Optimum Power Allocation for Multiuser OFDM with Arbitrary Signal Constellations

Angel Lozano,[*] Antonia M. Tulino,[†] and Sergio Verdú[‡]

January 25, 2012

Abstract

This paper formulates power allocation policies that maximize the region of mutual informations achievable in multiuser downlink OFDM channels. Arbitrary partitioning of the available tones among users and arbitrary modulation formats, possibly different for every user, are considered. Two distinct policies are derived, respectively for slow fading channels tracked instantaneously by the transmitter and for fast fading channels known only statistically thereby. With instantaneous channel tracking, the solution adopts the form of a multiuser mercury/waterfilling procedure that generalizes the single-user mercury/waterfilling introduced in [1, 2]. With only statistical channel information, in contrast, the mercury/waterfilling interpretation is lost. For both policies, a number of limiting regimes are explored and illustrative examples are provided.

Keywords: Gaussian channels; Power allocation; Waterfilling; Mercury/waterfilling; Mutual information; MMSE; OFDM; Downlink.

[*]Angel Lozano is with Bell Laboratories (Lucent Technologies), Holmdel, NJ07733, USA.

[†]Antonia M. Tulino is with Universita Degli Studi di Napoli "Federico II", 80125 Napoli, Italy

[‡]Sergio Verdú is with Princeton University, Princeton, NJ08544, USA.

I Introduction

There is, of late, great interest in OFDM (orthogonal frequency-division multiplexing) for multiuser wireless downlinks [3]. Although in general suboptimal in the face of instantaneous CSIT (channel state information at the transmitter), orthogonal multiplexing techniques are much more robust to CSIT inaccuracies than nonorthogonal schemes. Furthermore, OFDM has the added benefit of being naturally well suited to deal with frequency selectivity [4]. With the continued increase in signal bandwidths, OFDM is poised to be a central ingredient of most wireless systems to come [5, 6, 7].

The present paper formulates the optimum power allocation policy for multiuser OFDM downlinks with either instantaneous or statistical CSIT. These policies, which maximize the respective mutual information regions with no constraints on either the partitioning of the available tones among users or on their modulation formats, extends to the multiuser realm the single-user mercury/waterfilling policy presented in [1, 2] (see also [8]) enabling:

- Determination of the region of spectral efficiencies reliably achievable for any partition of the available tones and any modulation format.

- A benchmark against which suboptimal power allocation policies can be gauged.

- Assessment of the fundamental advantage of allocating power on the basis of either instantaneous or statistical CSIT.

For criteria other than the mutual information, the allocation of power in multiuser OFDM downlink channels is analyzed in [9]–[12].

II Models and Definitions

A Multiuser OFDM

Consider a downlink channel 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 active users, determines the signalling constellation to be used by each such user on its assigned tones, and establishes user priorities from the nonnegative set $\{w_j\}_{j=1}^K$ such that

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

so as to maximize the sum of the individual spectral efficiencies weighted by $\{w_j\}_{j=1}^K$.

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

$$Y_{i,j} = h_{i,j} X_{i,j} + W_{i,j} \qquad i = 1, \ldots, n_j \qquad j = 1, \ldots, K \tag{2}$$

where $X_{i,j}$ and $Y_{i,j}$ are the transmit and received complex signals, $h_{i,j}$ is a complex channel gain, and 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

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

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

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

$$\frac{1}{n} \sum_{j=1}^K \sum_{i=1}^{n_j} E\left[|X_{i,j}|^2\right] \le P \tag{3}$$

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

$$S_{i,j} = \frac{X_{i,j}}{\sqrt{E\left[|X_{i,j}|^2\right]}}, \tag{4}$$

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

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

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

$$\gamma_{i,j} = P\,|h_{i,j}|^2 \tag{6}$$

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 $p_{i,j}\gamma_{i,j}$.) With that, (2) under coherent reception is equivalent to

$$Y_{i,j} = \sqrt{\gamma_{i,j}\, p_{i,j}}\, S_{i,j} + W_{i,j} \qquad i = 1, \ldots, n_j \qquad j = 1, \ldots, K \tag{7}$$

subject to

$$\frac{1}{n} \sum_{j=1}^{K} \sum_{i=1}^{n_j} p_{i,j} \leq 1. \tag{8}$$

B   Fading Channels and CSIT

We consider two distinct fading scenarios, each with an associated level of CSIT:

- Slow fading, whereby $\{\gamma_{i,j}\}$ do not change appreciably during the span of a codeword. The transmitter can track the fading states and thus we associate this scenario with instantaneous CSIT. This is a valid model, e.g., for fixed broadband access and pedestrian traffic.

- Fast fading, with $\{\gamma_{i,j}\}$ exhibiting rapid variations within the span of a codeword. The fading states cannot be tracked by the transmitter, which is only privy to their distribution. The CSIT is thus statistical. This is a satisfactory model for vehicular traffic.

In Sections C and V, we need to model the statistical behavior of the channel coefficients. For every user j, the channel states $\{\gamma_{i,j}\}_{i=1}^{n_j}$ on the tones assigned to that user have the same marginal distribution, and are known by its receiver. In particular,

$$E\left[\gamma_{i,j}\right] = \bar{\gamma}_j \qquad\qquad i = 1, \ldots, n_j \tag{9}$$

where $\bar{\gamma}_j$ is a measure of the local-average signal-to-noise ratio at the location occupied by user j. The $\{\gamma_{i,j}\}_{i=1}^{n_j}$ for user j are mutually independent if the frequency separation of the corresponding tones exceeds some finite coherence bandwidth. Moreover, $\{\gamma_{i,j}\}_{i=1}^{n_j}$ are independent from their counterparts for all other users. It is assumed that $\gamma_{i,j} > 0$ with probability 1 (no mass points at zero), a very mild condition satisfied by every fading

4

distribution of practical interest. In the case of Rayleigh fading, which shall be invoked in many of our examples, each $\gamma_{i,j}$ has an exponential density

$$f_{\gamma_{i,j}}(\xi) = \frac{e^{-\xi/\bar{\gamma}_j}}{\bar{\gamma}_j} \qquad\qquad i = 1, \ldots, n_j. \tag{10}$$

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 of the form $Y = \sqrt{\rho}\, S + W$, 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 capacity-achieving, 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-PAM, m-QAM, m-PSK, etc) consisting of m points, $\{s_\ell\}_{\ell=1}^m$, taken with probabilities $\{q_\ell\}_{\ell=1}^m$ such that $\sum_{\ell=1}^m q_\ell = 1$,

$$\mathcal{I}(\rho) = -\log(\pi e) - \int f_m(y, \rho) \log f_m(y, \rho)\, dy \tag{11}$$

where the integration extends to the complex plane and the base of the logarithms determines the information units while

$$f_m(y, \rho) = \frac{1}{\pi} \sum_{\ell=1}^m q_\ell\, e^{-|y - \sqrt{\rho}\, s_\ell|^2}. \tag{12}$$

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

$$d = \min_{\substack{k,\ell \\ k \neq \ell}} |s_k - s_\ell|. \tag{13}$$

For square m-QAM specifically,

$$d = \sqrt{\frac{6}{m - 1}} \tag{14}$$

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 [13] affirms that, regardless of the distribution of the signal S,

$$\frac{d}{d\rho}\mathcal{I}(\rho) = \mathsf{MMSE}(\rho) \tag{15}$$

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. This minimum mean-square error is achieved by the conditional-mean estimator

$$\hat{S}(y, \rho) = E\left[S | \sqrt{\rho}\, S + W = y; \rho\right] \tag{16}$$

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

$$\mathsf{MMSE}(\rho) = E\left[\left|S - \hat{S}(\sqrt{\rho}\, S + W, \rho)\right|^2\right] \tag{17}$$

with expectation over both S and W. Since $E[|S|^2] = 1$, $\mathsf{MMSE}(\cdot) \in [0, 1]$. 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, (16) becomes

$$\hat{S}(y, \rho) = \frac{\sqrt{\rho}}{1 + \rho}\, y \tag{18}$$

leading to

$$\mathsf{MMSE}(\rho) = \frac{1}{1 + \rho} \tag{19}$$

and, in turn, to

$$\mathsf{MMSE}^{-1}(\zeta) = \frac{1}{\zeta} - 1. \tag{20}$$

For discrete constellations, (16) yields

$$\hat{S}(y, \rho) = \frac{\sum_{\ell=1}^{m} q_\ell\, s_\ell\, e^{-|y - \sqrt{\rho}\, s_\ell|^2}}{\sum_{\ell=1}^{m} q_\ell\, e^{-|y - \sqrt{\rho}\, s_\ell|^2}} \tag{21}$$

from which the $\mathsf{MMSE}(\cdot)$ follows via (17), 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 Slow Fading and Instantaneous CSIT

A Multiuser Mercury/waterfilling

In this slow fading scenario, recall, the channel states remain fixed over the duration of each codeword and are known by the transmitter. The power allocation is thus a function of $\{\gamma_{i,j}\}$. Every realization thereof gives rise to a k-dimensional region containing all of the feasible k-tuples $\{\mathcal{R}_1, \ldots, \mathcal{R}_k\}$ where[1]

$$\mathcal{R}_j(p_{1,j}, \ldots, p_{n_j,j}) = \frac{1}{n_j} \sum_{i=1}^{n_j} \mathcal{I}_j(p_{i,j}\gamma_{i,j}) \tag{22}$$

is the mutual information attained by user j on its assigned tones with the function $\mathcal{I}_j(\cdot)$ reflecting the signalling scheme used by that user. The boundary of this region can be fully characterized by means of the weighted function $\sum_{j=1}^{K} w_j \mathcal{R}_j(\cdot)$ for all priorities $\{w_j\}_{j=1}^{K}$ satisfying (1). We thus seek the power allocation $\{p_{i,j}^\star\}$ that solves

$$\{p_{i,j}^\star\} = \arg \max_{\{p_{i,j}\}: \frac{1}{n} \sum_{i,j} p_{i,j} \leq 1} \frac{1}{n} \sum_{j=1}^{K} \frac{w_j}{\beta_j} \sum_{i=1}^{n_j} \mathcal{I}_j(p_{i,j}\gamma_{i,j}). \tag{23}$$

yielding the optimum boundary $\{\mathcal{R}_1^\star, \ldots, \mathcal{R}_K^\star\}$.

Theorem 1 *With instantaneous CSIT, the unique power allocation that solves (23) is*

$$p_{i,j}^\star = 0 \qquad\qquad \gamma_{i,j} \leq \frac{\beta_j}{w_j}\eta \tag{24}$$

$$p_{i,j}^\star = \frac{1}{\gamma_{i,j}} \mathsf{MMSE}_j^{-1}\left(\frac{\beta_j \eta}{w_j \gamma_{i,j}}\right) \qquad\qquad \gamma_{i,j} > \frac{\beta_j}{w_j}\eta \tag{25}$$

with η such that (8) is met with strict equality.

Proof: *See Appendix A.*

The strategy spelled by Theorem 1 is as follows. No power is allocated to those tones whose channel state is below a threshold, $(\beta_j/w_j)\eta$, that is directly proportional to the

---

[1]For notational simplicity, we drop the explicit dependence of the spectral efficiencies on the channel coefficients.

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 render $\gamma_{i,j}\mathsf{MMSE}_j(p^\star_{i,j}\gamma_{i,j})$ equal to the threshold of their respective users.

Computationally, Theorem 1 boils down to solving a single nonlinear equation on η, from which $\{p^\star_{i,j}\}$ are then simply mapped via (24) and (25).

In order to provide a graphical interpretation, let us define the auxiliary function

$$G_j(\zeta) = \begin{cases} 1/\zeta - \mathsf{MMSE}_j^{-1}(\zeta) & \zeta \in [0,1] \\ 1 & \zeta > 1 \end{cases} \tag{26}$$

which, for a Gaussian signal, reduces to $G_j(\zeta) = 1$. Theorem 1 is tantamount to the following multiuser mercury/waterfilling procedure (cf. Fig. 1):

(a) For each of the n tones, set up a vessel of base $(w_j/\beta_j) \times 1$, solid up to a height $1/\gamma_{i,j}$.

(b) Choose η. Pour mercury onto each of the vessels until its height (including the solid) reaches $\frac{\beta_j}{w_j\gamma_{i,j}}G_j(\frac{\beta_j\eta}{w_j\gamma_{i,j}})$.

(c) Waterfill, keeping a common upper level of water, until the water level reaches $1/\eta$.

(d) The volume of water in the (i,j)th vessel gives $p_{i,j}$.

As in single-user mercury/waterfilling, the amount of poured mercury regulates the water admitted by each vessel thereby accounting for the respective signalling constellations. The new feature in multiuser mercury/waterfilling is the variability in vessel widths, which also regulates the water admission but in relation with the user priorities and assigned bandwidths. Since no mercury is poured onto vessels whose signal is Gaussian, multiuser mercury/waterfilling reverts to a multiuser waterfilling whenever all of the signals are Gaussian. Indeed, plugging (20) into Theorem 1 yields

$$p^\star_{i,j} = 0 \qquad\qquad \gamma_{i,j} \le \frac{\beta_j}{w_j}\eta \tag{27}$$

$$p^\star_{i,j} = \frac{w_j/\beta_j}{\eta} - \frac{1}{\gamma_{i,j}} \qquad\qquad \gamma_{i,j} > \frac{\beta_j}{w_j}\eta \tag{28}$$

In this particular case, we can invoke (19) to further obtain the following alternative form for the solution.

Corollary 1 *With Gaussian signals,*

$$p_{i,j}^\star = 0 \qquad\qquad \gamma_{i,j} \leq \frac{\beta_j}{w_j}\eta \qquad (29)$$

$$p_{i,j}^\star = \frac{\frac{w_j}{\beta_j}\left(1 - \mathsf{MMSE}_j(p_{i,j}^\star \gamma_{i,j})\right)}{\frac{1}{n}\sum_{\kappa=1}^{K}\frac{w_\kappa}{\beta_\kappa}\sum_{\ell=1}^{n_\kappa}\left(1 - \mathsf{MMSE}_\kappa(p_{\ell,\kappa}^\star \gamma_{\ell,\kappa})\right)} \qquad \gamma_{i,j} > \frac{\beta_j}{w_j}\eta \qquad (30)$$

Although in principle less appealing than (27)–(28), the fixed-point characterization in Corollary 1 has the advantage of generalizing to the case of nonorthogonal parallel channels [14].

B Low- and High-Power Regimes

In the low-power regime, a defining feature of a constellation is quadrature symmetry, which ensures first- and second-order optimality in terms of low-power mutual information [15]. The class of quadrature-symmetric signals (i.e., those that can be built by superimposing scaled and/or rotated QPSK constellations) encompasses ideal Gaussian signals as well as discrete constellations such as m-QAM and m-PSK.

Using the tools put forth in [15], we can synthesize the low-power behavior of multiuser mercury/waterfiling as follows.

Proposition 1 *For $P \to 0$, multiuser mercury/waterfilling allocates power only to the tone(s) with the largest factor $w_j\gamma_{i,j}/\beta_j$. If several tones share the largest such factor and the corresponding signals are quadrature-symmetric, then the power is uniformly distributed thereon.*

In contrast with the low-power regime, where two orders of optimality in terms of mutual information boil down to quadrature symmetry only, in the high-power regime the suboptimality of discrete constellations become overtly manifest.

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

$$p_{i,j}^\star = \frac{w_j}{\beta_j} + \mathcal{O}(\frac{1}{P}) \qquad\qquad i = 1,\ldots,n_j \qquad (31)$$

whereas, if the signals conform to discrete constellations with minimum distance d_j for user j, then

$$p_{i,j}^{\star} = \frac{\alpha}{|h_{i,j}|^2 d_j^2} + \mathcal{O}(\frac{\log P}{P}) \tag{32}$$

with

$$\frac{1}{\alpha} = \frac{1}{n} \sum_{\kappa=1}^{K} \sum_{\ell=1}^{n_\kappa} \frac{1}{|h_{\ell,\kappa}|^2 d_\kappa^2}. \tag{33}$$

Proof: *See Appendix B.*

Notice the discrepancy between the leading terms in the high-power expansion with ideal Gaussian signals and with discrete constellations. With the former, the power allocated to a tone is dominated by the priority and bandwidth fraction of the respective user. With the latter, in contrast, the user priority and bandwidth fraction become immaterial for large P. Only the channel states and the constellation minimum distances are of essence, with more power allocated to tones whose users are employing richer constellations but with less power on tones with stronger channel states.

The above discrepancy has its roots in the fact that, while for a Gaussian signal the mutual information grows without bound with the input power, for a discrete constellation it is bounded.

The cardinality of the constellation critically determines the point where the high-power behavior sets in [2]. The smaller the constellation, the less power required to bring it about.

C Ergodic Characterization

The optimum power allocation $\{p_{i,j}^{\star}\}$ and the corresponding user mutual informations $\{\mathcal{R}_j^{\star}\}_{j=1}^{K}$ can be regarded as random variables whose distributions are induced by the channel states, $\{\gamma_{i,j}\}$. For delay-tolerant applications, nonetheless, the time averages of the mutual informations acquire operational significance. Under ergodic fading, these time averages equal $\bar{\mathcal{R}}_j^{\star} = E[\mathcal{R}_j^{\star}]$. The corresponding average power allocation $\bar{p}_j^{\star} = E[p_{i,j}^{\star}]$, $i = 1, \ldots, n_j$, meaningfully conveys how the power is allocated on average. Note

that this average allocation is common to all the tones of a given user because of their identical marginal fading distribution and equal signalling constellation.

Example 1 *Consider an access point streaming data to $K = 2$ users over respective frequency-flat Rayleigh-faded channels with equal bandwidth assigned to each user (i.e., $\beta_1 = \beta_2 = 1/2$) and with $\bar{\gamma}_1|_{dB} = \bar{\gamma}_2|_{dB} = 5$.[2] The average multiuser mercury/waterfilling power allocation as function of the priority w_1, is depicted in Fig. 2 parameterized by the constellation used by both users. The corresponding ergodic mutual information region boundaries are shown in Fig. 3.*

With the average conditions in Example 1, the channel states are frequently in a range where 16-QAM is rich enough to resemble an ideal Gaussian signal whereas QPSK is not.

Example 2 *Consider the same scenario of Example 1, except with $\bar{\gamma}_1|_{dB} = 10$ and $\bar{\gamma}_2|_{dB} = 0$. The average multiuser mercury/waterfilling power allocation and the ergodic mutual information region boundaries are shown in Figs. 4 and 5, respectively.*

In Example 2, the various signalling constellations behave similarly whenever user $j = 2$ is prioritized because it is often in low-power conditions. When user $j = 1$ is favored, however, there is a large disparity between the power allocation and mutual informations for the several signalling formats.

Another instance in which, notwithstanding the delay tolerances, the average mutual informations provide a meaningful characterization is that of strong frequency selectivity per user, whereby (22) by itself provides an effective averaging mechanism. To gain insight, we consider the limiting regime where $n_j \to \infty$, $j = 1, \ldots, K$.[3] From the asymptotic independence of the channel states for each user, the $\{\mathcal{R}_j^\star\}$ converge in the mean-square sense to nonrandom limits that depend only on the signalling constellations and fading distributions:

$$\mathcal{R}_j^\star \to \int_{\frac{\beta_j}{w_j}\eta}^{\infty} \mathcal{I}_j \left(\mathsf{MMSE}^{-1} \left(\frac{\beta_j}{w_j} \frac{\eta}{\xi} \right) \right) f_{\gamma_{i,j}}(\xi) \, d\xi \tag{34}$$

---

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

[3]Note that, by virtue of (8), the total transmit power is also growing without bound as the system bandwidth increases. (If the bandwidth were increased with the total transmit power held fixed, the system would be pushed into the low-power regime and Proposition 1 would apply asymptotically.)

with η the solution of

$$\sum_{j=1}^{K} \beta_j \int_{\frac{\beta_j}{w_j}\eta}^{\infty} \frac{1}{\xi} \mathsf{MMSE}_j^{-1}\left(\frac{\beta_j}{w_j}\frac{\eta}{\xi}\right) f_{\gamma_{i,j}}(\xi)\, d\xi = 1. \tag{35}$$

The tone powers $\{p_{i,j}^{\star}\}$ remain random, but their empirical distributions for the various users converge in probability to nonrandom limits that can be found from η via (24)–(25).

Specifically for Gaussian signals and Rayleigh fading, we can invoke (10) and (20) to obtain more explicit versions of (35) and (34). Under these conditions

$$\mathcal{R}_j \to E_1\left(\frac{\beta_j}{w_j \bar{\gamma}_j}\eta\right) \qquad\qquad j = 1, \ldots, K \tag{36}$$

in nats/s/Hz, with

$$E_1(\zeta) = \int_{1}^{\infty} t^{-1} e^{-\zeta t} dt \tag{37}$$

an exponential integral and with η the solution to

$$\sum_{j=1}^{K} \left(\frac{w_j\, e^{-\frac{\beta_j}{w_j\bar{\gamma}_j}\eta}}{\eta} - \frac{\beta_j}{\bar{\gamma}_j} E_1\left(\frac{\beta_j}{w_j\bar{\gamma}_j}\eta\right) \right) = 1. \tag{38}$$

The expressions in (36) and (38), which are multiuser extensions of the single-user result found in [16], can be used to upper-bound the benefits that can be reaped—in Rayleigh fading—from exploiting frequency diversity with instantaneous CSIT.

V   Fast Fading and Statistical CSIT

A   Optimum Ergodic Power Allocation

In this fast-fading scenario, the channel states vary widely and ergodically over the code-word span. The transmitter is only privy to their statistical distributions and, therefore, the ergodic power allocation cannot be a function of $\{\gamma_{i,j}\}$ but only of their distributions. 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 can thus drop the tone index from the power allocation and emphasize the lack of instantaneous channel dependence on the power allocation by writing $\{\bar{p}_j\}_{j=1}^K$ to denote the ergodic power allocation. The performance measure is the K-dimensional region containing the feasible K-tuples $\{\bar{\mathcal{R}}_1, \ldots, \bar{\mathcal{R}}_K\}$ where

$$\bar{\mathcal{R}}_j(\bar{p}_j) \;=\; \frac{1}{n_j} \sum_{i=1}^{n_j} E\left[\mathcal{I}_j(\bar{p}_j \gamma_{i,j})\right] \tag{39}$$

$$\;=\; E\left[\mathcal{I}_j(\bar{p}_j \gamma_{i,j})\right] \tag{40}$$

is the ergodic mutual information attained by user j on its assigned tones and the index i in (40) corresponds to any of the those tones. We thus seek the power allocation $\{\bar{p}_j^\star\}$ that solves

$$\{\bar{p}_j^\star\} = \arg \max_{\{\bar{p}_j\}:\frac{1}{n}\sum_j n_j \bar{p}_j \leq 1} \sum_{j=1}^K w_j E\left[\mathcal{I}_j(\bar{p}_j \gamma_{i,j})\right]. \tag{41}$$

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 2 *With statistical CSIT, the unique power allocation that solves (41) satisfies the necessary and sufficient conditions*

$$\bar{p}_j^\star \;=\; 0 \qquad\qquad \bar{\gamma}_j \leq \frac{\beta_j}{w_j}\eta \tag{42}$$

$$E\left[\gamma_{i,j}\mathsf{MMSE}_j\left(\bar{p}_j^\star \gamma_{i,j}\right)\right] \;=\; \frac{\beta_j}{w_j}\eta \qquad\qquad \bar{\gamma}_j > \frac{\beta_j}{w_j}\eta \tag{43}$$

with η such that

$$\sum_{j=1}^K \beta_j\, \bar{p}_j^\star = 1. \tag{44}$$

Proof: *See Appendix C.*

The optimum strategy with statistical CSIT is thus 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. The active tones for user j are allocated the exact amount of power needed to make $E[\gamma_{i,j}\mathsf{MMSE}_j(\bar{p}_j^\star \gamma_{i,j})]$ (which does not depend on i) equal to $\beta_j\eta/w_j$.

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 [17]. Because of the expectations

involved in the optimality conditions, however, the mercury/waterfilling interpretation is lost.

Neither does the solution admit a waterfilling interpretation in the special case of Gaussian signals, although in that case it is possible—as it was with instantaneous CSIT—to manipulate the expressions into a fixed-point form.

Corollary 2 *With Gaussian signals,*

$$\bar{p}_j^\star = 0 \qquad\qquad \bar{\gamma}_j \leq \frac{\beta_j}{w_j}\eta \qquad (45)$$

$$\bar{p}_j^\star = \frac{w_j}{\beta_j} \frac{1 - \overline{\mathsf{MMSE}}_j(\bar{p}_j^\star)}{\sum_{\kappa=1}^K w_\kappa \left(1 - \overline{\mathsf{MMSE}}_\kappa(\bar{p}_\kappa^\star)\right)} \qquad\qquad \bar{\gamma}_j > \frac{\beta_j}{w_j}\eta \qquad (46)$$

where

$$\overline{\mathsf{MMSE}}_j(\bar{p}_j^\star) = E\left[\frac{1}{1 + \bar{p}_j^\star \gamma_{i,j}}\right]. \qquad (47)$$

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 (10) to further express $\overline{\mathsf{MMSE}}_j(\cdot)$ as

$$\overline{\mathsf{MMSE}}_j(\bar{p}_j^\star) = \frac{\exp\left\{\frac{1}{\bar{p}_j^\star \bar{\gamma}_j}\right\} E_1\left(\frac{1}{\bar{p}_j^\star \bar{\gamma}_j}\right)}{\bar{p}_j^\star \bar{\gamma}_j} \qquad (48)$$

where $E_1(\cdot)$ is the exponential integral in (37). Furthermore, from (40) and (10), the corresponding ergodic mutual information achieved by the jth user with ideal Gaussian signalling and Rayleigh fading is then

$$\bar{\mathcal{R}}_j^\star(\bar{p}_j^\star) = \exp\left\{\frac{1}{\bar{p}_j^\star \bar{\gamma}_j}\right\} E_1\left(\frac{1}{\bar{p}_j^\star \bar{\gamma}_j}\right). \qquad (49)$$

A similar expression for $\bar{\mathcal{R}}_j(\cdot)$ under uniform power allocation has been given in [18, 19].

Example 3 *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 $\bar{\gamma}_1|_{\mathrm{dB}} = 10$ and $\bar{\gamma}_2|_{\mathrm{dB}} = 0$. The optimum ergodic power allocation with statistical CSIT, as function of the priority w_1, is depicted in Fig. 6 parameterized by the constellation used by both users. The corresponding ergodic mutual information region boundaries are shown in Fig. 7.*

14

Notice, by comparing Figs. 5 and 7, that in this particular scenario there is no significant loss in having only statistical CSIT as opposed to instantaneous CSIT.

B Low- and High-Power Regimes

The low-power behavior with statistical CSIT bears close resemblance to its instantaneous CSIT counterpart.

Proposition 3 *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.*

Also the high-power allocation with Gaussian signals closely parallels its analogue with instantaneous CSIT, with the leading term dependent only on the priority and bandwidth fraction of each user.

Proposition 4 *If the signals are Gaussian, then for $P \to \infty$*

$$\bar{p}_j^\star = \frac{w_j}{\beta_j} + \mathcal{O}(\frac{\log P}{P}). \tag{50}$$

Proof: *See Appendix D.*

Although the corresponding behavior with discrete constellations is difficult to assess in full generality, for Rayleigh fading it can be compactly characterized. To that end, let us define $\tilde{\gamma}_j = E[|h_{i,j}|^2]$, $i = 1, \ldots, n_j$, which, from (6) and (9), is tantamount to $\tilde{\gamma}_j = \bar{\gamma}_j / P$.

Proposition 5 *If the k users experience Rayleigh fading and their signals conform to the same discrete constellation, then for $P \to \infty$*

$$\bar{p}_j^\star = \alpha \sqrt{\frac{w_j}{\beta_j \tilde{\gamma}_j}} + \mathcal{O}(\frac{1}{P}) \tag{51}$$

with

$$\frac{1}{\alpha} = \sum_{j=1}^{K} \sqrt{\frac{w_j \beta_j}{\tilde{\gamma}_j}}. \tag{52}$$

Proof: *See Appendix D.*

Example 4 *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. 8 are $\bar{p}_1^\star$ and $\bar{p}_2^\star$, obtained from Theorem 2, as function of $\bar{\gamma}_1$. For $P \to \infty$, the power allocation converges rapidly to its limit which, from Proposition 5, equals $\bar{p}_1^\star = 1.48$ and $\bar{p}_2^\star = 0.52$.*

As one would expect from the results in Section B, and Fig. 8 confirms, the limiting values do not depend on the constellation cardinalities but the speed of convergence does. (Naturally, with constellations of smaller cardinality the high-power behaviors are evidenced at lower power levels.) Notwithstanding such difference, for both QPSK and 16-QAM the limit is approached at very modest power levels. Note that, if the users where employing ideal Gaussian signals instead of discrete constellations, the limiting power allocation would be $\bar{p}_1^\star = \bar{p}_2^\star = 1$.

VI Summary

We have considered downlink multiuser OFDM systems where each tone is assigned to one user. For any given tone-user assignment, we have obtained the optimum power allocation policies with both instantaneous and statistical CSIT. These policies hold for arbitrary modulation formats, possibly different for every user. Although the modulation formats of the tones assigned to each user have been assumed to be equal, this restriction can be easily removed. The interesting associated combinatorial optimization problem of assignment of users and constellations to tones is left for future work.

The formulation also readily accommodates multiple transmit and receive antennas as long as the signal transmitted on every tone remains scalar. Techniques such as diversity and beamforming can indeed be straightforwardly included by simply modifying as appropriate the distribution of the channel states $\{\gamma_{i,j}\}$.

With instantaneous CSIT, the more general problem of jointly assigning tones and allocating power could be explored [20, 21]. While with Gaussian signalling and equal priorities on all users this may reduce to simply assigning each tone to the strongest user followed

by a multiuser waterfilling, in general this appears to be a challenging optimization. With statistical CSIT, on the other hand, the problem is comparatively simpler as only the bandwidth fractions $\{\beta_j\}_{j=1}^K$ need to be assigned.

Although the formulation has been set up within a multitone framework, the main results remain valid for single-carrier frequency-division multiplexing (FDM). In this more rigid scenario, each signal becomes equivalent to a number of adjacent tones and the bandwidth fractions $\{\beta_j\}_{j=1}^K$ can be taken to signify the fractional widths of the respective signals rather than the portion of assigned tones. With ideal Gaussian signalling, analyses of the FDM broadcast channel can be found, e.g., in [22]. Our results enable extending the analysis to arbitrary constellation formats.

Appendices

A   Proof of Theorem 1

Since the functions $\mathcal{I}_j(\cdot)$ increase monotonically with the input power, the region of mutual informations is maximized when the power constraint (8) is met with strict equality. At the same time, the objective function in (23) is strictly concave over the set of inputs associated with the feasible power allocations $\{p_{i,j}\}$ because:

- It represents a weighted average of $\mathcal{I}_j(\cdot)$ functions.

- Each $\mathcal{I}_j(\cdot)$ is strictly concave[4] since the corresponding $\mathsf{MMSE}_j(\cdot)$ decreases monotonically and the latter is the derivative of the former by virtue of (15).

Consequently, there is a unique $\{p_{i,j}^\star\}$ that maximizes the function in (23) while satisfying (8) with strict equality. Such optimal power allocation can be determined by building the Lagrangian

$$L(\{p_{i,j}\}) = \frac{1}{n} \sum_{j=1}^K \frac{w_j}{\beta_j} \sum_{i=1}^{n_j} \mathcal{I}_j(p_{i,j}\gamma_{i,j}) - \eta \left(\frac{1}{n} \sum_{i=1}^{n_j} \sum_{j=1}^K p_{i,j} - 1 \right) \tag{53}$$

---

[4]Recall that $\gamma_{i,j} > 0$ for every i and j.

whose maximization for $p_{i,j} \geq 0$, invoking (15), yields the necessary and sufficient conditions

$$\frac{w_j}{\beta_j}\gamma_{i,j}\mathsf{MMSE}_j(p^\star_{i,j}\gamma_{i,j}) = \eta \qquad\qquad p^\star_{i,j} > 0 \tag{54}$$

$$\frac{w_j}{\beta_j}\gamma_{i,j}\mathsf{MMSE}_j(p^\star_{i,j}\gamma_{i,j}) < \eta \qquad\qquad p^\star_{i,j} = 0 \tag{55}$$

Using the definition of $\mathsf{MMSE}_j^{-1}(\cdot)$, (54) and (55) can be rearranged into (24) and (25) as claimed.

B Proof of Proposition 2

Eq. (31) follows straightforwardly from Theorem 1 and from the high-power expansion for Gaussian signals

$$\mathsf{MMSE}_j(\rho) = \frac{1}{\rho} + \mathcal{O}(1/\rho^2). \tag{56}$$

For discrete constellations, in turn, we have [2, Theorem 4]

$$\mathsf{MMSE}_j(\rho) = K(\rho)\, e^{-d_j^2\rho/4} \tag{57}$$

with

$$K_1(\rho) \leq K(\rho) \leq K_2 \tag{58}$$

where $K_1(\rho) = \mathcal{O}(1/\sqrt{\rho})$ and K_2 is a constant. Because of the exponential decay of $\mathsf{MMSE}_j(\cdot)$, we find that $\eta \to 0$ in Theorem 1. Hence, for sufficiently large P, on every tone of every user $p^\star_{i,j} > 0$ strictly. Then, Theorem 1 reduces to (54) and, since the $\log_e(\cdot)$ function is strictly monotonic, to

$$\log_e \frac{w_j}{\beta_j} + \log_e |h_{i,j}|^2 + \log_e P + \log_e \mathsf{MMSE}_j(p^\star_{i,j}|h_{i,j}|^2 P) = \log_e \eta. \tag{59}$$

where we have used $\gamma_{i,j} = P|h_{i,j}|^2$. Dividing both sides of (59) by P and capitalizing on (57)–(58) we find, for large P,

$$\frac{d_j^2}{4}p^\star_{i,j}|h_{i,j}|^2 = \frac{\log_e 1/\eta}{P} + \mathcal{O}\left(\frac{\log P}{P}\right) \tag{60}$$

where, as justified below, $\log_e(1/\eta)/P = \mathcal{O}(1)$. For $P \to \infty$, therefore, $\{p^\star_{i,j}\}$ converges to the solution of

$$d_j^2 p^\star_{i,j}|h_{i,j}|^2 = \alpha \tag{61}$$

18

where α is determined by the power constraint in (8) with strict equality. This leads directly to the claimed limiting power allocation.

To determine that $\log_e(1/\eta)/P = \mathcal{O}(1)$, we first turn (57) into

$$\log_e \mathsf{MMSE}(\rho) = \log_e K(\rho) - \frac{d_j^2}{4}\rho. \tag{62}$$

Plugging (62) into (59) and dividing both sides by $-P$ we obtain, for $P \to \infty$,

$$\frac{d_j^2}{4}p_{i,j}^\star|h_{i,j}|^2 + \mathcal{O}(\frac{\log_e P}{P}) = -\frac{\log_e \eta}{P} \tag{63}$$

which evidences the claimed order for $\log_e(1/\eta)/P$.

C Proof of Theorem 2

The arguments supporting the uniqueness of $\{\bar{p}_j^\star\}_{j=1}^K$ parallel those put forth in Appendix A with regards to $\{p_{i,j}^\star\}$ in Theorem 1 and are not repeated here for the sake of conciseness. Furthermore, and because of the monotonicity of the functions $\mathcal{I}_j(\cdot)$ with the input powers, $\{\bar{p}_j^\star\}_{j=1}^K$ is found when the power constraint is met with strict equality. Hence (44). The optimum $\{\bar{p}_j^\star\}_{j=1}^K$ can be identified by means of the Lagrangian

$$L(\bar{p}_1, \ldots, \bar{p}_K) = \sum_{j=1}^K w_j E\left[\mathcal{I}_j(\bar{p}_j\gamma_{i,j})\right] - \eta\left(\sum_j \beta_j\bar{p}_j - 1\right) \tag{64}$$

whose maximization for $\bar{p}_j \geq 0$, using (15), yields the necessary and sufficient conditions

$$\frac{w_j}{\beta_j}E\left[\gamma_{i,j}\mathsf{MMSE}_j(\bar{p}_j^\star\gamma_{i,j})\right] = \eta \qquad \bar{p}_j^\star > 0 \tag{65}$$

$$\frac{w_j}{\beta_j}E\left[\gamma_{i,j}\mathsf{MMSE}_j(\bar{p}_j^\star\gamma_{i,j})\right] < \eta \qquad \bar{p}_j^\star = 0 \tag{66}$$

These equations are equivalent to (42) and (43) as claimed.

D Proof of Propositions 4 and 5

Proposition 4 is obtained by plugging (56) into Theorem 2.

For Proposition 5, the starting point is the high-power expansion of $\mathsf{MMSE}_j(\cdot)$ given in (57)–(58) with $d_j = d$ for every $j = 1, \ldots, K$. Because of the exponential decay of $\mathsf{MMSE}_j(\cdot)$ for $P \to \infty$, we have $\eta \to 0$ in Theorem 2 and thus the limiting power allocation is strictly positive. It follows that, under Rayleigh fading, Theorem 2 reduces to

$$\frac{P}{\tilde{\gamma}_j} \int_0^\infty \mathsf{MMSE}_j(\bar{p}_j^\star \xi P) \, e^{-\xi/\tilde{\gamma}_j} \xi \, d\xi = \frac{\beta_j}{w_j} \eta \tag{67}$$

where we have used $\gamma_{i,j} = P|h_{i,j}|^2$ and $\bar{\gamma}_j = P\tilde{\gamma}_j$ and where η must ensure that (44) is satisfied. Plugging the upper expansion

$$\mathsf{MMSE}_j(\rho) = \frac{K_1}{\sqrt{\rho}} e^{-\frac{d^2}{4}\rho} \tag{68}$$

into (67) and integrating we obtain

$$\frac{4\sqrt{\pi}K_1}{\tilde{\gamma}_j d^3 (\bar{p}_j^\star)^2} \frac{1}{P} + \mathcal{O}(\frac{1}{P^2}) = \frac{\beta_j}{w_j} \eta \tag{69}$$

from which, defining

$$\alpha = \sqrt{\frac{4\sqrt{\pi}K_1}{d^3 \eta P}}, \tag{70}$$

we obtain

$$\bar{p}_j^\star = \alpha \sqrt{\frac{w_j}{\beta_j \tilde{\gamma}_j}} + \mathcal{O}(\frac{1}{P}) \tag{71}$$

with α such that (44) is satisfied. Note that α is $\mathcal{O}(1)$ with respect to P because, as can be verified using argumentation similar to that in Appendix B, $\eta = \mathcal{O}(1/P)$.

In turn, plugging the lower expansion

$$\mathsf{MMSE}_j(\rho) = K_2 \, e^{-\frac{d^2}{4}\rho} \tag{72}$$

into (67) and integrating we obtain

$$\frac{16K_2}{\tilde{\gamma}_j d^4 (\bar{p}_j^\star)^2} \frac{1}{P} + \mathcal{O}(\frac{1}{P^2}) = \frac{\beta_j}{w_j} \eta \tag{73}$$

which leads to the same parametric solution in (71). Thus, the parametric solutions corresponding to the lower and upper expansions coincide. Given the structure of (67) and the monotonicity of the function $\mathsf{MMSE}_j(\cdot)$, this parametric solution stands and, using (44), the claimed limiting power allocation follows.

References

[1] A. Lozano, A. M. Tulino, and S. Verdú, "Mercury/waterfilling: Optimum power allocation with arbitrary input constellations," *Proc. of IEEE Intern. Symp. on Inform. Theory (ISIT'05)*, Sept. 2005.

[2] A. Lozano, A. M. Tulino, and S. Verdú, "Power allocation over parallel Gaussian channels with arbitrary input distributions," *IEEE Trans. on Inform. Theory*, vol. 52, no. 7, pp. 3033–3051, July 2006.

[3] E. Lawrey, "Multiuser OFDM," *5th Int'l Symp. on Signal Processing and Applications (ISSPA'99)*, pp. 761–764, Aug. 1999.

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

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

[7] 3rd Generation Partnership Project (3GPP), *UTRA-UTRAN Long Term Evolution (LTE)*, Nov. 2004.

[8] 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.

[9] 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.

[10] 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.

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

[12] S. Thoen, L. Van der Perre, M. Engels, and H. De Man, "Adaptive loading for OFDM/SDMA-based wireless networks," *IEEE Trans. on Communications*, vol. 50, no. 11, pp. 1798–1810, Nov. 2002.

[13] 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.

[14] A. M. Tulino, A. Lozano, and S. Verdú, "Capacity-achieving input covariance for single-user multi-antenna channels," *IEEE Trans. on Wireless Communications*, vol. 5, no. 2, pp. 662–671, Mar. 2006.

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

[16] M.-S. Alouini and A. J. Goldsmith, "Capacity of Rayleigh fading channels under different adaptive transmission and diversity-combining techniques," *IEEE Trans. Veh. Technol.*, vol. 48, no. 4, pp. 1165–1181, July 1999.

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

[18] W. C. Y. Lee, "Estimate of channel capacity in Rayleigh fading environments," *IEEE Trans. Veh. Technol.*, vol. 39, pp. 187–189, Aug. 1990.

[19] L. H. Ozarow, S. Shamai, and A. D. Wyner, "Information theoretic considerations for cellular mobile radio," *IEEE Trans. Veh. Technol.*, vol. 43, no. 2, pp. 359–378, May 1994.

[20] D. Kivanc, G. Li, and H. Liu, "Computationally efficient bandwidth allocation and power control for OFDMA," *IEEE Trans. on Wireless Communications*, vol. 2, pp. 1150–1158, Nov. 2003.

[21] Z. Shen, J. G. Andrews, and B. L. Evans, "Adaptive resource allocation in multiuser OFDM systems with proportional rate constraints," *IEEE Trans. on Wireless Communications*, vol. 4, no. 6, pp. 2726–2737, Nov. 2005.

[22] L. Li and A. J. Goldsmith, "Capacity and optimal resource allocation for fading broadcast channels - Part I: Ergodic capacity," *IEEE Trans. on Inform. Theory*, vol. 47, pp. 1083–1102, Mar. 2001.
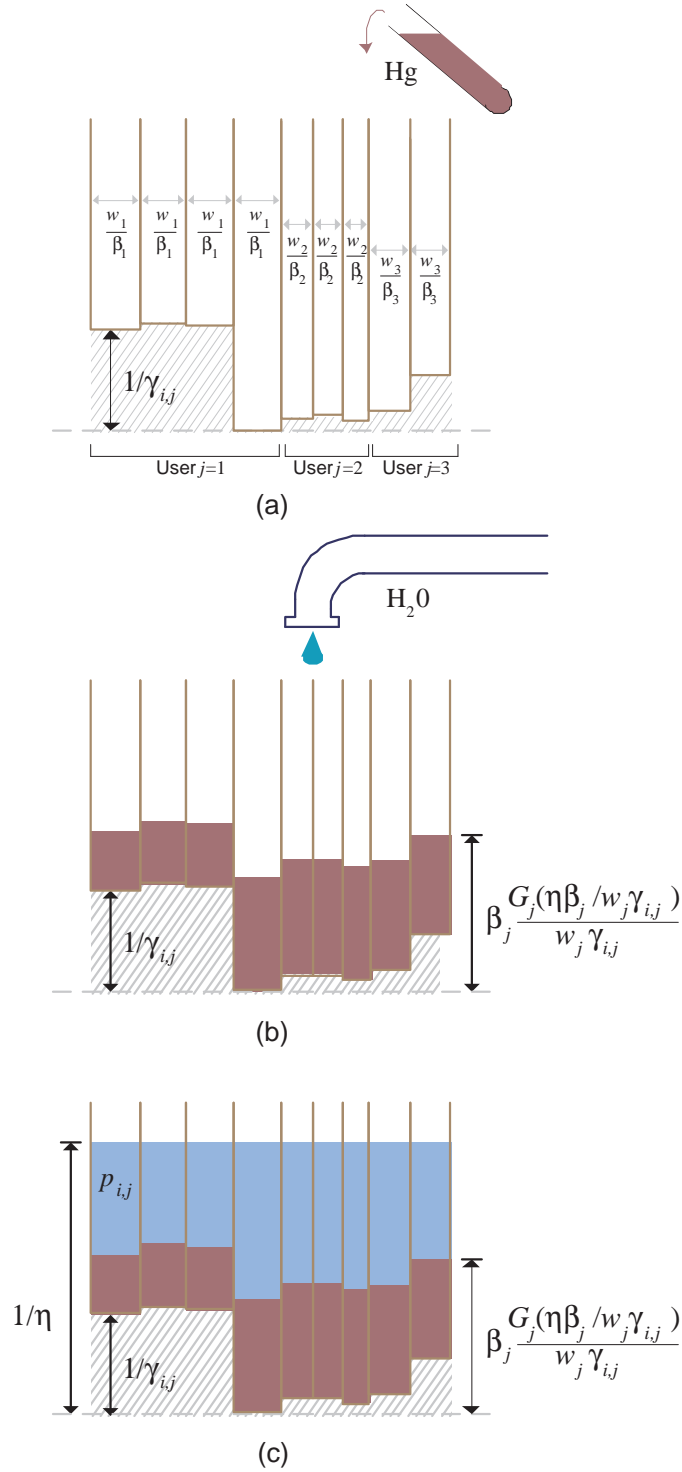
Hg

$\frac{w_1}{\beta_1}$ $\frac{w_1}{\beta_1}$ $\frac{w_1}{\beta_1}$ $\frac{w_1}{\beta_1}$ $\frac{w_2}{\beta_2}$ $\frac{w_2}{\beta_2}$ $\frac{w_2}{\beta_2}$ $\frac{w_3}{\beta_3}$ $\frac{w_3}{\beta_3}$

$1/\gamma_{i,j}$

User $j=1$ User $j=2$ User $j=3$

(a)

H_2O

$1/\gamma_{i,j}$

$\beta_j \dfrac{G_j(\eta\beta_j/w_j\gamma_{i,j})}{w_j\gamma_{i,j}}$

(b)

$p_{i,j}$

$1/\eta$

$1/\gamma_{i,j}$

$\beta_j \dfrac{G_j(\eta\beta_j/w_j\gamma_{i,j})}{w_j\gamma_{i,j}}$

(c)

Figure 1: Multiuser mercury/waterfilling with $K = 3$ users and with adjacent tones assigned to each user for clarity. (a) On vessels solid up to heights $1/\gamma_{i,j}$, pour mercury as shown. (b) Waterfill until the water levels reach $1/\eta$. (c) The volume of water in the (i, j)th vessel gives $p_{i,j}$.

23
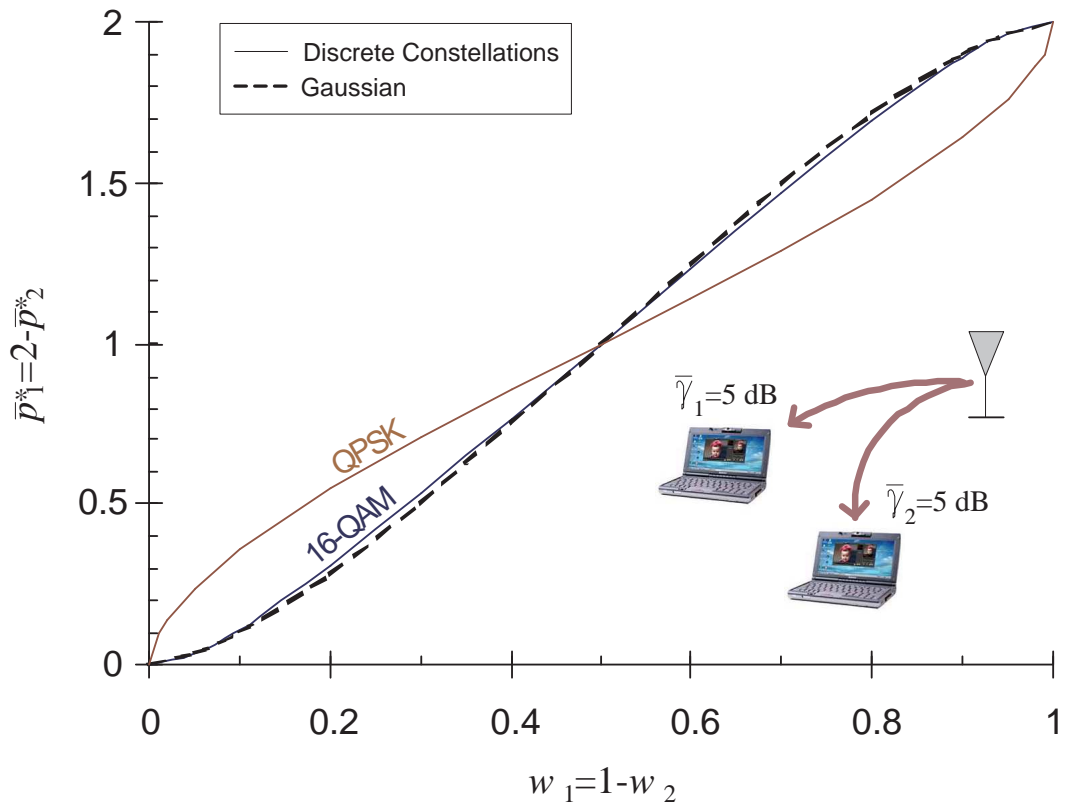
Figure 2: Average power allocation as function of w_1, for $K = 2$ users having $\bar{\gamma}_1|_{\mathrm{dB}} = \bar{\gamma}_2|_{\mathrm{dB}} = 5$. Both channels are frequency-flat Rayleigh-faded with bandwidth partitioning $\beta_1 = \beta_2 = 1/2$.
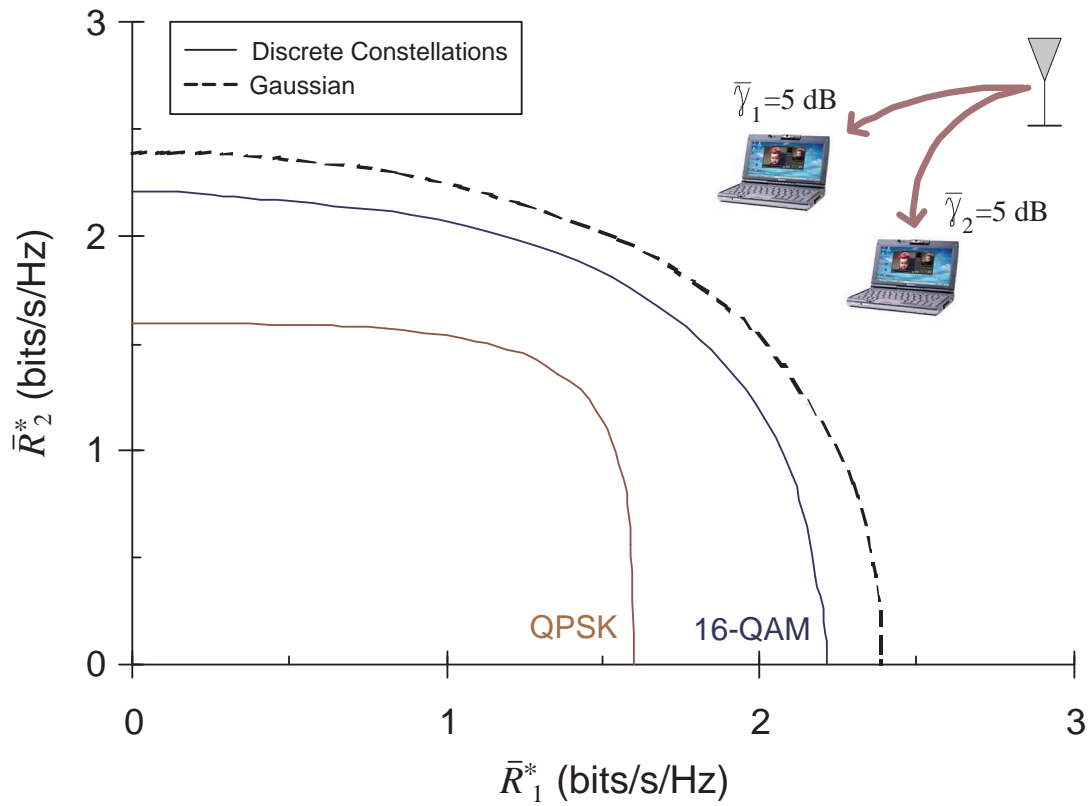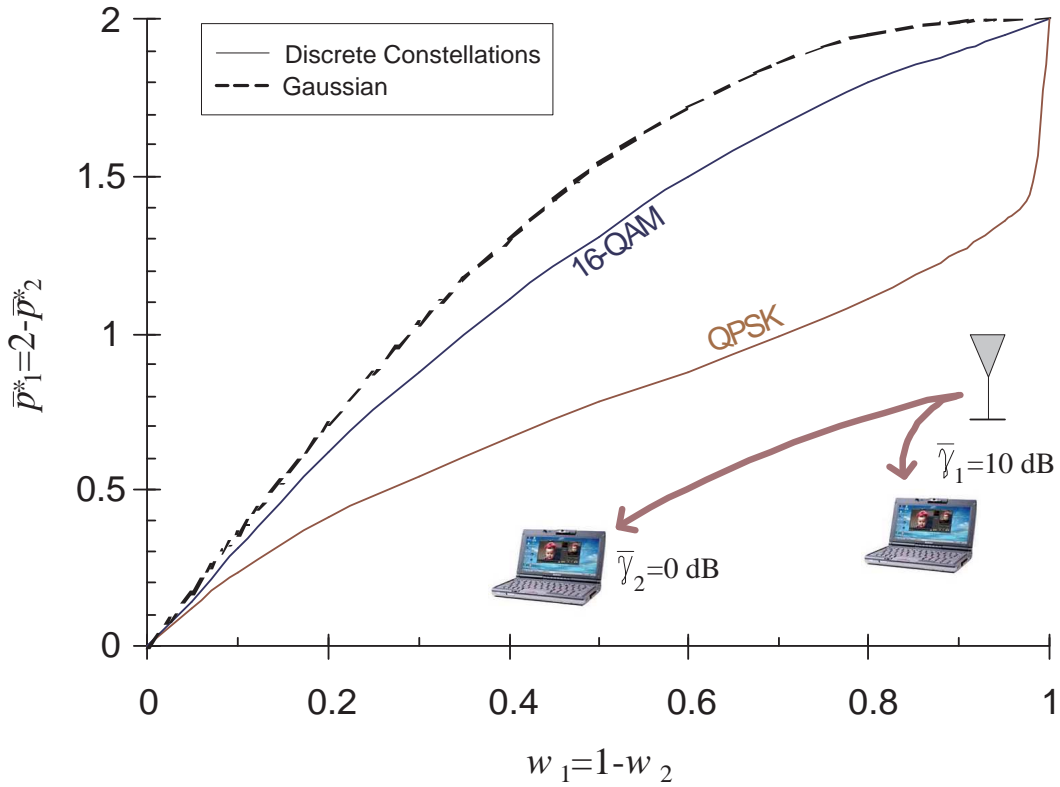
Figure 3: Average mutual information regions for $K = 2$ users having $\bar{\gamma}_1|_{\text{dB}} = \bar{\gamma}_2|_{\text{dB}} = 5$. Both channels are frequency-flat Rayleigh-faded with bandwidth partitioning $\beta_1 = \beta_2 = 1/2$.

Figure 4: Ergodic power allocation with instantaneous CSIT, as function of the priority w_1, for $K = 2$ users having $\bar{\gamma}_1|_{dB} = 10$ and $\bar{\gamma}_2|_{dB} = 0$. 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.
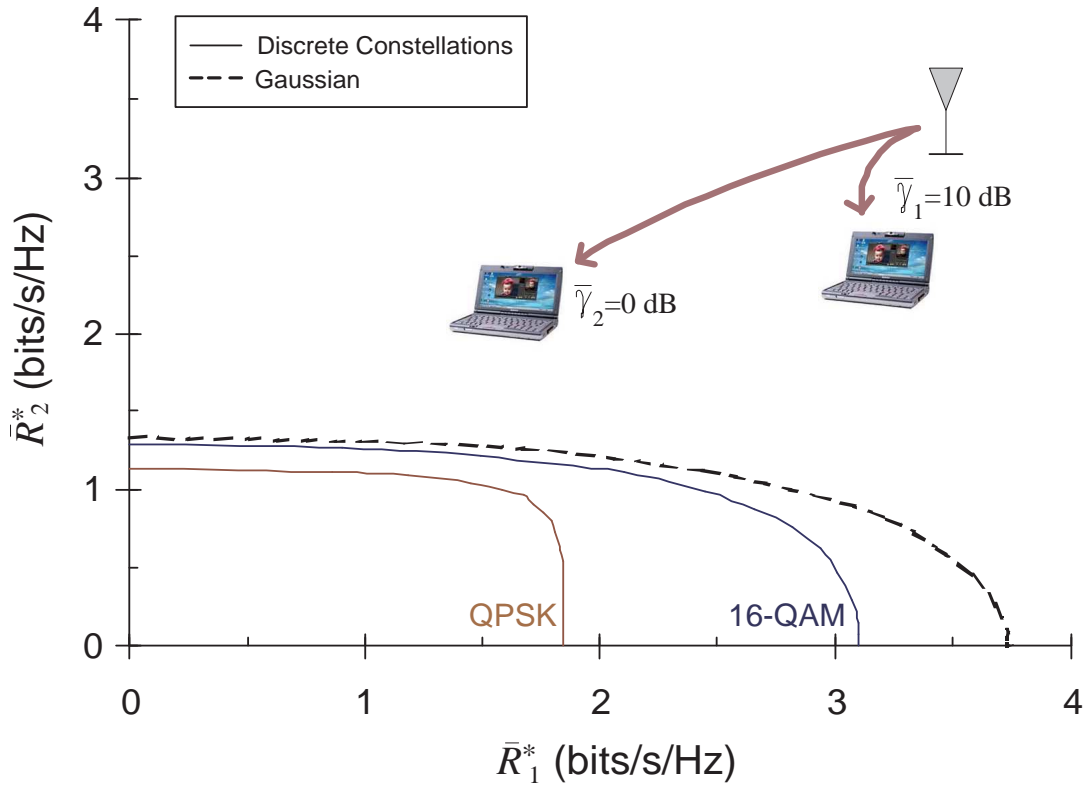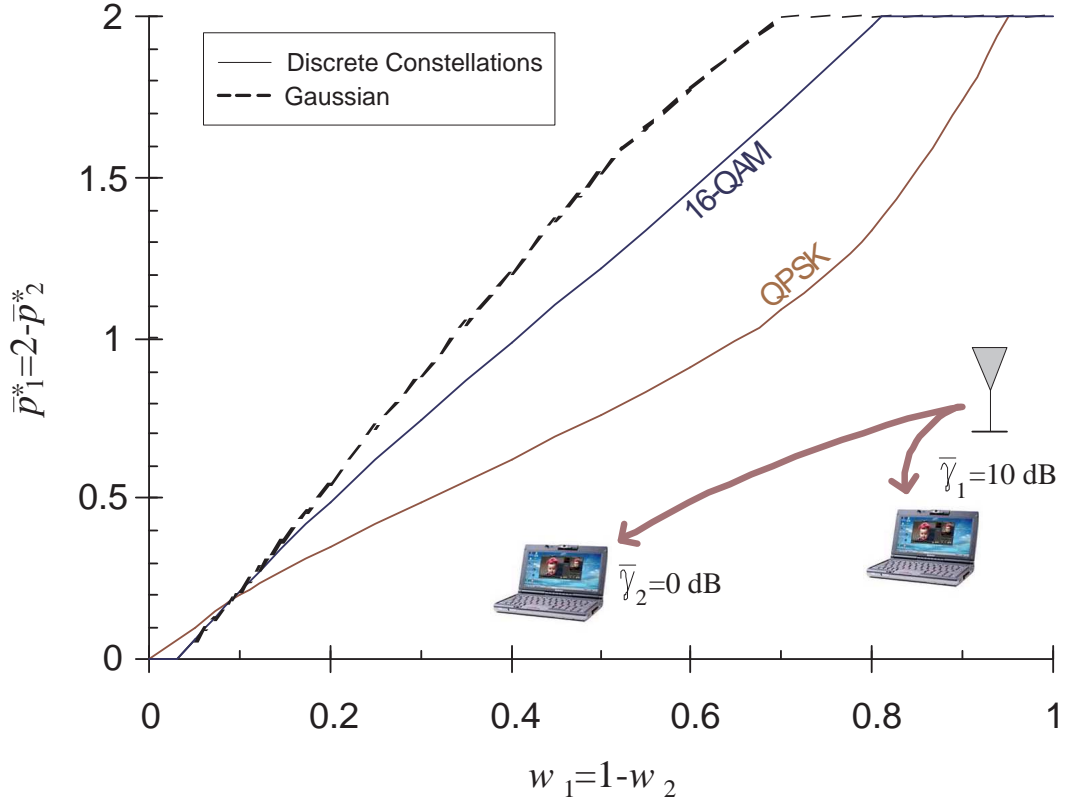
Figure 5: Ergodic mutual information regions with instantaneous CSIT for $K = 2$ users having $\bar{\gamma}_1|_{dB} = 10$ and $\bar{\gamma}_2|_{dB} = 0$. Both channels are frequency-flat Rayleigh-faded with bandwidth partitioning $\beta_1 = \beta_2 = 1/2$. Continuous lines correspond to discrete constellations (QPSK and 16-QAM) while the dashed line corresponds to a Gaussian signal.

Figure 6: Ergodic optimum power allocation with statistical CSIT, as function of the priority w_1, for $K = 2$ users having $\bar{\gamma}_1|_{\text{dB}} = 10$ and $\bar{\gamma}_2|_{\text{dB}} = 0$. 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.
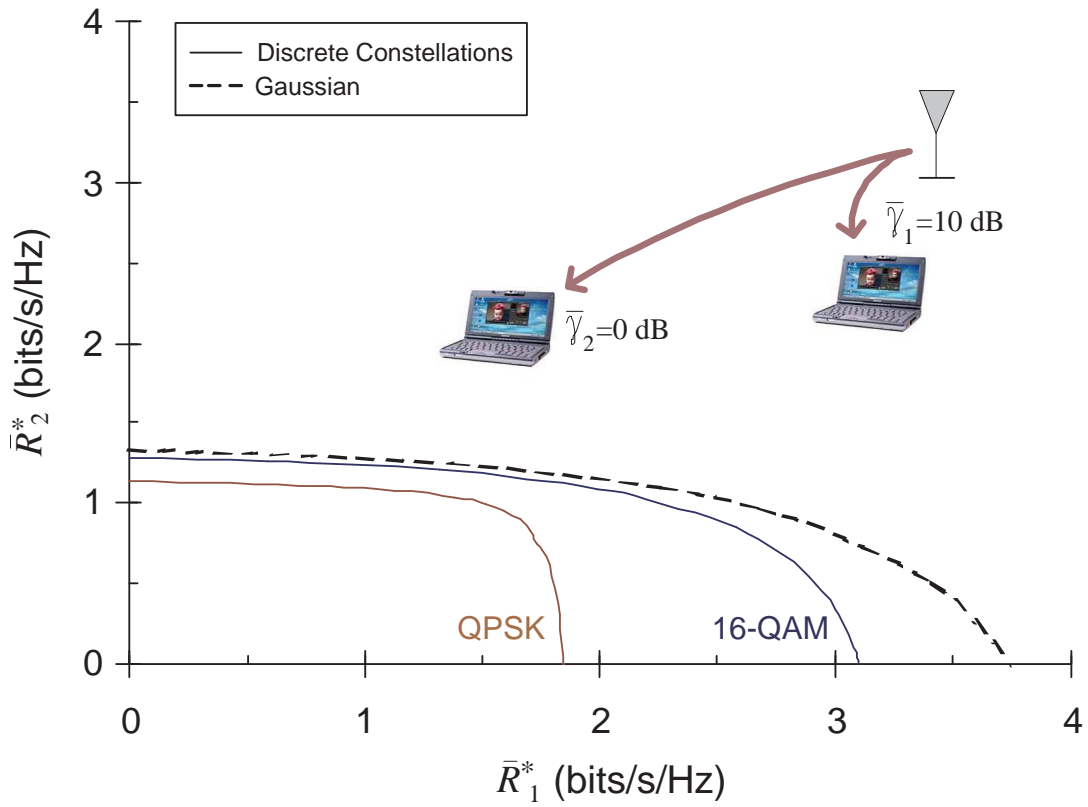
Figure 7: Ergodic mutual information region with statistical CSIT for $K = 2$ users having $\bar{\gamma}_1|_{\mathrm{dB}} = 10$ and $\bar{\gamma}_2|_{\mathrm{dB}} = 0$. Both channels are frequency-flat Rayleigh-faded. Continuous lines correspond to discrete constellations (QPSK and 16-QAM) while the dashed line corresponds to a Gaussian signal.
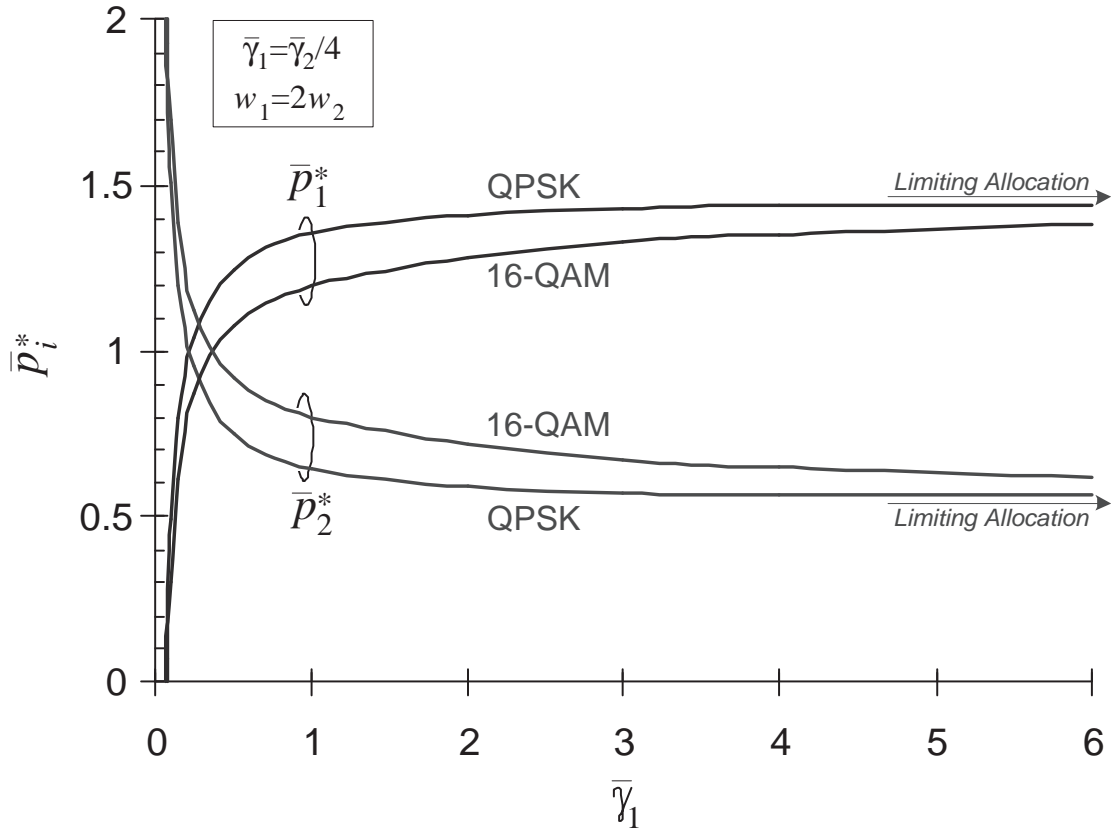
Figure 8: For $K = 2$ with $\bar{\gamma}_1|_{dB} = \bar{\gamma}_2|_{dB} - 6$ and $w_1 = 2w_2$, exact $\bar{p}_1^\star$ and $\bar{p}_2^\star$ as function of $\bar{\gamma}_1$ and parameterized by the constellation used by both users (QPSK or 16-QAM). The limiting power allocation for $P \to \infty$ is also indicated.