Optimum Ergodic Power Allocation for Multiuser OFDM with Arbitrary Signal Constellations

Angel Lozano
Bell Labs (Lucent Technologies)
Holmdel, NJ 07733, USA
Email: aloz@lucent.com

Antonia M. Tulino
Universitá di Napoli "Federico II"
Naples 80125, Italy
Email: atulino@ee.princeton.edu

Sergio Verdú
Princeton University
Princeton, NJ 08544, USA
Email: verdu@princeton.edu

***Abstract*— This paper formulates the power allocation policy that maximizes the region of ergodic mutual informations achievable in multiuser downlink OFDM channels known only statistically by the base station. Arbitrary partitioning of the available tones among users and arbitrary modulation formats, possibly different for every user, are considered. The derivation relies on the nexus between the mutual information of Gaussian channels and the minimum mean-square error incurred in the nonlinear estimation of the transmit constellation points given their noisy receive observations.**

I. INTRODUCTION

OFDM (orthogonal frequency-division multiplexing) has been adopted for the downlink of nascent wireless systems such as IEEE 802.16 [1]. It is also the multiplexing scheme of choice for the long-term evolution of 3rd Generation UMTS [2] and for most other 4th Generation proposals [3]. Being an orthogonal multiplexing technique, OFDM is particularly enticing in mobile systems where instantaneous CSIT (channel state information at the transmitter) may be unreliable or wholly unavailable. Moreover, OFDM is naturally well suited to deal with frequency selectivity [4]. With the sustained increase in signalling bandwidths, these features have justifiably positioned OFDM as a chief ingredient of most wireless systems in the making.

The present paper formulates the optimum (in the sense of maximizing the region of achievable ergodic mutual informations) power allocation policy for multiuser OFDM downlinks with statistical CSIT. The formulation places no constraints on either the partitioning of the available tones among users or on their modulation formats, which are not limited to ideal Gaussian signals. The counterpart of this problem with instantaneous CSIT is analyzed in [5]–[7] under various performance criteria. When the performance criterion is the mutual information, the optimal solution with instantaneous CSIT adopts the form of a multiuser mercury/waterfilling [8] that generalizes the single-user result given in [9], [10] (see also [11]). As we will see, however, with only statistical CSIT the mercury/waterfilling interpretation is lost.

II. MODELS AND DEFINITIONS

A. Multiuser OFDM

Consider a downlink channel whose bandwidth is partitioned into n orthogonal tones, sized such that each experiences (approximately) frequency-flat fading. A scalar signal is transmitted on every tone. A scheduler at the base station assigns each tone to one of k users, determines the signalling constellation to be used by each user on its assigned tones, and establishes user priorities from the nonnegative set $\{w_j\}_{j=1}^{k}$ such that

$$\sum_{j=1}^{k} w_j = 1. \tag{1}$$

The number of tones per user and the corresponding priorities may be assigned on the basis on the requirements of the respective types of traffic (rate, latency, etc).

Denoting by n_j the number of tones assigned to user j, the input-output relationship on the ith tone of the jth user is

$$Y_{i,j} = h_{i,j} X_{i,j} + W_{i,j} \qquad i = 1, \dots, n_j \quad j = 1, \dots, k \tag{2}$$

where $h_{i,j}$ is a complex gain while the noise $W_{i,j}$ is a zero-mean unit-variance complex Gaussian random variable independent of the noise on the other tones. Noting that the tones assigned to a given user may be nonadjacent, we define

$$\beta_j = \frac{n_j}{n} \tag{3}$$

as the fraction of the total bandwidth assigned to user j.

Since there is no instantaneous CSIT, the values of $\{w_j\}_{j=1}^{k}$ and $\{\beta_j\}_{j=1}^{k}$ cannot depend on the realization of $\{h_{i,j}\}$ but only on their distribution.

The complex signals $\{X_{i,j}\}$, zero-mean and mutually independent, must satisfy the power constraint

$$\frac{1}{n} \sum_{j=1}^{k} \sum_{i=1}^{n_j} E\left[|X_{i,j}|^2\right] \le P \tag{4}$$

where P is not a function of time. It is convenient to introduce normalized unit-power signals

$$S_{i,j} = \frac{X_{i,j}}{\sqrt{E\left[|X_{i,j}|^2\right]}}, \tag{5}$$

whose distribution is dictated by the modulation scheme used by the corresponding user, and a normalized power allocation

$$p_{i,j} = \frac{E\left[|X_{i,j}|^2\right]}{P}. \tag{6}$$

We can then define, for each tone, the channel state

$$\gamma_{i,j} = P\,|h_{i,j}|^2 \tag{7}$$

1-4244-0357-X/06/$20.00 ©2006 IEEE
This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE GLOBECOM 2006 proceedings.

which represents the receive signal-to-noise ratio on the ith tone of the jth user when the power allocation is uniform, i.e., $p_{i,j} = 1 \; \forall i, j$. (More generally, the receive signal-to-noise ratio is $p_{i,j}\gamma_{i,j}$.)

B. Fading Channels and Statistical CSIT

For every user j, the channel states $\{\gamma_{i,j}\}_{i=1}^{n_j}$ on the tones assigned to that user have an arbitrary joint distribution but the same marginal distribution, determined by the location of that particular user, and are known by its receiver. In particular,

$$\bar{\gamma}_j = E\left[\gamma_{i,j}\right] \qquad i = 1, \ldots, n_j \tag{8}$$

where $\bar{\gamma}_j$ is a measure of the local-average signal-to-noise ratio at the location of user j.

Because of their different locations, the channel states of different users are independent.

In the case of Rayleigh fading, which shall be invoked in many of our examples, every $\gamma_{i,j}$ has the exponential density

$$f_{\gamma_{i,j}}(\xi) = \frac{e^{-\xi/\bar{\gamma}_j}}{\bar{\gamma}_j} \qquad i = 1, \ldots, n_j. \tag{9}$$

Without instantaneous CSIT, the power allocation cannot depend on channel states but only on their distribution. Furthermore, since the fading distribution is identical on tones assigned to a given user, the constellation format on those tones must be equal and the power allocation must be uniform thereon. We emphasize the lack of instantaneous channel dependence on the power allocation by writing

$$\bar{p}_j = p_{i,j} \qquad i = 1, \ldots, n_j. \tag{10}$$

With that, (2) under coherent reception becomes equivalent to

$$Y_{i,j} = \sqrt{\gamma_{i,j}\,\bar{p}_j}\, S_{i,j} + W_{i,j} \tag{11}$$

subject to

$$\sum_{j=1}^{k} \beta_j \bar{p}_j \leq 1. \tag{12}$$

Note that the total power allocated to user j on its assigned tones equals $n_j \bar{p}_j$.

III. Mutual Information and MMSE

Our measure of performance is the mutual information, which specifies the maximum spectral efficiency achievable with arbitrary reliability for a given modulation format. Given a scalar Gaussian-noise channel $Y = \sqrt{\rho}\, S + W$, we denote its mutual information by $\mathcal{I}(\rho) = I(S; \sqrt{\rho}\, S + W)$ which is maximized when S is Gaussian, for which $\mathcal{I}(\rho) = \log(1+\rho)$ where the base of the logarithm determines the information units. While ideal, Gaussian signals cannot be realized in practice because of their continuous and unbounded support. Rather, signals usually conform to discrete constellations with limited peak-to-average ratios. For a discrete constellation (m-QAM, m-PSK, etc) consisting of m points, $\{s_\ell\}_{\ell=1}^m$, taken with probabilities $\{q_\ell\}_{\ell=1}^m$ such that $\sum_{\ell=1}^m q_\ell = 1$,

$$\mathcal{I}(\rho) = -\log(\pi e) - \int f_m(y,\rho) \log f_m(y,\rho)\, dy \tag{13}$$

where the integration extends to the complex plane while

$$f_m(y,\rho) = \frac{1}{\pi} \sum_{\ell=1}^{m} q_\ell \, e^{-|y - \sqrt{\rho}\, s_\ell|^2}. \tag{14}$$

A defining feature of any discrete constellation is the minimum distance between any two of its points, which we indicate by

$$d = \min_{\substack{k,\ell \\ k \neq \ell}} |s_k - s_\ell|. \tag{15}$$

For square m-QAM specifically,

$$d = \sqrt{\frac{6}{m-1}} \tag{16}$$

which gives $d = \sqrt{2}$ in the case of QPSK and $d = \sqrt{2/5}$ in the case of 16-QAM. For BPSK, in turn, $d = 2$.

The fact that, for non-Gaussian signals, the mutual information cannot in general be expressed in closed form greatly complicates optimization procedures that entail its differentiation. Propitiously, a recently unveiled relationship [12] affirms that, regardless of the distribution of the signal S,

$$\frac{d}{d\rho}\mathcal{I}(\rho) = \mathsf{MMSE}(\rho) \tag{17}$$

where $\mathcal{I}(\cdot)$ is in nats/s/Hz and the function $\mathsf{MMSE}(\cdot)$ returns the minimum mean-square error in estimating S by observing Y. As it is well known, this minimum mean-square error is achieved by the conditional-mean estimator

$$\hat{S}(y,\rho) = E\left[S | \sqrt{\rho}\, S + W = y; \rho\right] \tag{18}$$

which is, in general, a nonlinear function of the observation y. (It becomes linear if S is Gaussian.) Therefore,

$$\mathsf{MMSE}(\rho) = E\left[\left|S - \hat{S}(\sqrt{\rho}\, S + W, \rho)\right|^2\right] \tag{19}$$

with expectation over both S and W. Since S is unit power, $\mathsf{MMSE}(\cdot) \in [0,1]$. The inverse of $\mathsf{MMSE}(\cdot)$ with respect to the composition of functions is denoted by $\mathsf{MMSE}^{-1}(\cdot) \in [0,\infty)$.

For a Gaussian signal, (18) becomes

$$\hat{S}(y,\rho) = \frac{\sqrt{\rho}}{1+\rho}\, y \tag{20}$$

leading to $\mathsf{MMSE}(\rho) = 1/(1+\rho)$ and, in turn, to

$$\mathsf{MMSE}^{-1}(\zeta) = \frac{1}{\zeta} - 1. \tag{21}$$

For discrete constellations, (18) yields

$$\hat{S}(y,\rho) = \frac{\sum_{\ell=1}^m q_\ell\, s_\ell\, e^{-|y-\sqrt{\rho}\, s_\ell|^2}}{\sum_{\ell=1}^m q_\ell\, e^{-|y-\sqrt{\rho}\, s_\ell|^2}} \tag{22}$$

from which the $\mathsf{MMSE}(\cdot)$ follows via (19), which can be easily implemented as a low-pass filter driven by the estimation error $|S - \hat{S}|^2$. Alternatively, the $\mathsf{MMSE}(\cdot)$ can be tabulated and stored in memory for each of the constellations in use.

IV. Optimum Ergodic Power Allocation

We consider the channel states to vary ergodically over the codeword span. Note that, in general, codewords may extend to all of the tones assigned to a given user and thus the codewords have span in both time and frequency (and possibly in space if antenna diversity is used). Ergodic conditions are encountered under various operational situations:

- Fast fading, induced in the time domain by user motion and complemented by the use of hybrid ARQ (automatic repeat request).
- Frequency selectivity, induced by delay spread and reinforced by assigning non-adjacent tones to each user.
- Transmit and/or receive antenna diversity, which is not part of our formulation but could be easily incorporated.

Under such ergodic conditions, the key information-theoretic performance measure is the k-dimensional region containing the feasible k-tuples $\{\bar{\mathcal{R}}_1, \ldots, \bar{\mathcal{R}}_k\}$ where

$$\begin{aligned} \bar{\mathcal{R}}_j(\bar{p}_j) &= \frac{1}{n_j}\sum_{i=1}^{n_j} E\left[\mathcal{I}_j(\bar{p}_j \gamma_{i,j})\right] && (23)\\ &= E\left[\mathcal{I}_j(\bar{p}_j \gamma_{i,j})\right] && (24)\end{aligned}$$

is the ergodic mutual information attained by user j on its assigned tones and the index i in (24) corresponds to any of those tones. We thus seek the power allocation $\{\bar{p}_j^\star\}$ that solves

$$\{\bar{p}_j^\star\} = \arg \max_{\{\bar{p}_j\}: \sum_j \beta_j \bar{p}_j \leq 1} \sum_{j=1}^{k} w_j E\left[\mathcal{I}_j(\bar{p}_j \gamma_{i,j})\right]. \quad (25)$$

for all feasible $\{w_j\}_{j=1}^k$ thereby yielding the optimum boundary $\{\bar{\mathcal{R}}_1^\star, \ldots, \bar{\mathcal{R}}_k^\star\}$.

Theorem 1: The unique power allocation $\{\bar{p}_j^\star\}$ that solves (25) satisfies the necessary and sufficient conditions

$$\begin{aligned} \bar{p}_j^\star &= 0 && \bar{\gamma}_j \leq \frac{\beta_j}{w_j}\eta && (26)\\ E\left[\gamma_{i,j} \mathsf{MMSE}_j\left(\bar{p}_j^\star \gamma_{i,j}\right)\right] &= \frac{\beta_j}{w_j}\eta && \bar{\gamma}_j > \frac{\beta_j}{w_j}\eta && (27)\end{aligned}$$

with η such that

$$\sum_{j=1}^{k} \beta_j \, \bar{p}_j^\star = 1. \quad (28)$$

The strategy spelled by Theorem 1 is as follows. No power is allocated to the tones of users whose average channel state is below a threshold, $(\beta_j/w_j)\eta$, which is directly proportional to the bandwidth fraction of the corresponding user and inversely proportional to its priority. Active tones, in turn, are allocated the exact amount of power needed to equalize $E[\gamma_{i,j}\mathsf{MMSE}_j(\bar{p}_j^\star \gamma_{i,j})]$ for any of the tones $i = 1, \ldots, n_j$ at the threshold of the corresponding user j.

Computationally, finding $\{\bar{p}_j^\star\}$ requires solving a system of k nonlinear equations. This can be readily performed, e.g., via Broyden's method [13].

It is interesting to contrast Theorem 1 with its counterpart in the case that instantaneous CSIT is available, which adopts the form of a multiuser mercury/waterfilling [8]. Because of the expectations involved in the optimality conditions, when the CSIT is statistical rather than instantaneous the mercury/waterfilling interpretation (cf. [9]) is lost. Neither does the solution with statistical CSIT admit a waterfilling interpretation in the special case of ideal Gaussian signals, although in that case it is possible to manipulate the expressions into an alternative fixed-point form that is worth noting.

Corollary 1: With Gaussian signals,

$$\begin{aligned} \bar{p}_j^\star &= 0 && \bar{\gamma}_j \leq \frac{\beta_j}{w_j}\eta \\ \bar{p}_j^\star &= \frac{w_j}{\beta_j} \frac{1 - \overline{\mathsf{MMSE}}_j(\bar{p}_j^\star)}{\sum_{\kappa=1}^{k} w_\kappa \left(1 - \overline{\mathsf{MMSE}}_\kappa(\bar{p}_\kappa^\star)\right)} && \bar{\gamma}_j > \frac{\beta_j}{w_j}\eta \end{aligned}$$

where

$$\overline{\mathsf{MMSE}}_j(\bar{p}_j^\star) = E\left[\frac{1}{1 + \bar{p}_j^\star \gamma_{i,j}}\right]. \quad (29)$$

is the minimum mean-square error corresponding to a Gaussian signal, averaged over the fading distribution.

If the fading is Rayleigh distributed, we can use (9) to further express $\overline{\mathsf{MMSE}}_j(\cdot)$ as

$$\overline{\mathsf{MMSE}}_j(\bar{p}_j^\star) = \frac{\exp\left\{\frac{1}{\bar{p}_j^\star \bar{\gamma}_j}\right\} E_1\left(\frac{1}{\bar{p}_j^\star \bar{\gamma}_j}\right)}{\bar{p}_j^\star \bar{\gamma}_j} \quad (30)$$

where

$$E_1(\zeta) = \int_1^\infty t^{-1} e^{-\zeta t} dt \quad (31)$$

is an exponential integral.

Example 1: Consider an access point serving $k = 2$ users over Rayleigh-faded channels with equal bandwidth assigned to each user (i.e., $\beta_1 = \beta_2 = 1/2$) and with[1] $\bar{\gamma}_1|_{\mathrm{dB}} = \bar{\gamma}_2|_{\mathrm{dB}} = 10$. The optimum ergodic power allocation, as function of the priority w_1, is depicted in Fig. 1 parameterized by the constellation used by both users. The corresponding ergodic mutual information region boundaries are shown in Fig. 2.

Example 2: Consider the same scenario of Example 1, except with $\bar{\gamma}_1|_{\mathrm{dB}} = 10$ and $\bar{\gamma}_2|_{\mathrm{dB}} = 0$. The optimum ergodic power allocation and the ergodic mutual information region boundaries are shown in Figs. 3 and 4, respectively.

V. Limiting Power Regimes

A. Low-power regime

In the low-power regime, the optimum ergodic power allocation behaves as follows for quadrature-symmetric signals, a class that encompasses ideal Gaussian signals as well as discrete constellations such as m-QAM and m-PSK [14].

Proposition 1: For $P \to 0$, the optimum policy with statistical CSIT is to allocate power only to the tones assigned to the user(s) with the largest factor $w_j \bar{\gamma}_j / \beta_j$. If several users share the largest such product and the corresponding signals are quadrature symmetric, then power is uniformly distributed thereupon.

[1] $x|_{\mathrm{dB}} = 10 \log_{10} x$.

1-4244-0357-X/06/$20.00 ©2006 IEEE
This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE GLOBECOM 2006 proceedings.

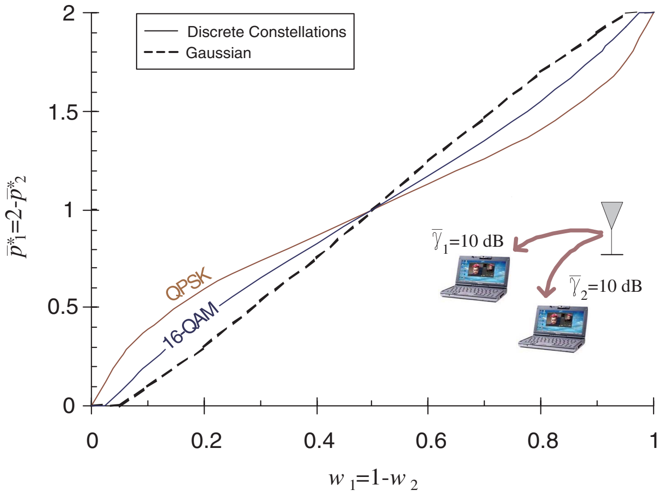

Fig. 1. Ergodic optimum power allocation with statistical CSIT, as function of the priority w_1, for $k = 2$ users having $\bar{\gamma}_1|_{\text{dB}} = \bar{\gamma}_2|_{\text{dB}} = 10$. Both channels are frequency-flat Rayleigh-faded with bandwidth partitioning $\beta_1 = \beta_2 = 1/2$. Continuous lines correspond to discrete constellations (QPSK and 16-QAM) while the dashed line corresponds to a Gaussian signal.

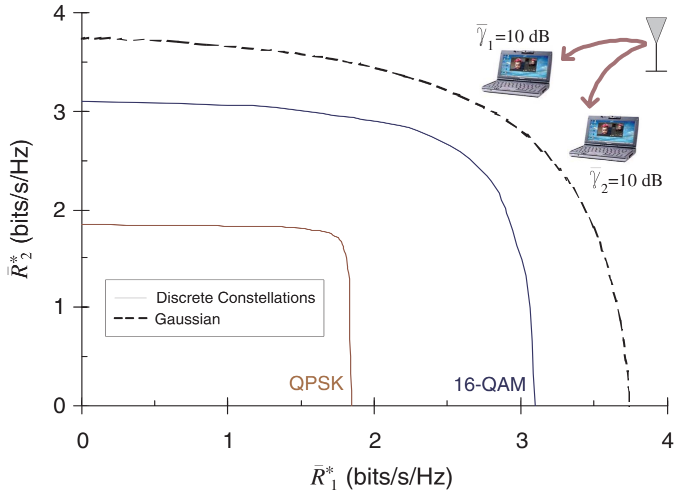

Fig. 2. Ergodic mutual information region with statistical CSIT for $k = 2$ users having $\bar{\gamma}_1|_{\text{dB}} = \bar{\gamma}_2|_{\text{dB}} = 10$. Both channels are frequency-flat Rayleigh-faded with bandwidth partitioning $\beta_1 = \beta_2 = 1/2$. Continuous lines correspond to discrete constellations (QPSK and 16-QAM) while the dashed line corresponds to a Gaussian signal.

B. High-power Regime

When the signals are ideal Gaussian, the leading term in the high-power expansion of the optimum ergodic power allocation depends only on the priority and bandwidth fraction of each user.

Proposition 2: If the signals are Gaussian, then for $P \to \infty$

$$\bar{p}_j^\star = \frac{w_j}{\beta_j} + \mathcal{O}(\frac{\log P}{P}). \tag{32}$$

This limiting behavior is already evident in Fig. 1, where despite the power level being only modestly high, the power allocation for Gaussian signals closely follows (32).

Although the corresponding behavior with discrete constellations is difficult to assess in full generality, for Rayleigh fading specifically the corresponding limit can be compactly characterized. To that end, let us define

$$\tilde{\gamma}_j = E[|h_{i,j}|^2] \qquad\qquad i = 1, \ldots, n_j \tag{33}$$

which, from (7) and (8), is tantamount to $\tilde{\gamma}_j = \bar{\gamma}_j / P$.

Proposition 3: If all the user signals conform to the same discrete constellation and the channels are Rayleigh-faded, then for $P \to \infty$

$$\bar{p}_j^\star = \alpha \sqrt{\frac{w_j}{\beta_j \tilde{\gamma}_j}} + \mathcal{O}(\frac{1}{P^2}) \tag{34}$$

with

$$\frac{1}{\alpha} = \sum_{j=1}^{k} \sqrt{\frac{w_j \beta_j}{\tilde{\gamma}_j}}. \tag{35}$$

Example 3: Let $k = 2$ with $\bar{\gamma}_1|_{\text{dB}} = \bar{\gamma}_2|_{\text{dB}} - 6$, with priorities $w_1 = 2/3$ and $w_2 = 1/3$, and with equal bandwidth assigned to each user (i.e., $\beta_1 = \beta_2 = 1/2$). Both users signal with the same constellation (either QPSK or 16-QAM). Shown in Fig. 5 are $\bar{p}_1^\star$ and $\bar{p}_2^\star$, obtained from Theorem 1, as function of $\bar{\gamma}_1$. For $P \to \infty$, the power allocation converges rapidly to its limit which, from Proposition 3, equals $\bar{p}_1^\star = 1.48$ and $\bar{p}_2^\star = 0.52$.

Note that the limiting values do not depend on the constellation cardinalities but the convergence does. (Naturally, with constellations of smaller cardinality the high-power behaviors are evidenced at lower power levels.) Notwithstanding such difference, in Example 3 the limit is closely approached at very modest power levels for both QPSK and 16-QAM.

TABLE I
LIMITING HIGH-POWER ALLOCATION IN EXAMPLE 4.

User (j)	$(\bar{\gamma}_j/\bar{\gamma}_1)|_{\text{dB}}$	$\bar{p}_j^\star$
1	0	0.43
2	-6	0.86
3	-12	1.71

Example 4: Consider $k = 3$ users with equal priorities and bandwidth fractions (i.e., $w_1 = w_2 = w_3 = 1/3$ and $\beta_1 = \beta_2 = \beta_3 = 1/3$) and whose signals conform to the same discrete constellation. The fading of the first user is, on average, 6 dB stronger than that of the second and 12 dB stronger than that of the third. The limiting power allocation for $P \to \infty$ is listed in Table I. Note the high nonuniformity of this allocation: the first user is allocated 3 dB less power than the second and 6 dB less power than the third. If the users where employing ideal Gaussian signals instead of discrete constellations, in contrast, the limiting power allocation would be uniform.

VI. SUMMARY

We have formulated the optimum (in the sense of maximizing the region of ergodic mutual informations for any

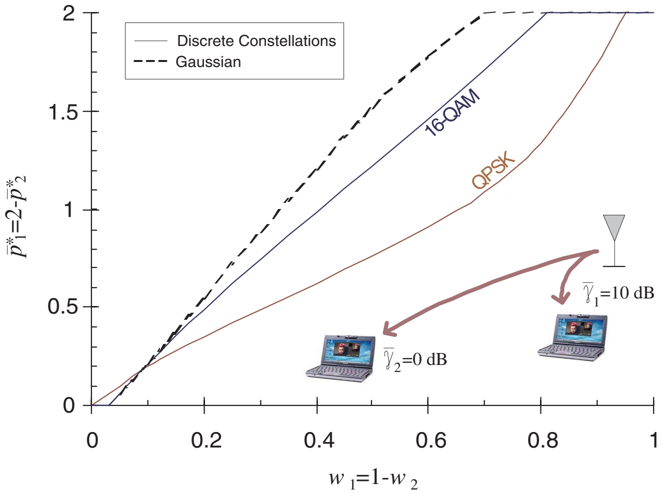

Fig. 3. Ergodic optimum power allocation with statistical CSIT, as function of the priority w_1, for $k = 2$ users having $\bar{\gamma}_1|_{\mathrm{dB}} = 0$ and $\bar{\gamma}_2|_{\mathrm{dB}} = 10$. Both channels are frequency-flat Rayleigh-faded. Continuous lines correspond to discrete constellations (QPSK and 16-QAM) while the dashed line corresponds to a Gaussian signal.

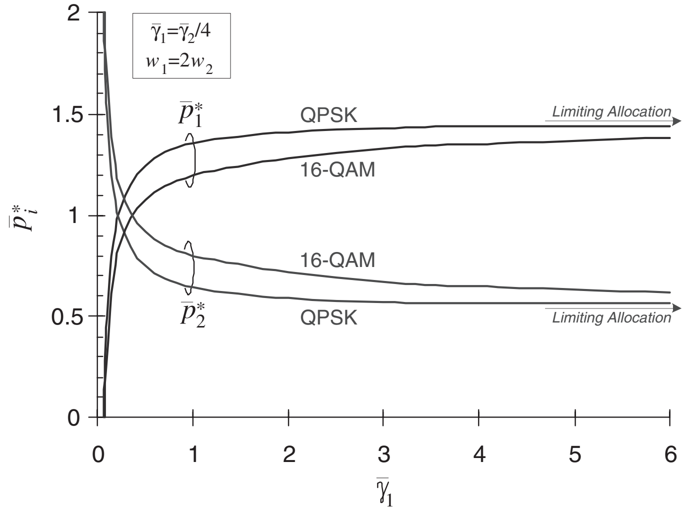

Fig. 5. For $k = 2$ with $\bar{\gamma}_1|_{\mathrm{dB}} = \bar{\gamma}_2|_{\mathrm{dB}} - 6$ and $w_1 = 2w_2$, exact $\bar{p}_1^\star$ and $\bar{p}_2^\star$ as function of $\bar{\gamma}_1$ and parameterized by the constellation used by both users (QPSK or 16-QAM). The limiting power allocation for $P \to \infty$ is also indicated.

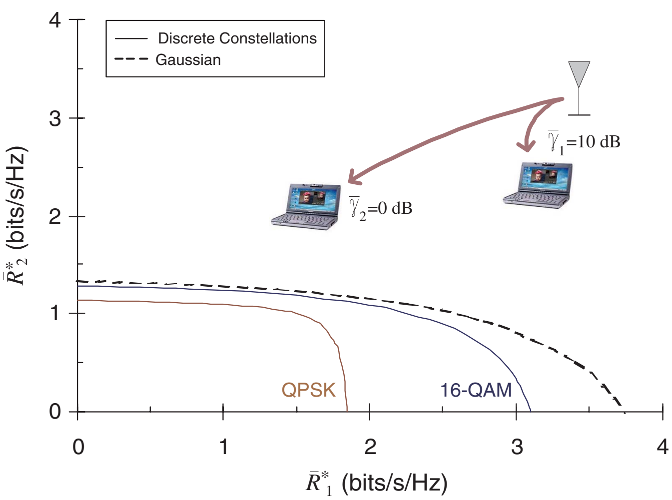

Fig. 4. Ergodic mutual information region with statistical CSIT for $k = 2$ users having $\bar{\gamma}_1|_{\mathrm{dB}} = 0$ and $\bar{\gamma}_2|_{\mathrm{dB}} = 10$. Both channels are frequency-flat Rayleigh-faded. Continuous lines correspond to discrete constellations (QPSK and 16-QAM) while the dashed line corresponds to a Gaussian signal.

partitioning of the available tones) power allocation policy for multiuser OFDM downlinks with statistical CSIT. This policy holds for arbitrary modulation formats, possibly different for every user. It further holds, except for a periodic re-mapping of the tone indices, if frequency hopping is used to randomize the tone assignments.

References

[1] *IEEE Standard for Local and Metropolitan Area Networks—Part 16: Air Interface for Fixed Broadband Wireless Access Systems*, IEEE 802 LAN/MAN Standards Committee, Nov. 2005.

[2] *UTRA-UTRAN Long Term Evolution (LTE)*, 3rd Generation Partnership Project (3GPP), Nov. 2004.

[3] R. Tafazolli, *Technologies for the Wireless Future: Wireless World Research Forum*. New York, Wiley, 2004.

[4] J. A. C. Bingham, "Multicarrier modulation for data transmission: An idea whose time has come," *IEEE Commun. Magazine*, vol. 28, no. 5, pp. 5–14, May 1990.

[5] C. Y. Wong, R. S. Cheng, K. B. Letaief, and R. D. Murch, "Multiuser OFDM with adaptive subcarrier, bit, and power allocation," *IEEE J. Sel. Areas in Communications*, vol. 17, no. 10, pp. 1747–1758, Oct. 1999.

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

[7] S. Z. Shen, J. G. Andrews, and B. L. Evans, "Optimal power allocation in multiuser OFDM systems," in *IEEE Global Telecommunications Conference (GLOBECOM'03)*, vol. 1, Dec. 2003, pp. 337–341.

[8] A. Lozano, A. M. Tulino, and S. Verdú, "Multiuser mercury/waterfilling for downlink OFDM with arbitrary signal constellations," in *IEEE Int'l Symp. on Spread Spectrum Techn. & Applications (ISSSTA'06)*, Aug. 2006.

[9] ——, "Mercury/waterfilling: Optimum power allocation with arbitrary input constellations," *Proc. of IEEE Intern. Symp. on Inform. Theory (ISIT'05)*, Sept. 2005.

[10] ——, "Optimum power allocation for parallel Gaussian channels with arbitrary input distributions," *IEEE Trans. on Inform. Theory*, vol. 52, no. 7, pp. 3033–3051, July 2006.

[11] M. Taouk, M. J. M. Peacock, and I. B. Collings, "Statistical power allocation and coded bit allocation optimization in mercury/waterfilling," in *IEEE 7th Australian Communications Theory Workshop*, Feb. 2006, pp. 157–162.

[12] D. Guo, S. Shamai, and S. Verdú, "Mutual information and minimum mean-square error in Gaussian channels," *IEEE Trans. on Inform. Theory*, vol. 51, no. 4, pp. 1261–1283, Apr. 2005.

[13] C. G. Broyden, "A class of methods for solving nonlinear simultaneous equations," *Mathematics of Computation*, vol. 19, pp. 577–593, 1965.

[14] S. Verdú, "Spectral efficiency in the wideband regime," *IEEE Trans. on Inform. Theory*, vol. 48, no. 6, pp. 1319–1343, June 2002.