely to a limit which has the density

$$g(\lambda) = \begin{cases} \pi^{-1} \sqrt{1/\lambda - 1/4} & 0 \leq \lambda \leq 4 \\ 0 & \text{else.} \end{cases} \quad (14)$$

Since  $\mathbf{H}$  is equal in distribution with  $\sqrt{2}\sigma\mathbf{G}$ , the eigenvalues of  $\mathbf{H}\mathbf{H}^\dagger$  are scaled versions (by  $2\sigma^2$ ) of the eigenvalues of  $\mathbf{G}\mathbf{G}^\dagger$ . Thus, the scaled empirical eigenvalue distribution of  $\mathbf{H}\mathbf{H}^\dagger$  converges to a limit with density

$$h(\lambda) = \begin{cases} (\sqrt{2}\sigma\pi)^{-1} \sqrt{1/\lambda - 1/8\sigma^2} & 0 \leq \lambda \leq 8\sigma^2 \\ 0 & \text{else.} \end{cases} \quad (15)$$

Similarly, the scaled empirical eigenvalue distribution of  $\hat{\mathbf{H}}\hat{\mathbf{H}}^\dagger$  converges to a limit with density

$$\hat{h}(\lambda) = \begin{cases} \frac{1}{\sqrt{2(\sigma^2 + \sigma_e^2)}\pi} \sqrt{\frac{1}{\lambda} - \frac{1}{8(\sigma^2 + \sigma_e^2)}} & 0 \leq \lambda \leq 8(\sigma^2 + \sigma_e^2) \\ 0 & \text{else.} \end{cases} \quad (16)$$

#### A. No CSIT

Since the empirical distribution of  $\lambda_i/n$  in (1) converges to a limit with density as in (15), as  $n \rightarrow \infty$ , using the law of large numbers (LLN), we have [1, Section IV.A.]

$$\frac{I_n^N}{n} = \frac{1}{n} \sum_{i=1}^n \log(1 + P\lambda_i/n) \rightarrow I^N, \quad (17)$$

where the convergence is almost sure [1]. Denoting  $\mathbf{E}[\cdot]$  as the expectation of a random variable,

$$I^N = \mathbf{E}[\log(1 + P\lambda)] = \int_0^{8\sigma^2} \log(1 + P\lambda) h(\lambda) d\lambda. \quad (18)$$

As  $P \rightarrow 0$ , following the same approach as in [1], a first-order approximation gives<sup>2</sup>

$$I^N \approx P \log e \int_0^{8\sigma^2} \lambda h(\lambda) d\lambda = 2P\sigma^2 \log e. \quad (19)$$

<sup>2</sup>Note that the result in equation (12) of [1] is obtained in nats as opposed to (19) which is in bits/sec/Hz.

### B. Perfect CSIT (best-case scenario)

As in [1, Section IV.A.], by relabeling  $\mu$  as  $\mu_n/n$  in (2) and (3) respectively, and using the LLN along with (15), we find

$$\frac{C^n}{n} = \frac{1}{n} \sum_{i=1}^n (\log(\mu_n \lambda_i/n))^+ \rightarrow C^P = \int_0^{8\sigma^2} (\log(\mu^* \lambda))^+ h(\lambda) d\lambda, \quad (20)$$

and

$$\frac{1}{n} \sum_{i=1}^n (\mu_n - (\lambda_i/n)^{-1})^+ \rightarrow \int_0^{8\sigma^2} (\mu^* - 1/\lambda)^+ h(\lambda) d\lambda = P, \quad (21)$$

as  $n \rightarrow \infty$  where  $\mu^* = \lim_{n \rightarrow \infty} \mu_n$  is the asymptotic scaled water-filling level and the convergence is almost sure. In (20) and (21), following the same approach as in [1], we can find  $dC^P/dP = \log e/\mu^*$  and  $\mu^* \rightarrow 1/8\sigma^2$  as  $P \rightarrow 0$  respectively, and thus obtain the first-order approximation<sup>3</sup>

$$C^P \approx 8\sigma^2 P \log e. \quad (22)$$

Hence, as previously shown in [1], at low SNR, availability of perfect CSIT provides significant performance improvement for the MIMO channel asymptotically compared to the no CSIT scenario since  $C^P/I^N \approx 4$ .

### C. Imperfect CSIT

By relabeling  $\hat{\mu}$  as  $\hat{\mu}_n/n$  in (8), and using the LLN along with (16), as  $n \rightarrow \infty$ , we obtain

$$\frac{1}{n} \sum_{i=1}^n (\hat{\mu}_n - (\hat{\lambda}_i/n)^{-1})^+ \rightarrow \int_0^{8(\sigma^2 + \sigma_e^2)} (\hat{\mu}^* - 1/\lambda)^+ \hat{h}(\lambda) d\lambda = P, \quad (23)$$

where the convergence is almost sure. Note that  $\hat{\mu}^*$  in (23) is the asymptotic scaled water-filling level.

Since  $\tilde{\Sigma}$  and  $\tilde{\mathbf{Q}}$  are diagonal with the  $i$ th diagonal entry respectively as  $\hat{\lambda}_i$  and (7), the matrix  $\mathbf{Y}$  in (11) is diagonal and the first term in (10) is thus

$$\log \det \mathbf{Y} = \sum_{i=1}^n \log(1 + \eta^2 (\hat{\lambda}_i \hat{\mu} - 1)^+). \quad (24)$$

Relabeling  $\hat{\mu}$  as  $\hat{\mu}_n/n$  in (24), we get the scaled result

$$\frac{1}{n} \log \det \mathbf{Y} = \frac{1}{n} \sum_{i=1}^n \log(1 + \eta^2 (\hat{\mu}_n \hat{\lambda}_i/n - 1)^+). \quad (25)$$

Since  $\{\hat{\lambda}_1, \hat{\lambda}_2, \dots, \hat{\lambda}_n\}$  are the eigenvalues of  $\hat{\mathbf{H}}\hat{\mathbf{H}}^\dagger$ , as  $n \rightarrow \infty$ , the empirical distribution of  $\hat{\lambda}_i/n$  converges to a limit with density as in (16). Thus, using the LLN,

$$\frac{1}{n} \log \det \mathbf{Y} \rightarrow \int_0^{8(\sigma^2 + \sigma_e^2)} \log(1 + \eta^2 (\hat{\mu}^* \lambda - 1)^+) \hat{h}(\lambda) d\lambda, \quad (26)$$

as  $n \rightarrow \infty$ , where  $\hat{\mu}^*$  satisfies (23).

Denoting  $L_{i,j}$  as the entry in the  $i$ th row and the  $j$ th column of the matrix  $\mathbf{L}$ , for  $i, j = 1, \dots, n$ , we find from (12)

$$L_{i,j} = \eta \sqrt{\hat{\lambda}_j} \tilde{Q}_j K_{i,j} + \eta \sqrt{\hat{\lambda}_i} \tilde{Q}_i K_{j,i}^*, \quad (27)$$

where  $K_{j,i}^*$  is the complex conjugate of  $K_{j,i}$ . Using (7),

$$L_{i,j} = \eta \sqrt{\hat{\lambda}_j} (\hat{\mu} - 1/\hat{\lambda}_j)^+ K_{i,j} + \eta \sqrt{\hat{\lambda}_i} (\hat{\mu} - 1/\hat{\lambda}_i)^+ K_{j,i}^*,$$

<sup>3</sup>The corresponding result in [1] is obtained in nats as opposed to (22) which is in bits/sec/Hz.

and relabeling  $\hat{\mu}$  as  $\hat{\mu}_n/n$ , we find

$$L_{i,j} = \eta \sqrt{\hat{\lambda}_j/n} (\hat{\mu}_n - (\hat{\lambda}_j/n)^{-1})^+ K_{i,j}/\sqrt{n} + \eta \sqrt{\hat{\lambda}_i/n} (\hat{\mu}_n - (\hat{\lambda}_i/n)^{-1})^+ K_{j,i}^*/\sqrt{n}. \quad (28)$$

The empirical distribution of  $\hat{\lambda}_j/n$  for any  $j = 1, 2, \dots, n$ , converges to a limit with density as in (16) as  $n \rightarrow \infty$ . Thus,  $\hat{\lambda}_j/n$  in (28) does not scale with  $n$  for any  $j$ . In addition,  $(\hat{\mu}_n - (\hat{\lambda}_j/n)^{-1})^+$  does not scale with  $n$  either. Moreover, the entries of  $\mathbf{K} = \mathbf{U}^\dagger \mathbf{X} \mathbf{V}$  are i.i.d. and ZMCSCG with finite variance  $2\sigma^2/(1 + \zeta)$ . This is because  $\mathbf{U}$  and  $\mathbf{V}$  are unitary matrices obtained from the SVD of  $\hat{\mathbf{H}}$  and the entries of  $\mathbf{X}$  are i.i.d. and ZMCSCG with variance  $2\sigma^2/(1 + \zeta)$ . Therefore, having the  $\sqrt{n}$  term in the denominator of  $L_{i,j}$  in (28), for any  $i, j = 1, 2, \dots, n$ ,  $L_{i,j} \rightarrow 0$  as  $n \rightarrow \infty$ . It is worth noting that the lower the SNR, the smaller the term  $(\hat{\mu}_n - (\hat{\lambda}_j/n)^{-1})^+$  in (28), and thus the faster the convergence.

For the matrix  $\mathbf{T}$  in (13), we have

$$T_{i,j} = \sum_{l=1}^n (\hat{\mu} - 1/\hat{\lambda}_l)^+ K_{i,l} K_{j,l}^* = \frac{1}{n} \sum_{l=1}^n (\hat{\mu}_n - (\hat{\lambda}_l/n)^{-1})^+ K_{i,l} K_{j,l}^*, \quad (29)$$

which is obtained by relabeling  $\hat{\mu}$  as  $\hat{\mu}_n/n$ . Using the LLN, as  $n \rightarrow \infty$ ,  $T_{i,j}$  converges to

$$T_{i,j} \rightarrow \mathbb{E}[(\hat{\mu}^* - 1/\hat{\lambda})^+ K_{i,l} K_{j,l}^*]. \quad (30)$$

Since  $\hat{\mathbf{H}} \rightarrow (\mathbf{U}, \mathbf{V}) \rightarrow \mathbf{K}$  is a Markov chain, thus  $\mathbf{K}$  and  $\hat{\mathbf{H}}$  are independent given  $\mathbf{U}$  and  $\mathbf{V}$ . Therefore, we can write

$$\begin{aligned} & \mathbb{E}[(\hat{\mu}^* - 1/\hat{\lambda})^+ K_{i,l} K_{j,l}^*] \\ &= \mathbb{E}[\mathbb{E}[(\hat{\mu}^* - 1/\hat{\lambda})^+ K_{i,l} K_{j,l}^* | \mathbf{U}, \mathbf{V}]] \\ &= \mathbb{E}[\mathbb{E}[(\hat{\mu}^* - 1/\hat{\lambda})^+ | \mathbf{U}, \mathbf{V}] \mathbb{E}[K_{i,l} K_{j,l}^* | \mathbf{U}, \mathbf{V}]]. \end{aligned} \quad (31)$$

The water-filling level  $\hat{\mu}^*$  was chosen in (8) such that (8) holds for each realization of  $\hat{\mathbf{H}}$ . Thus,

$$\mathbb{E}[(\hat{\mu}^* - 1/\hat{\lambda})^+ | \mathbf{U}, \mathbf{V}] = P. \quad (32)$$

Hence, from (30), (31), and (32) we obtain

$$T_{i,i} \rightarrow PE[\mathbb{E}[K_{i,l}^2]] = 2P\sigma^2/(1 + \zeta),$$

for  $i = 1, 2, \dots, n$ , and

$$T_{i,j} \rightarrow PE[K_{i,l} \mathbb{E}[K_{j,l}^*]] = 0,$$

for  $i, j = 1, 2, \dots, n$ , and  $i \neq j$ .

Although, elementwise, off-diagonal entries  $Z_{i,j}$  of  $\mathbf{Z} = \mathbf{I}_n + \mathbf{Y}^{-1/2} (\mathbf{L} + \mathbf{T}) \mathbf{Y}^{-1/2}$  for  $i, j = 1, \dots, n$ ,  $i \neq j$  are such that  $Z_{i,j} \rightarrow 0$  as  $n \rightarrow \infty$ , this does not necessarily imply that  $\det \mathbf{Z} \rightarrow \prod_{i=1}^n Z_{i,i}$  as  $n \rightarrow \infty$ . This is because the effect of off-diagonal entries of  $\mathbf{Z}$  may not necessarily be neglected since the number of these entries are increasing with  $n$ . However, for a fixed  $n$ , as  $P \rightarrow 0$ , the off-diagonal entries of  $\mathbf{Z}$  vanish and thus  $\det \mathbf{Z} \rightarrow \prod_{i=1}^n Z_{i,i}$ . For example, it is easy to verify by cofactor expansion for an  $n \times n$  matrix  $\mathbf{A}$  with constant entries and  $\epsilon \rightarrow 0$  that  $\det[\mathbf{I}_n + \epsilon \mathbf{A}] = \prod_{i=1}^n (1 + \epsilon A_{i,i}) + O(\epsilon^2)$ . Therefore, in the low SNR regime,  $\det \mathbf{Z} \approx \prod_{i=1}^n Z_{i,i} = \det[\mathbf{I}_n + (2P\sigma^2/(1 + \zeta)) \mathbf{Y}^{-1}]$  as  $n \rightarrow \infty$ . Thus, knowing  $\mathbf{Y}$  from (11) and using the LLN, we can write

$$\frac{1}{n} \log \det \mathbf{Z} \rightarrow \int_0^{8(\sigma^2 + \sigma_e^2)} \log \left( 1 + \frac{2P\sigma^2/(1 + \zeta)}{1 + \eta^2 (\hat{\mu}^* - 1)^+} \right) \hat{h}(\lambda) d\lambda. \quad (33)$$

Hence, from (26) and (33), we find  $R_n/n \rightarrow R$  as  $n \rightarrow \infty$ . Therefore, at low SNR,  $R_n$  scales linearly with  $n$  as  $n \rightarrow \infty$  with a proportionality constant (the asymptotic growth rate) that is not random and is given by  $R$  as

$$R \approx \int_0^{8\sigma^2(1+1/\zeta)} \log(1 + (\zeta^2/(1+\zeta)^2)(\lambda\hat{\mu}^* - 1)^+) \hat{h}(\lambda) d\lambda + \int_0^{8\sigma^2(1+1/\zeta)} \log\left(1 + \frac{2P\sigma^2/(\zeta+1)}{1 + (\zeta^2/(1+\zeta)^2)(\lambda\hat{\mu}^* - 1)^+}\right) \hat{h}(\lambda) d\lambda, \quad (34)$$

where  $\hat{\mu}^*$  satisfies (23). Using (34), one can evaluate the effect of channel estimation error on the performance of water-filling power allocation at low SNR. Note that for a fixed  $\sigma$ , when  $\zeta = \infty$  dB (high SER), the second term in (34) is zero and

$$R = \int_0^{8\sigma^2} (\log(\lambda\hat{\mu}^*))^+ h(\lambda) d\lambda = C^P,$$

which is the asymptotic growth rate for water-filling with perfect CSIT as found in (20). Furthermore, for a fixed  $\sigma$  and  $\zeta = -\infty$  dB (low SER), the first term in (34) is zero and as  $P \rightarrow 0$  we have

$$R \approx \int_0^\infty \log(1 + 2P\sigma^2) \hat{h}(\lambda) d\lambda \approx 2P\sigma^2 \log e \int_0^\infty \hat{h}(\lambda) d\lambda = 2P\sigma^2 \log e,$$

which is equal to  $I^N$  in (19). This result implies that, at low SER, water-filling based on erroneous channel estimates achieves the same throughput as equal power allocation asymptotically in the low SNR regime.

## VI. NUMERICAL RESULTS

In this section, we provide numerical results when  $\sigma = 1$  and  $P = -20$  dB. Fig. 1 shows the convergence of the actual scaled throughputs (averaged over multiple realizations)  $I_n^N/n$ ,  $C_n^P/n$ , and  $R_n/n$  to the asymptotic (analytic) results  $I^N$ ,  $C^P$ , and  $R$  presented in (18), (20), and (34) respectively, as  $n$  increases. For  $n = 50$ , Fig. 2 plots the actual (averaged over multiple realizations) and the asymptotic scaled throughputs versus the SER. Based on the figures, water-filling with imperfect CSIT monotonically interpolates between water-filling with perfect CSIT and equal power allocation. Furthermore, one can compare the result for water-filling with imperfect CSIT with the result for water-filling with perfect CSIT as a function of the SER. Particularly, with imperfect CSIT and for SER values such as 5 dB, 0 dB, and  $-5$  dB,  $R$  is 86%, 70%, and 52% of  $C^P$  respectively. In other words, at low SNR, water-filling based on erroneous channel estimates can still achieve significant throughputs asymptotically for moderate values of the SER. For imperfect CSIT, the gap between the actual and the asymptotic (analytic) results in the figures is very small. When  $P$  is above  $-15$  dB, the approximation in (34) starts to deviate notably. However, numerically, we find that at higher SNR values, the same conclusion holds that moderate SER of 5 dB provides performance close to  $C^P$ .

## VII. CONCLUDING REMARKS

We have derived the water-filling throughput of a point-to-point MIMO channel in the presence of imperfect CSIT. Consequently, we have found the asymptotic growth rate for

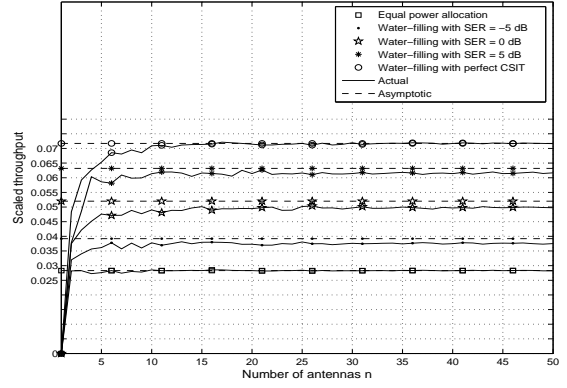


Fig. 1. Scaled throughputs when  $P = -20$  dB and  $\sigma = 1$ .

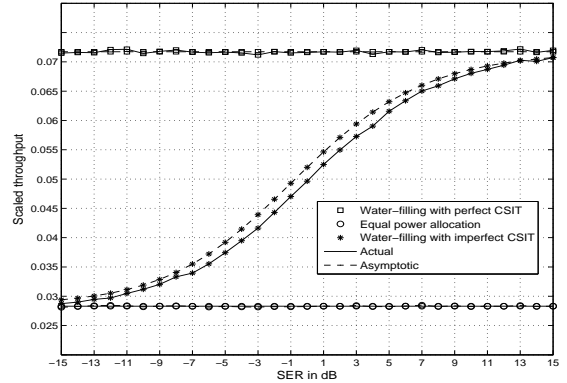


Fig. 2. Scaled throughputs when  $P = -20$  dB,  $\sigma = 1$ , and  $n = 50$ .

the throughput with imperfect CSIT at low SNR and compared it with the asymptotic growth rate for the capacity with perfect CSIT as a function of the SER. We have shown that at low SNR and for moderate values of the SER, water-filling even based on erroneous channel estimates can still achieve significant throughputs asymptotically.

## REFERENCES

- [1] C. N. Chuah, D. N. C. Tse, J. M. Kahn, and R. A. Valenzuela, "Capacity scaling in MIMO wireless systems under correlated fading," *IEEE Trans. Inform. Theory*, vol. 48, pp. 637-650, Mar. 2002.
- [2] G. J. Foschini and M. J. Gans, "On limits of wireless communication in a fading environment when using multiple antennas," *Wireless Personal Commun.*, vol. 6, no. 3, pp. 311-335, Mar. 1998.
- [3] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE J. Sel. Areas Commun.*, vol. 51, no. 6, pp. 684-702, Jun. 2003.
- [4] E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecommun.*, vol. 10, pp. 585-598, Nov. 1999.
- [5] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Proces. Mag.*, vol. 30, no. 1, pp. 40-46, Jan. 2013.
- [6] 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.
- [7] J. Hoydis, K. Hosseini, S. ten Brink, and M. Debbah, "Making smart use of excess antennas: Massive MIMO, small cells, and TDD," *Bell Labs Tech. Journal*, vol. 18, no. 2, pp. 5-21, Sep. 2013.
- [8] H.L. Van Trees, *Detection, Estimation, and Modulation Theory, Part I*, New York: John Wiley and Sons, 1968.
- [9] D. Gu and C. Leung, "Performance analysis of transmit diversity scheme with imperfect channel estimation," *Elect. Lett.*, vol. 39, pp. 402-403, Feb. 2003.