On the Optimal Sum Capacity for OFDM With On/Off Power Allocation and Imperfect Channel Estimation

Wiroonsak Santipach

Department of Electrical Engineering

Faculty of Engineering, Kasetsart University

Bangkok, Thailand 10900

Email: `wiroonsak.s@ku.ac.th`

Abstract—We consider a point-to-point frequency-selective orthogonal frequency division multiplexing (OFDM) channel. Based on a linear minimum mean square error (MMSE) channel estimates from pilot signals, a receiver determines a set of 'on' or active subchannels that maximizes a lower bound on a sum capacity. The capacity bound optimization is constrained to a fixed total power, which is a sum of total training power and a sum of equal transmission power on activated subchannels. We show that as number of subchannels N increases, the optimal training power converges to a deterministic constant while the optimal number of active subchannels increases at the rate of $(\log N - 2 \log \log N)^2$. The associated capacity bound tends to infinity at the rate of $\log N - 2 \log \log N$ while that with uniform power allocation of all subchannels converges to a constant.

I. INTRODUCTION

Orthogonal frequency division multiplexing (OFDM) is widely used in current wireless communications systems such as digital audio broadcast (DAB), digital video broadcast (DVB), IEEE Std. 802.11 (WiFi), IEEE Std. 802.16 (WiMAX) and is a core technique for the fourth generation (4G) cellular telephony. OFDM provides a high spectral efficiency over a frequency selective channel. In OFDM system, data stream is divided and transmitted over many low-rate parallel subchannels, which help increase symbol duration and reduce intersymbol interference [1], [2].

For coherent detection, a receiver must track a time-varying channel. Channel information is also useful at the transmitter that adapts transmitted power according to channel variations. Several pilot-aided channel estimation schemes for OFDM have been studied [3]–[5]. Here we assume that known pilots are transmitted to estimate time-domain channel taps and that channel fades slowly so that channel estimates are valid for several OFDM blocks. A linear minimum mean square error (MMSE) estimator is employed at the receiver. Accuracy of the obtained channel estimates depends on power allocated for pilots.

Based on channel estimates, the receiver can compute the optimal transmitted power allocation, which is water-pouring across subchannels and then can relay set of power levels via a feedback channel to the transmitter. Since a number of subchannels denoted by N is usually large, a number of feedback bits required to specify power levels can be significantly large. To reduce amount of feedback, we choose a suboptimal on/off threshold-based power allocation. If a subchannel gain exceeds a threshold, then equal power is allocated for that subchannel. Otherwise, no power will be allocated for transmission and thus, that subchannel is not used. The on/off scheme requires only 1 feedback bit per subchannel, which is substantially less than water-pouring algorithm, and the scheme has been shown to perform close to the optimal water-pouring solution [6], [7]. References [7], [8] considered similar problem. However, [7] assumes perfect channel knowledge at the receiver while [8] chose to estimate N subchannel gains instead of channel taps whose number is much smaller.

We derive the lower bound on a sum capacity over all subchannels with channel estimation and the on/off power allocation and which maximizes the bound, subject to a total power constraint. The total power consists of total power allocated for training and that for transmission. As N tends to infinity, we show that the optimal training power converges to the constant that depends on number of channel taps while the number of active subchannels increases as $(\log N - 2 \log \log N)^2$. The lower bound on the sum capacity grows as $\log N - 2 \log \log N$. We note that the results obtained here is different from those derived in [8] due to different assumption on channel estimation. Numerical examples illustrate a performance tradeoff between training power, which increases accuracy of channel estimates and transmission power.

II. SYSTEM MODEL

We consider a point-to-point discrete-time OFDM system with N subchannels. The channel $\boldsymbol{h} = [h_0 \quad h_2 \quad \ldots \quad h_{M-1}]^T$ is assumed to be a frequency selective Rayleigh fading channel with order M and each channel tap h_m is independent complex Gaussian distributed with zero mean and variance σ_m^2. Also

$\sigma_m, 0 \le m \le M-1$ represents a power delay profile and is normalized

$$\sum_{m=0}^{M-1} \sigma_m^2 = 1. \tag{1}$$

Thus, $\sigma_m^2 = \frac{1}{M}$ for all m.

We assume independent block fading in which channel taps remain relatively unchanged for several OFDM symbols before changing to the next independent set. This slow fading channel model allows a meaningful adaptation of transmission power allocation. Let

$$H(i) = \sum_{m=0}^{M-1} h_m \mathrm{e}^{-\frac{j2\pi mi}{N}} \tag{2}$$

denote a discrete Fourier transform (DFT) of $\boldsymbol{h}$ for the ith subchannel. Thus, the output signal of a DFT unit at the receiver is given by

$$y(i) = H(i)x(i) + n(i), \quad 0 \le i \le N-1 \tag{3}$$

where $x(i)$ is a transmitted symbol on the ith subchannel and $n(i)$ is an additive white Gaussian noise (AWGN) with zero mean and variance σ_n^2.

A. Channel Estimation

To adapt a transmission power over subchannels, current channel coefficients are needed. We assume that the transmitter sends to the receiver pilot signals on N_T subchannels during a training period. N_T pilots are inserted in the OFDM block at known location $\{i_n; 0 \le n \le N_T - 1\}$ with each pilot allocated power $\bar{P}_T$. The received pilot signals at the output of a DFT unit is given by

$$\boldsymbol{y}_T = [y_T(i_0) \quad y_T(i_1) \quad \cdots \quad y_T(i_{N_T-1})]^T. \tag{4}$$

Applying MMSE estimation, we obtain the estimate of the $M \times 1$ vector of channel taps as follows [9]

$$\hat{\boldsymbol{h}} = (\sigma_n^2 (E[\boldsymbol{h}\boldsymbol{h}^\dagger])^{-1} + \boldsymbol{B}^\dagger \boldsymbol{B})^{-1} \boldsymbol{B}^\dagger \boldsymbol{y}_T \tag{5}$$

$$= (\sigma_n^2 M \boldsymbol{I} + \boldsymbol{B}^\dagger \boldsymbol{B})^{-1} \boldsymbol{B}^\dagger \boldsymbol{y}_T \tag{6}$$

where (6) follows from the earlier assumption on channel tap statistics ($E[\boldsymbol{h}\boldsymbol{h}^\dagger] = \frac{1}{M}\boldsymbol{I}$), $\boldsymbol{I}$ is an $M \times M$ identity matrix, and $\boldsymbol{B}$ is an $N_T \times M$ matrix whose entry

$$[\boldsymbol{B}]_{n,m} = \mathrm{e}^{-\frac{j2\pi mi_n}{N}}, \quad 0 \le n \le N_T-1, 0 \le m \le M-1. \tag{7}$$

Reference [3] shows that the total MMSE is achieved when pilots are uniformly placed among N subchannels. Computing the DFT of $\hat{\boldsymbol{h}}$, we obtain an estimate of $H(i)$ denoted by

$$\hat{H}(i) = \sum_{m=0}^{M-1} \hat{h}_m \mathrm{e}^{-\frac{j2\pi mi}{N}}. \tag{8}$$

The resulting MSE for each subchannel is given by [10]

$$E|\Delta H(i)|^2 \triangleq E|\hat{H}(i) - H(i)|^2 \tag{9}$$

$$= \frac{\sigma_n^2 M}{\bar{P}_T N_T + \sigma_n^2 M} \tag{10}$$

$$= \frac{\sigma_n^2 M}{P_T + \sigma_n^2 M} \tag{11}$$

where $P_T \triangleq \bar{P}_T N_T$ is a total power used for pilots. We note from (11) that the estimation variance is the same for all subchannels and that the MSE decreases as either number of channel taps M to estimate increases or power allocated for training channel decreases. With an orthogonal property of MMSE estimation, the variance of the channel estimate is given by

$$\sigma_{\hat{H}}^2 \triangleq E|\hat{H}(i)|^2 \tag{12}$$

$$= E|H(i)|^2 - E|\Delta H(i)|^2 \tag{13}$$

$$= \frac{P_T}{P_T + \sigma_n^2 M}. \tag{14}$$

B. On/Off Power Allocation

The MMSE estimate of the subchannel $\hat{H}(i)$ is a complex Gaussian random variable with zero mean and variance $\sigma_{\hat{H}}^2$. Based on the channel estimates, the receiver employs on/off transmission power allocation. Namely, if $|\hat{H}(i)|^2$ exceeds a threshold μ, the ith subchannel will be allocated an equal power or will be 'on' during a transmission period. Otherwise, the ith subchannel will not be allocated power and hence will be 'off'. The on/off power scheme is suboptimal, however, was shown to perform close to the optimal water-pouring scheme [6], [7].

Let N_D denote a number of subchannels whose channel gains exceed or equal a threshold, which is given by

$$N_D = \sum_{i=1}^{N} \mathbf{1}_\mu(|\hat{H}(i)|^2) \tag{15}$$

where a indicator function

$$\mathbf{1}_\mu(x) = \begin{cases} 1 & : & x \ge \mu \\ 0 & : & x < \mu \end{cases}, \tag{16}$$

N_D is random and depends on the threshold μ. For a fixed total transmission power, as N_D increases, the power per active subchannel decreases. Although channel taps $\{h_m\}$ are independent, $\hat{H}(i), \forall i$ are *correlated* complex Gaussian sequence with zero mean and variance $\sigma_{\hat{H}}^2$. ($\Re\{\hat{H}(i)\}$ and $\Im\{\hat{H}(i)\}$ are independent Gaussian with zero mean and variance $\sigma_{\hat{H}}^2/2$, where $\Re\{x\}$ and $\Im\{x\}$ are real and imaginary parts of x, respectively.) Thus, $|\hat{H}(i)|^2$ is exponential distributed with probability density function (pdf)

$$f_{|\hat{H}|^2}(x) = \frac{1}{\sigma_{\hat{H}}^2} \mathrm{e}^{-\frac{x}{\sigma_{\hat{H}}^2}}. \tag{17}$$

For given μ, we can compute expected number of active subchannel as follows

$$E[N_D] = \sum_{i=1}^{N} E[\mathbf{1}_\mu(|\hat{H}(i)|^2)] \tag{18}$$

$$= N \int_\mu^\infty f_{|\hat{H}|^2}(x) \, \mathrm{d}x \tag{19}$$

$$= N \mathrm{e}^{-\frac{\mu}{\sigma_{\hat{H}}^2}}. \tag{20}$$

From (20) we note that N_D is decreasing exponentially with μ. The variance is given by

$$\text{var}[N_D] = \sum_{i=1}^{N} \text{var}[\mathbf{1}_\mu(|\hat{H}(i)|^2)]$$
$$+ \sum_{i=1}^{N} \sum_{\substack{k=1-i, \\ k \neq 0}}^{N-i} \text{cov}[\mathbf{1}_\mu(|\hat{H}(i)|^2)\mathbf{1}_\mu(|\hat{H}(i+k)|^2)]. \tag{21}$$

The variance term in (21) can be easily computed using the pdf in (17) and is given by

$$\text{var}[\mathbf{1}_\mu(|\hat{H}(i)|^2)] = \text{e}^{-\frac{\mu}{\sigma_{\hat{H}}^2}} - \text{e}^{-\frac{2\mu}{\sigma_{\hat{H}}^2}}. \tag{22}$$

The covariance computation requires a joint pdf for channel gains $\hat{H}(i)$ and $\hat{H}(i+k)$. Reference [11] shows that correlations between subchannel gains are given by

$$E[\Re\{\hat{H}(i)\}\Re\{\hat{H}(j)\}] = E[\Im\{\hat{H}(i)\}\Im\{\hat{H}(j)\}] \tag{23}$$
$$= \frac{\sigma_{\hat{H}}^2}{2M} \sum_{m=0}^{M-1} \cos\frac{2\pi m(i-j)}{N} \tag{24}$$

and

$$E[\Re\{\hat{H}(i)\}\Im\{\hat{H}(j)\}] = \frac{\sigma_{\hat{H}}^2}{2M} \sum_{m=0}^{M-1} \sin\frac{2\pi m(i-j)}{N} \tag{25}$$

and pdf for $|\hat{H}(i)|^2$ and $|\hat{H}(i+k)|^2$ is jointly exponential and is given by

$$f_{|\hat{H}(i)|^2,|\hat{H}(i+k)|^2}(x_1, x_2) = \frac{1}{\sigma_{\hat{H}}^4(1-\rho_k^2)} \text{e}^{-\frac{x_1+x_2}{\sigma_{\hat{H}}^2(1-\rho_k^2)}}$$
$$\times I_0\left(\frac{2\rho_k}{\sigma_{\hat{H}}^2(1-\rho_k^2)}\sqrt{x_1 x_2}\right) \tag{26}$$

where $I_0(\cdot)$ is a modified Bessel function of the first kind and correlation coefficient

$$\rho_k^2 = \frac{4}{\sigma_{\hat{H}}^4}\left(E^2[\Re\{\hat{H}(i)\}\Re\{\hat{H}(i+k)\}]\right.$$
$$\left. + E^2[\Re\{\hat{H}(i)\}\Im\{\hat{H}(i+k)\}]\right) \tag{27}$$
$$= \frac{1}{M^2}\left(\sum_{m=0}^{M-1}\cos\frac{2\pi mk}{N}\right)^2 + \frac{1}{M^2}\left(\sum_{m=0}^{M-1}\sin\frac{2\pi mk}{N}\right)^2 \tag{28}$$
$$= \left(\frac{\sin\frac{\pi k}{N}M}{M\sin\frac{\pi k}{N}}\right)^2. \tag{29}$$

Deducing (29) requires some trigonometry manipulation.

Thus, the covariance term in (21) can be computed as follows

$$\text{cov}[\mathbf{1}_\mu(|\hat{H}(i)|^2)\mathbf{1}_\mu(|\hat{H}(i+k)|^2)]$$
$$= E[\mathbf{1}_\mu(|\hat{H}(i)|^2)\mathbf{1}_\mu(|\hat{H}(i+k)|^2)] \tag{30}$$
$$- E[\mathbf{1}_\mu(|\hat{H}(i)|^2)]E[\mathbf{1}_\mu(|\hat{H}(i+k)|^2)] \tag{31}$$
$$= \int_\mu^\infty \int_\mu^\infty f_{|\hat{H}(i)|^2,|\hat{H}(i+k)|^2}(x_1, x_2) \, \text{d}x_1 \text{d}x_2 - \text{e}^{-\frac{2\mu}{\sigma_{\hat{H}}^2}}. \tag{32}$$

After some algebraic manipulation, we can show that

$$\int_\mu^\infty \int_\mu^\infty f_{\hat{H}(i),\hat{H}(i+k)}(x_1, x_2) \, \text{d}x_1 \text{d}x_2$$
$$= (1-\rho_k^2)\text{e}^{-\frac{2\mu}{\sigma_{\hat{H}}^2(1-\rho_k^2)}} \sum_{n=0}^\infty \left[\frac{\rho_k^2}{\sigma_{\hat{H}}^4(1-\rho_k^2)^2}\right]^n$$
$$\times \frac{1}{n\Gamma(n+1)}\left[\sum_{l=0}^n (\sigma_{\hat{H}}^2(1-\rho_k^2))^l \frac{n}{(n-l)}\mu^{n-l}\right]^2 \tag{33}$$

where $\Gamma(\cdot)$ is the Gamma function. Substituting (33) and (29) into (32) gives the covariance between $|\hat{H}(i)|^2$ and $|\hat{H}(i+k)|^2$. However, if number of subchannels N is increasing with fixed number of channel taps M, the frequency-selective channel will converge to a flat fading channel with gains $H(i) = \sum_{m=0}^{M-1} h_m, \forall i$. Namely, as $N \to \infty$,

$$\text{cov}[\mathbf{1}_\mu(|\hat{H}(i)|^2)\mathbf{1}_\mu(|\hat{H}(i+k)|^2)] \to \text{var}[\mathbf{1}_\mu(|\hat{H}(i)|^2)] \tag{34}$$

where $\text{var}[\mathbf{1}_\mu(|\hat{H}(i)|^2)]$ is given in (22).

Let P_D denote total power allocated for transmission. Therefore, for on/off allocation, the power allocated for each *active* subchannel is equal to $\bar{P}_D \triangleq P_D/N_D$.

III. CAPACITY BOUND OPTIMIZATION

In this work, a sum capacity on N subchannels is a performance metric for the system. Since analyzing the sum capacity with channel estimation is not tractable, instead we derive its lower bound, which is obtained by assuming that the sum of channel estimation error and AWGN is Gaussian. Thus, the sum capacity

$$C \geq \sum_{i=1}^{N} E\left[\log\left(1 + \frac{\bar{P}_D|\hat{H}(i)|^2}{\bar{P}_D E|\Delta H(i)|^2 + \sigma_n^2}\right)\mathbf{1}_\mu(|\hat{H}(i)|^2)\right] \tag{35}$$

where $\log$ is a natural logarithm and the expectation is over both $\hat{H}(i)$ and $\bar{P}_D$. Conditioned on $\bar{P}_D$ (and hence N_D), the expectation in (35) can be explicitly expressed as follows

$$C_{L|\bar{P}_D}$$
$$\triangleq \int_\mu^\infty \log\left(1 + \frac{\bar{P}_D x}{\bar{P}_D E|\Delta H(i)|^2 + \sigma_n^2}\right) f_{|\hat{H}|^2}(x) \, \text{d}x \tag{36}$$
$$= \text{e}^{\frac{1}{W}}\text{Ei}_1(W^{-1}+\mu') + \text{e}^{-\mu'}\log(1+W\mu') \tag{37}$$

where an exponential integral

$$\text{Ei}_m(z) = \int_1^\infty \text{e}^{-tz}t^{-m} \, \text{d}t \tag{38}$$

and

$$W \triangleq \frac{\bar{P}_D}{\bar{P}_D \frac{\sigma_n^2 M}{P_T} + \frac{\sigma_n^2}{\sigma_{\hat{H}}^2}}, \tag{39}$$

$$\mu' \triangleq \frac{\mu}{\sigma_{\hat{H}}^2}. \tag{40}$$

The exponential integral term in (37) has tight lower and upper bounds [12]

$$\frac{1}{2}e^{-\mu'}\log\left(1 + \frac{2}{W^{-1} + \mu'}\right) \le e^{\frac{1}{W}}\mathrm{Ei}_1(W^{-1} + \mu')$$
$$\le e^{-\mu'}\log\left(1 + \frac{1}{W^{-1} + \mu'}\right). \tag{41}$$

We would like to maximize the lower bound of the sum capacity subject to the total power constraint as follows

$$\max_{P_T, \bar{P}_D, \bar{N}_D} \left\{ C_L \triangleq N E[C_{L|\bar{P}_D}] \right\} \tag{42}$$
$$\text{subject to} \quad P = P_T + \bar{P}_D \bar{N}_D N \tag{43}$$

where $\bar{N}_D \triangleq N_D/N$ denotes the normalized number of active subchannels and the optimal set of solutions and the optimized capacity lower bound are denoted by $\{P_T^*, \bar{P}_D^*, \bar{N}_D^*\}$ and C_L^*, respectively. We resort to numerical method to solve this optimization problem for finite system parameters. However, numerical solutions obtained do not give us much insights into the system performance. By letting $N \to \infty$, we can derive asymptotic solutions as stated in Theorem 1. The limit considered is reasonable since in a practical OFDM system, N is usually large (e.g. 512).

Theorem 1: As $N \to \infty$, the solutions to (42)-(43) converge as follows

$$P_T^* \longrightarrow \sigma_n^2 M \left(\sqrt{\frac{P}{\sigma_n^2 M} + 1} - 1 \right) \tag{44}$$

$$\bar{P}_D^* \log^2 \bar{N}_D^* \longrightarrow 2\sigma_n^2 \left(1 + \frac{\sigma_n^2 M}{P_T} \right). \tag{45}$$

The associated sum capacity is given by

$$\frac{C_L^*}{\log N - 2\log\log N} \longrightarrow \frac{P_T(P - P_T)}{\sigma_n^2(P_T + \sigma_n^2 M)}. \tag{46}$$

To derive these results, we first define the Lagrangian given by

$$L(P_T, \bar{P}_D, \bar{N}_D, \lambda) = C_L + \lambda(P - P_T - \bar{P}_D \bar{N}_D N). \tag{47}$$

Applying the Karush–Kuhn–Tucker conditions, the bounds in (41), and the asymptotic mean and variance of N_D derived in previous section, we obtain Theorem 1. (Due to limited space, the detailed derivation is omitted here.)

From Theorem 1, the optimal training power converges to the constant, which is a function of number of channel taps, noise variance, and total power. We note that for large P and N, P_T^* scales with $\sqrt{P\sigma_n^2 M}$. However, the optimal power per active subchannel $\bar{P}_D$ decreasing to zero implies that N_D must tend to infinity. With (45), we can show that the optimal N_D scales with $(\log N - 2\log\log N)^2$ as $N \to \infty$. The associated

optimized capacity bound also increases at the rate of $\log N - 2\log\log N$. The results derived should be a good indicator for the solutions for the *actual* sum capacity, which remains an open question.

IV. NUMERICAL RESULTS

In this section, we show some simulation results and compare them with analytical ones from Theorem 1. In Fig. 1, we consider an OFDM channel with number of channel taps $M = 20$, total available power $P = 20$, and noise variance $\sigma_n^2 = 0.1$ and find the optimal set of parameters that maximizes the capacity bound. As we increase number of subchannels N, we observe that the optimal training power converges to a constant while the optimal power per subchannel and number of active subchannels decreases to zero and increases to infinity, respectively. At $N = 512$, $N_D^* \approx 184$ or approximately one third of all subchannels should be activated and $P_T^* = 6.1$ or approximately 30% of total power should be used to estimate 10 channel taps.
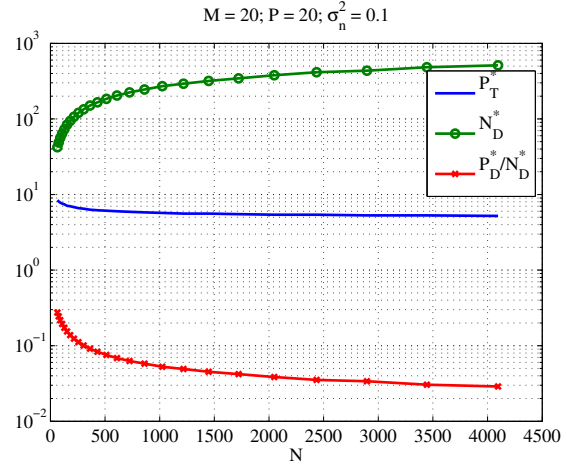

Fig. 1. Shown is the optimal training power P_T^*, transmission power allocated for each active subchannel $\bar{P}_D^*$, and number of active subchannel N_D^* with number of total subchannels N.

In Fig. 2, we compare the simulation results with asymptotic results in Theorem 1 for $M = 10$, $P = 10$, and $\sigma_n^2 = 0.1$. Both P_T^* and $\bar{P}_D^* \log^2 \bar{N}_D^*$ converge to their limiting values as $N \to \infty$. We note that the rate of convergence is quite slow. Thus, the asymptotic results can be a good approximation for a system performance with large N.

The capacity bound associated with the optimal parameters in Fig. 1 is plotted with a solid line in Fig. 3. We see that the lower bound increases with N. We also show the sum capacity with perfect channel estimates with a dashed line. The gap between the two performances is substantial. A more advanced estimation method should narrow the gap between the two. Finally we plot a capacity lower bound with uniform power allocation with a dashed-dot line. Namely, transmission power is uniformly allocated on all subchannels. Uniform power scheme performs much worse than the on/off one. We also note that the capacity of the uniform power allocation
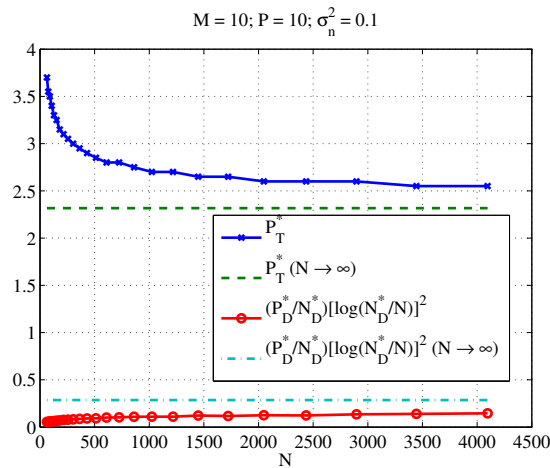
Fig. 2. Simulation results for P_T^* and $\bar{P}_D^* \log^2 \bar{N}_D^*$ is shown with their limiting values derived in Theorem 1.

converges as N tends to infinity [13]. However, the uniform power scheme does not require any feedback while the on/off power scheme requires 1 feedback bit per subchannel.
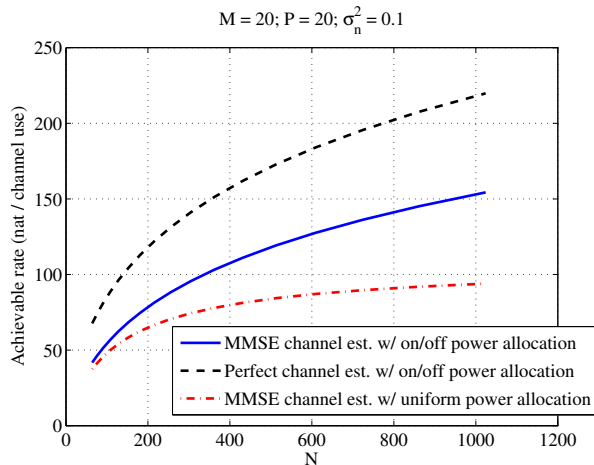

Fig. 3. Rates (nat per channel use) for different assumptions on channel estimation and power allocation are shown with number of subchannels.

Fig. 4 shows the capacity lower bound with a ratio between power used for training and total power for different M, P, and σ_n^2. $\bar{P}_D$ and $\bar{N}_D$ are set approximately to the asymptotic value in (45). We see that for $M = 20$, $P = 20$, and $\sigma_n^2 = 0.1$, the maximum performance is achieved when P_T/P is about 30%, which is approximately the same to what the simulation predicts in Fig. 1. Thus, the asymptotic results can give a good approximation for optimal solutions for finite-size system. When M is increased to 20, the optimal P_T/P also increases.

V. CONCLUSIONS

In this work, we derived the asymptotic optimal training power, transmission power, and number of active subchannels
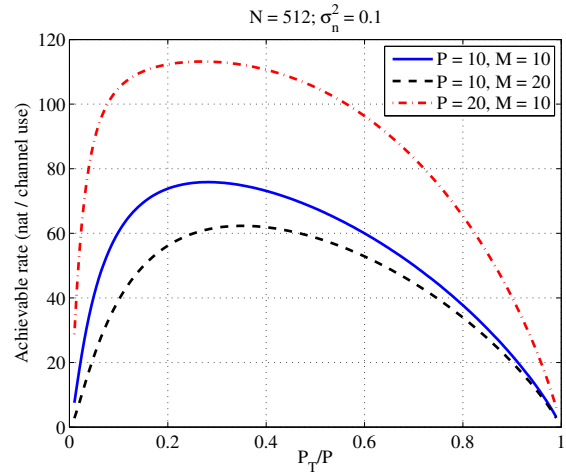

Fig. 4. The lower bound is shown with ratio between training power and total power.

in a point-to-point OFDM system. As number of total sub-channels increases, the total power for training converges to a deterministic constant, which is a function of total power, noise variance, and number of channel taps. The number of active subchannels increases at the rate of $(\log N - 2 \log \log N)^2$. The numerical results showed that channel estimation error has a significant impact on the performance and the on/off power allocation performs much better than a uniform power allocation. However, the performance gain from the on/off power allocation comes with the cost of N feedback bits. We also showed that operating at the optimal parameters can give significant performance gain.

However, our results here only apply to channels with uniform power delay profile. A more practical channel to consider may have exponential-decaying profile. Future work includes extending the current results to a multiantenna channel.

REFERENCES

[1] S. B. Weinstein and P. M. Ebert, "Data transmission by frequency-division multiplexing using the discrete Fourier transform," *IEEE Trans. Commun.*, vol. 19, no. 5, pp. 628–634, Oct. 1971.

[2] L. J. Cimini, "Analysis and simulation of a digital mobile channel using orthogonal frequency division multiplexing," *IEEE Trans. Commun.*, vol. 33, no. 7, pp. 665–675, Jul. 1985.

[3] R. Negi and J. Cioffi, "Pilot tone selection for channel estimation in a mobile OFDM system," *IEEE Trans. Consum. Electron.*, vol. 44, no. 3, pp. 1122–1128, Aug. 1998.

[4] L. Tong, B. M. Sadler, and M. Dong, "Pilot-assisted wireless transmissions: General model, design criteria, and signal processing," *IEEE Signal Process. Mag.*, vol. 21, no. 6, pp. 12 – 25, Nov. 2004.

[5] Y. Le, "Pilot-symbol-aided channel estimation for OFDM in wireless systems," *IEEE Trans. Veh. Technol.*, vol. 49, no. 4, pp. 1207–1215, Jul. 2000.

[6] J. Jang and K.-B. Lee, "Transmit power adaptation for multiuser OFDM systems," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 2, pp. 171–178, Feb. 2003.

[7] Y. Sun and M. L. Honig, "Asymptotic capacity of multicarrier transmission with frequency-selective fading and limited feedback," *IEEE Trans. Inf. Theory*, vol. 54, no. 7, pp. 2879–2902, Jul. 2008.

[8] M. Agarwal and M. L. Honig, "Wideband fading channel capacity with training and partial feedback," in *Allerton Conf. on Commun., Control, and Computing*, Monticello, Illinois, Sep. 2005, pp. 65–74.

[9] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, New Jersey: Prentice-Hall, 1993.

[10] M. Morelli and U. Mengali, "A comparison of pilot-aided channel estimation methods for OFDM systems," *IEEE Trans. Signal Process.*, vol. 49, no. 12, pp. 3065–3073, Dec. 2001.

[11] M. Mittelbach, C. Müller, and K. Schubert, "Ergodic capacity of discrete- and continuous-time, frequency-selective Rayleigh fading channels with correlated scattering," in *IEEE Global Commun. Conf. (GLOBECOM)*, Washington, DC, Nov. 2007, pp. 3332–3337.

[12] M. Abramovitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. New York: Dover, 1964.

[13] A. Clark, P. J. Smith, and D. P. Taylor, "Instantaneous capacity of OFDM on Rayleigh-fading channels," *IEEE Trans. Inf. Theory*, vol. 53, no. 1, pp. 355 –361, Jan. 2007.