On the Information Throughput and Optimized Power Allocation for MIMO Wireless Systems With Imperfect Channel Estimation

Enzo Baccarelli, Mauro Biagi, *Member, IEEE*, and Cristian Pelizzoni

***Abstract*—In this paper, we focus on the throughput analysis, outage evaluation and optimized power allocation for Multiple-Input Multiple-Output (MIMO) *pilot-based* wireless systems subject to short-term constraints on the radiated power and equipped with a feedback-path for communicating back to the transmitter the *imperfect* MIMO channel estimates available at the receiver. The case of the ergodic throughput for Gaussian distributed input signals is analyzed, and the conditions for the (asymptotical) achievement of the Shannon capacity are pointed out. The main contributions of this work may be so summarized. First, we develop closed-form analytical expressions for the computation of the ergodic information throughput conveyed by the considered MIMO system for the case of ideal feedback link. Second, we present an iterative algorithm for the optimized power allocation over the transmit antennas that explicitly accounts for the *imperfect* MIMO channel estimates available at the receiver. Third, after relaxing the assumption of ideal feedback link, we test the sensitivity of the proposed power allocation algorithm on errors possibly introduced by the feedback channel, and then, we numerically evaluate the resulting throughput loss. Finally, we develop closed-form upper and lower bounds on the outage probability that are asymptotically tight.**

***Index Terms*—Ergodic throughput, imperfect channel estimation, MIMO systems, multiple antennas, outage probability, power allocation.**

I. Motivations and Outline

BROADBAND wireless access along with evolving mobile Internet and multimedia applications are driving the development of fourth-Generation Wireless Local Area Networks (4GWLANs) and the Wireless Metropolitan Area Network (WMAN). Since the main goal of emerging 4GWLANs is to provide reliable services to nomadic users at throughput exceeding 100 Mb/s, new wireless architectures are being planned to attain large spectral and power efficiencies [8]. Both these efficiencies may be substantially improved by resorting to MIMO systems that are able to simultaneously exploit the temporal and the spatial diversity of fading channels for increasing the conveyed throughput without wasting the available power and bandwidth [18].

Manuscript received March 18, 2004; revised July 30, 2004. This work was supported by the New-Internet Project funded by the autonomous province of Trento. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Markus Rupp.

The authors are with INFO-COM Department, University of Rome "La Sapienza," 00184 Rome, Italy (e-mail: enzobac@infocom.uniroma1.it; m.biagi@ieee.org; pelcris@infocom.uniroma1.it).

Digital Object Identifier 10.1109/TSP.2005.849165

Thus far, a first line of work focused on the evaluation of the ultimate information throughput sustained by multiantenna systems when perfect channel estimates are available at the receiver (see [1], [8], and references therein) and, possibly, even at the transmitter (see [7], [24], and references therein). In particular, in [19], the throughput sustained by quadrature amplitude modulated (QAM) MIMO systems is analytically evaluated and compared against the corresponding channel capacity achieved by Gaussian distributed input signals. However, in mobile environments, channel coefficients may change fast so that their reliable estimation becomes difficult to achieve, especially for systems with a large number of antenna elements. Therefore, in such operating scenarios, it may be of interest to explore the ultimate information throughput supported by MIMO wireless architectures when no channel estimates are available both at the receiver and transmitter. A line of work along this direction was initiated in [3] and refined, for example, in [4], [11], and [30]. These contributions lead to the development of unitary (possibly, differential) space-time codes for fully incoherent applications that, unfortunately, suffer from non-negligible performance loss with respect to MIMO systems equipped with perfect channel estimates [18] and references therein. Hence, to achieve the improved bandwidth-versus-power tradeoff demanded by emerging 4GWLANs, it may be of interest to also explore the information throughput conveyed by MIMO systems equipped with partial (e.g., imperfect) channel information at the transmitter (CSIT) and/or at the receiver (CSIR). About this topic, the recent contribution [28] introduces a general MIMO system with partial CSIR and CSIT, and then, it studies its capacity in two different specific cases, namely, perfect CSIR and partial CSIR given by the "instantaneous SNR." However, the topic of MIMO channel estimation is not addressed in [28], and the capacity loss induced by errors possibly affecting the MIMO channel estimates is not investigated either.

Pilot-based MIMO systems relaying on short training sequences for acquiring rough (e.g., imperfect) channel estimates are examined in [9]–[11], where a lower bound on the resulting system capacity is developed, and then, it is exploited for optimization purposes [10].

A. Main Contributions and Organization of the Paper

In this work, we present new results about the ergodic information throughput, outage probability, and optimized power allocation for pilot-based MIMO wireless systems aided by *imperfect* channel estimates at the receiver when "short-term"

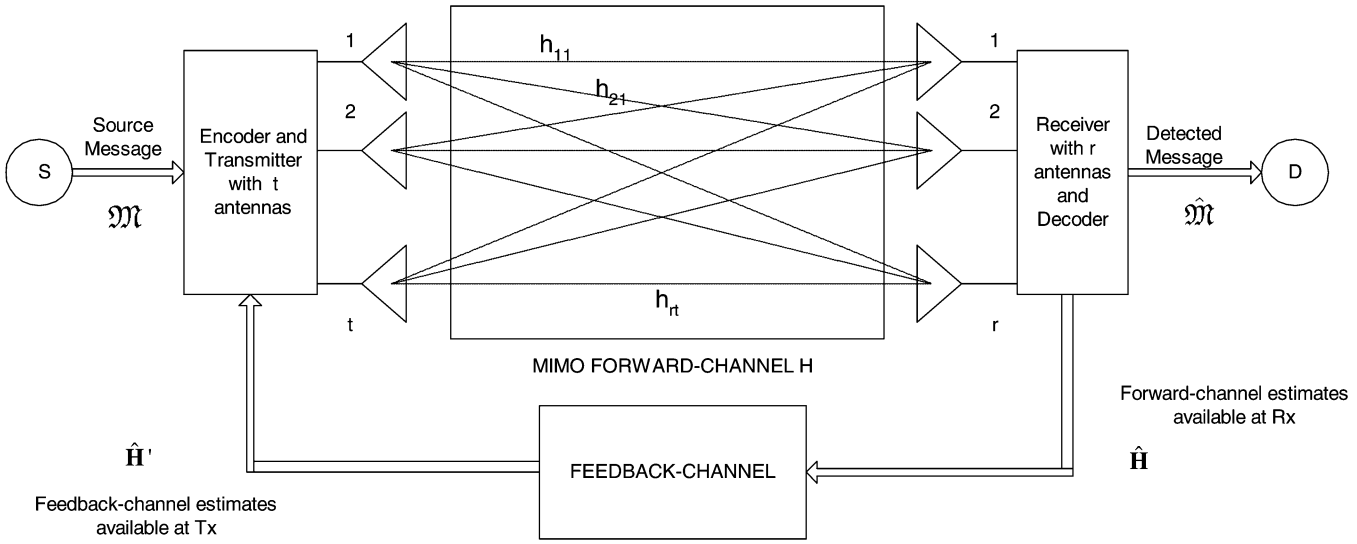

Fig. 1. Sketch of the considered multiantenna system with feedback channel. $\hat{\mathbf{H}} = \hat{\mathbf{H}}'$ for an ideal feedback link.

constraints are imposed on the overall radiated power and the input signals are Gaussian distributed. Thus, the application scenario we consider *generalizes* those previously covered in [5], [6], [9], and [10] and, de facto, comprehends them as limit cases. Furthermore, the systems on which we focus are the emerging 4GWLANs adopting packet-based transmissions to provide broadband services to nomadic users. Thus, in this case, it is reasonable to resort to the (usual) block-fading frequency-flat model [3], [4], [11] for describing the behavior of the underlying MIMO channel, where the fading blocks may be considered as separated in time (e.g., time-division system), in frequency (e.g., multicarrier systems), or both in time and frequency (e.g., time-frequency hopping systems).

The remainder of the paper is organized as follows. After introducing in Section II the considered system model, in Section III, we carry out the corresponding throughput analysis for Gaussian distributed input signals, and then, we provide the resulting optimized power allocation algorithm for the (ideal) case of noiseless feedback link. Afterwards, in the first part of Section IV, we present numerical plots testing the actual effectiveness of the proposed power allocation algorithm, whereas in Section IV-A, we test its sensitivity on errors possibly affecting the feedback link. Thus, Section IV-B is devoted to the presentation of upper and lower bounds on the information outage probability of the system, whereas in Section IV-C, we summarize some practical guidelines for an optimized system design. Analytical details and proofs are included in the final Appendices.

Before proceeding, some remarks about the adopted notations are in order. Bold capital letters denote matrices and underlined bold lower case symbols indicate vectors, while bold characters overlined by an arrow $\vec{(.)}$ denote block-matrices and block-vectors. Apexes $(.)^*$, $(.)^T$, $(.)^\dagger$ mean conjugation, transposition, and conjugate-transposition respectively, and lowercase letters are used for scalar quantities. Furthermore, $\det[\mathbf{A}]$ and $\mathrm{TR}[\mathbf{A}]$ denote determinant and trace of the matrix $\mathbf{A} \triangleq [\underline{\mathbf{a}}_1 \dots \underline{\mathbf{a}}_m]$, whereas $vect(\mathbf{A})$ indicates the (block vector built up by the ordered stacking of the columns of $\mathbf{A}$. Finally, $\mathbf{I}_m$ is the $(m \times m)$ unit matrix, $||\mathbf{A}||_E$ indicates the Euclidean norm of $\mathbf{A}$, $\mathbf{A} \otimes \mathbf{B}$ is the Kroenecker product of matrix $\mathbf{A}$ by matrix $\mathbf{B}$, $\underline{\mathbf{0}}_m$ is the m-dimensional zero-vector, and $\lg$ denotes natural logarithm.

II. System Modeling

Let us consider the (complex baseband equivalent) point-to-point link composed by $t \geq 1$ transmit and $r \geq 1$ receive antennas impaired by slow-variant Rayleigh flat-fadings sketched in Fig. 1. The path gain h_{ji} from the transmit antenna i to the receive one j may be modeled as a complex zero-mean unit-variance proper random variable (r.v.) [1], [3], [4], and for sufficiently spaced antennas, these gains $\{h_{ji} \in \mathbb{C}, 1 \leq j \leq r, 1 \leq i \leq t\}$ may be considered mutually independent [1]. Furthermore, for low-mobility applications like those served by emerging 4GWLANs, the path gains $\{h_{ji}\}$ may be assumed constant over $T \geq 1$ signaling periods, after which, they change to new *statistically independent* values held for another T signaling periods, and so on.[1] The resulting "block-fading" model generally gives rise to adequate representations of several TDMA, frequency-hopping, and packet-based interleaved systems of practical interest, where each transmitted frame is independently detected [3], [4], [19]. More in particular, we assume that the coded and modulated streams radiated by the transmit antennas of Fig. 1 are split into packets constituted by $T \equiv T_{\mathrm{tr}} + T_{\mathrm{pay}}$ slots, where the first $T_{\mathrm{tr}} \geq 0$ slots (e.g., the packet header) are used for the transmission of known pilot sequences, whereas the last $T_{\mathrm{pay}} = T - T_{\mathrm{tr}}$ slots convey payload data.

A. Training Phase and Minimum Mean Square Error (MMSE) Estimation of the Path Gains

Therefore, the $T_{\mathrm{tr}} \times r$ (complex) samples gathered at the outputs of the r receive antennas during the training phase may be

[1]This is equivalent to assuming that the coherence-time T_{coh} of the MIMO channel of Fig. 1 equates to the length T of the transmitted packets (e.g., $T_{\mathrm{coh}} = T$). This assumption will be relaxed in Section V, where the outage probability of no ergodic delay-limited MIMO systems will be considered.

organized into the $\mathbb{C}^{T_{\rm tr}\times r}$ observed matrix $\widetilde{\mathbf{Y}} \triangleq [\underline{\widetilde{\mathbf{y}}}_1 \ \dots \ \underline{\widetilde{\mathbf{y}}}_r]$ given by [3], [4], [10]

$$\widetilde{\mathbf{Y}} = \sqrt{\frac{\widetilde{\gamma} T_{\rm tr}}{t}}\widetilde{\mathbf{X}}\mathbf{H} + \widetilde{\mathbf{W}} \tag{1}$$

where $\widetilde{\mathbf{X}} \triangleq [\underline{\widetilde{\mathbf{x}}}_1 \ \dots \ \underline{\widetilde{\mathbf{x}}}_t]$ is the $\mathbb{C}^{T_{\rm tr}\times t}$ matrix of known transmitted pilots, $\mathbf{H} \triangleq [\underline{\mathbf{h}}_1 \ \dots \ \underline{\mathbf{h}}_r]$ is the $\mathbb{C}^{t\times r}$ complex matrix of the above introduced MIMO channel path gains $\{h_{ji}\}$, and the $\mathbb{C}^{T_{\rm tr}\times r}$ noise matrix $\widetilde{\mathbf{W}}$ is composed by independent zero-mean unit-variance proper complex Gaussian samples. It is also assumed that the pilot matrix $\widetilde{\mathbf{X}}$ in (1) meets the following normalizing relationship:[2]

$$\mathrm{TR}\left[\widetilde{\mathbf{X}}^{\dagger}\widetilde{\mathbf{X}}\right] = t \tag{1.1}$$

so that $\widetilde{\gamma}$ in (1) represents the (average) SNR per received sample at the output of each receive antenna during the training phase. The observations $\widetilde{\mathbf{Y}}$ in (1) are employed by the receiver to build up the MMSE matrix estimate $\widehat{\mathbf{H}} \triangleq E\{\mathbf{H}|\widetilde{\mathbf{Y}}\}$ of the corresponding channel matrix $\mathbf{H}$. This last topic has been debated in depth in [9], [11], [22] so that, for the sake of brevity, we will only summarize some key results exploited in Sections II-B and C.

First, in [21], it is proved that the MMSE estimates $\widehat{\mathbf{H}}$ and related error-covariances constitute a *sufficient statistic* for the maximum likelihood (ML) detection $\widehat{\mathfrak{M}}$ of the transmitted source message $\mathfrak{M}$ of Fig. 1. Hence, no information loss is experienced by the considered receiver composed by an MMSE channel-estimator cascaded to an ML detector. Second, since $\widetilde{\mathbf{W}}$ in (1) is Gaussian and $\widetilde{\mathbf{X}}$ is known to the receiver, the MMSE estimates $\widehat{\mathbf{H}}$ may be computed via a direct application of the Principle of Orthogonality. Therefore, a key problem is, indeed, how to choose the pilot matrix $\widetilde{\mathbf{X}}$ in order to minimize the total mean square error $\mathrm{sqer}(\widetilde{\mathbf{X}}) \triangleq \|\mathbf{H} - \widehat{\mathbf{H}}\|_E^2$ experienced when $\widetilde{\mathbf{X}}$ is employed for the MMSE estimates of $\mathbf{H}$. In this regard, it is proved [10], [21] that the optimal pilot matrix $\widetilde{\mathbf{X}}$ must be unitary so that the resulting MMSE channel estimates $\widehat{\mathbf{H}}$ are given by the simple expression [10], [21]

$$\widehat{\mathbf{H}} = \sigma_{\varepsilon}^2\sqrt{\frac{\widetilde{\gamma} T_{\rm tr}}{t}}\widetilde{\mathbf{X}}^{\dagger}\widetilde{\mathbf{Y}}, \quad (\widetilde{\mathbf{X}} \text{ unitary matrix}) \tag{2}$$

where

$$\sigma_{\varepsilon}^2 \triangleq E\left\{\|\widehat{h}_{ji} - h_{ji}\|^2\right\} \equiv E\left\{\|\varepsilon_{ji}\|^2\right\} = \left(1 + \frac{T_{\rm tr}\widetilde{\gamma}}{t}\right)^{-1} \tag{2.1}$$

is the mean squared error affecting each estimate $\widehat{h}_{ji}$. Furthermore, from the uncorrelation property of $\widehat{\mathbf{H}}$ with the corresponding error matrix $\mathbf{E} \triangleq \mathbf{H} - \widehat{\mathbf{H}}$, it follows that all the $\mathbb{C}^{t\times r}$ elements of matrices $\widehat{\mathbf{H}}$ and $\mathbf{E}$ are mutually independent equally distributed zero-mean complex Gaussian proper r.v.s with variances $E\{|\widehat{h_{ji}}|^2\} \equiv 1 - \sigma_{\varepsilon}^2$ and $E\{\|\varepsilon_{ji}\|^2\} \equiv \sigma_{\varepsilon}^2$, respectively (see [10] and [21] for the formal proofs of these properties). Therefore, since $\lim_{\sigma_{\varepsilon}^2\to 0}\widehat{\mathbf{H}} = \mathbf{H}$ and $\lim_{\sigma_{\varepsilon}^2\to 1}\widehat{\mathbf{H}} = \mathbf{0}$, thereinafter, we refer to the limit cases of $\sigma_{\varepsilon}^2 = 0$ and $\sigma_{\varepsilon}^2 = 1$ as those of Perfect Channel-State Information (PCSI) and No CSI (NCSI), respectively.[3] Similarly, we qualify the case of $0 < \sigma_{\varepsilon}^2 < 1$ as that of Imperfect CSI (ICSI). Interesting results about the optimized setting of the training sequences may be found in [10].

B. Some System Considerations About the Feedback Channel

After computing the channel estimates $\widehat{\mathbf{H}}$, the receiver communicates them back to the transmitter via the feedback channel, so that $\widehat{\mathbf{H}}' \triangleq \left[\underline{\hat{\mathbf{h}}}_1' \dots \underline{\hat{\mathbf{h}}}_r'\right]$ is the $\mathbb{C}^{t\times r}$ matrix composed by the MIMO channel estimates $\{\hat{h}_{ji}', 1 \le j \le r, 1 \le i \le t\}$ that is available at the transmitter of Fig. 1. About the modeling of the feedback link, we point out that Time-Division-Duplex (TDD) systems planned for low-mobility applications may exploit pilot tones on the downlink to estimate the uplink and *vice versa*, so that in this case, the assumption of noiseless and delay-free feedback link may be considered reasonable [7]. However, this assumption falls short when the Doppler spread of the MIMO (forward) channel is of the same order of the transmission rate of the feedback one [5], [6] or when the capacity of the feedback link is limited, so that the estimates $\widehat{\mathbf{H}}'$ fed back to the transmitter are affected by quantization noise [12].[4] Therefore, for sake of clarity, in Section III, we focus on the case of ideal feedback link, and then, we debate and test the validity limit of this assumption in Section IV-A.

C. Payload Phase

On the basis of the channel estimates $\widehat{\mathbf{H}}$ and message $\mathfrak{M}$ to be communicated, the transmitter of Fig. 1 suitably shapes the signal streams $\{\phi_i(n) \in \mathbb{C}, T_{\rm tr}+1 \le n \le T\}$, $1 \le i \le t$, radiated by the transmit antennas. The corresponding (sampled) signals $\{y_j(n) \in \mathbb{C}, T_{\rm tr}+1 \le n \le T\}$, $1 \le j \le r$ received during the payload phase at the outputs of the receive antennas may be modeled as

$$y_j(n) = \sqrt{\frac{\gamma T_{\rm pay}}{t}}\sum_{i=1}^{t} h_{ji}\phi_i(n) + w_j(n)$$
$$T_{\rm tr}+1 \le n \le T, \ \ 1 \le j \le r \tag{3}$$

where $\{w_j(n)\}$ are mutually independent zero-mean unit-variance proper Gaussian noise samples. Therefore, after assuming the following "short-term" constraint [3], [4]:

$$\sum_{i=1}^{t} E\left\{\|\phi_i(n)\|^2\right\} = \frac{t}{T_{\rm pay}}, \quad T_{\rm tr}+1 \le n \le T \tag{3.1}$$

on the average power radiated by the transmit antennas during *each slot-period*, it is recognized that γ in (3) represents the resulting average SNR at the output of each receive antenna, regardless of the number t of the transmit antennas. Hence, from (3), we deduce that the $\mathbb{C}^{r\times 1}$ column vector

[2]This is automatically satisfied when $\widetilde{\mathbf{X}}$ is a unitary matrix so that the defining relationship $\widetilde{\mathbf{X}}^{\dagger}\widetilde{\mathbf{X}} = \mathbf{I}_t$ is met.

[3]For large $\widetilde{\gamma} T_{\rm tr}$, σ_{ε}^2 in (2.1) vanishes, while $\widetilde{\mathbf{Y}}$ in (1) converges to $\sqrt{(\widetilde{\gamma} T_{\rm tr}/t)}\widetilde{\mathbf{X}}\mathbf{H}$ so that $\widehat{\mathbf{H}}$ in (2) approaches $\mathbf{H}$. On the other hand, when σ_{ε}^2 in (2.1) approaches unit, we have that $\widetilde{\gamma} T_{\rm tr}$ vanishes so that the right-hand-side (r.h.s.) of (2) vanishes as well, thus meaning that the pdf of $\widehat{\mathbf{H}}$ converges to the Dirac-delta. This is in agreement with the fact that when σ_{ε}^2 in (2.1) approaches unit, the observation $\widetilde{\mathbf{Y}}$ in (1) converges to the noise $\widetilde{\mathbf{W}}$ so that no information about $\mathbf{H}$ is carried by $\widetilde{\mathbf{Y}}$.

[4]An analysis of the effects of nonideal feedback links are carried out, for example, in [5], [6], and [12] under the assumption of perfect channel estimates available at the receiver.

$\underline{\mathbf{y}}(n) \triangleq [y_1(n)\dots y_r(n)]^T$ collecting the outputs of the r receive antennas during the nth payload slot is linked to the $\mathbb{C}^{t\times 1}$ column vector $\underline{\boldsymbol{\phi}}(n) \triangleq [\phi_1(n)\dots\phi_t(n)]^T$ of the signals radiated by the transmit antennas via the relationship

$$\underline{\mathbf{y}}(n) = \sqrt{\frac{\gamma T_{\rm pay}}{t}}\mathbf{H}^T\underline{\boldsymbol{\phi}}(n) + \underline{\mathbf{w}}(n), \quad T_{\rm tr}+1 \le n \le T \quad (4)$$

where $\underline{\mathbf{w}}(n) \triangleq [w_1(n)\dots w_t(n)]^T$ is the corresponding noise vector. Thus, the resulting $\mathbb{C}^{t\times t}$ correlation matrix $\mathbf{R}_{\boldsymbol{\phi}} \triangleq E\{\underline{\boldsymbol{\phi}}(n)\underline{\boldsymbol{\phi}}(n)^\dagger\}$ of the radiated signals is constrained as (see (3.1))

$$\mathrm{TR}[\mathbf{R}_{\boldsymbol{\phi}}] \equiv E\left\{\underline{\boldsymbol{\phi}}(n)^\dagger\underline{\boldsymbol{\phi}}(n)\right\} = \frac{t}{T_{\rm pay}}, \quad T_{\rm tr}+1 \le n \le T. \quad (4.1)$$

Finally, after stacking the $T_{\rm pay}$ observed vectors $\{\underline{\mathbf{y}}(n)\}$ in (4) into the corresponding $\mathbb{C}^{T_{\rm pay}r\times 1}$ block vector $\vec{\underline{\mathbf{y}}} \triangleq \left[\underline{\mathbf{y}}^T(T_{\rm tr}+1)\dots\underline{\mathbf{y}}^T(T)\right]^T$, we may collect the $T_{\rm pay}$ relationships in (4) as in

$$\vec{\underline{\mathbf{y}}} = \sqrt{\frac{\gamma T_{\rm pay}}{t}}\left[\mathbf{I}_{T_{\rm pay}}\otimes\mathbf{H}\right]^T\vec{\underline{\boldsymbol{\phi}}} + \vec{\underline{\mathbf{w}}} \quad (5)$$

where $\vec{\underline{\mathbf{w}}} \triangleq \left[\underline{\mathbf{w}}^T(T_{\rm tr}+1)\dots\underline{\mathbf{w}}^T(T)\right]^T$. Hence, the $\mathbb{C}^{T_{\rm pay}t\times 1}$ block vector $\vec{\underline{\boldsymbol{\phi}}} \triangleq [\underline{\boldsymbol{\phi}}^T(T_{\rm tr}+1)\dots\underline{\boldsymbol{\phi}}^T(T)]^T$ of the radiated signals in (5) must satisfy the following relationship (see (3.1)):

$$E\left\{\vec{\underline{\boldsymbol{\phi}}}^\dagger\vec{\underline{\boldsymbol{\phi}}}\right\} = t. \quad (5.1)$$

III. Throughput Analysis and Optimized Power Allocation

Since the above considered block-fading model of Section II guarantees that the MIMO channel path gains $\mathbf{H}$ change to new independent realizations every T signaling periods, the MIMO-channel is information stable and ergodic [23], and then, in principle, the payload data can be reliably transmitted over the MIMO channel at any rate below the corresponding ergodic capacity C. Following standard approaches [19], [24, p. 361], this last can be directly expressed as

$$C = E\{C(\widehat{\mathbf{H}})\} \equiv \int C(\widehat{\mathbf{H}})p(\widehat{\mathbf{H}})d\widehat{\mathbf{H}}, \quad \left(\frac{\text{nats}}{\text{payload slot}}\right) \quad (6)$$

where

$$p(\widehat{\mathbf{H}}) = \left(\frac{1}{(\pi(1-\sigma_\varepsilon^2))}\right)^{rt}\exp\left(-\frac{1}{(1-\sigma_\varepsilon^2)}\mathrm{TR}[\widehat{\mathbf{H}}^\dagger\widehat{\mathbf{H}}]\right) \quad (6.1)$$

is the joined pdf of the MMSE channel estimates in (2), whereas the random variable

$$C(\widehat{\mathbf{H}}) \triangleq \sup_{\vec{\underline{\boldsymbol{\phi}}}:E\left\{\vec{\underline{\boldsymbol{\phi}}}^\dagger\vec{\underline{\boldsymbol{\phi}}}\right\}\le t}\left(\frac{1}{T_{\rm pay}}\right) I(\vec{\underline{\mathbf{y}}};\vec{\underline{\boldsymbol{\phi}}}|\widehat{\mathbf{H}}) \quad (7)$$

is the capacity of the MIMO channel in (5), *conditioned* on the realization $\widehat{\mathbf{H}}$ of the channel estimates available at both transmitter and receiver. Finally, $I(\cdot;\cdot|\cdot)$ in (7) is the mutual information conveyed by the MIMO channel of (5) during the payload phase.

Although the probability density functions (pdfs) of the input signals achieving the supremum in (7) are known for the limit cases of PCSI [1], [7] and NCSI [3], until now, they are *unknown* for the general case of IPCSI [9], [10], [18]. As a consequence, at present, *no* analytical, semi-analytical, or numerical tools are available for the evaluation of the capacity in (7) in the general case of IPCSI (e.g., for $0 < \sigma_\varepsilon^2 < 1$). Therefore, as in [25, Sec. IV], we proceed to evaluate the conditional information throughput $I(\cdot;\cdot|\cdot)$ in (7) by assuming Gaussian distributed input signals. This means that the $T_{\rm pay}$ components $\left\{\underline{\boldsymbol{\phi}}(n)\in\mathbb{C}^t,\ T_{\rm tr}+1\le n\le T\right\}$ of the signal vector $\vec{\underline{\boldsymbol{\phi}}}$ in (5) are assumed mutually independent equally distributed zero-mean proper complex Gaussian t-dimensional vectors with correlation matrix $\mathbf{R}_{\boldsymbol{\phi}}$ satisfying the constraint in (4.1). Hence, after indicating as

$$\mathrm{TR}_G(\widehat{\mathbf{H}}) \triangleq \sup_{\mathrm{Tr}[\mathbf{R}_{\boldsymbol{\phi}}]\le(t/T_{\rm pay})} I(\vec{\underline{\mathbf{y}}};\vec{\underline{\boldsymbol{\phi}}}|\widehat{\mathbf{H}}) \quad \left(\frac{\text{nats}}{\text{payload-slot}}\right) \quad (8)$$

the maximal ergodic throughput conveyed by the MIMO channel of Fig. 1 for the above-considered input Gaussian pdf, in general, we have $\mathrm{TR}_G(\widehat{\mathbf{H}}) \le C(\widehat{\mathbf{H}})$. However, the above inequality is satisfied as equality when the available channel estimates $\widehat{\mathbf{H}}$ are perfect (that is, for $\sigma_\varepsilon^2 \to 0$ [1], [7]) or when no CSI is available at the receiver but the length $T_{\rm pay}$ of the payload phase largely exceeds the number t of the transmit antennas (see [3, Sec. V] for more details on this important asymptotic result). Therefore, when at least one of the above operating condition is met, the assumed Gaussian pdf is the one achieving capacity. An analysis of the ergodic throughput sustained by MIMO channels with imperfect channel estimates for isotropically distributed unitary input signals may be found in [2]. Due to space limitations, it will be not reported here.

A. Analytical Evaluation of $TR_G(\widehat{\mathbf{H}})$

Although the conditional mutual information $I(\vec{\underline{\mathbf{y}}};\vec{\underline{\boldsymbol{\phi}}}|\widehat{\mathbf{H}})$ in (8) generally resists closed-form computation, the considerations developed in the first part of the Appendix A lead to the following result.

Proposition 1: Let us assume that we are assigned the correlation matrix $\mathbf{R}_{\boldsymbol{\phi}}$ of the mutually independent equally distributed Gaussian input signals $\left\{\underline{\boldsymbol{\phi}}(n)\in\mathbb{C}^t, T_{\rm tr}+1\le n\le T\right\}$ in (4). Thus, the resulting conditional mutual information conveyed by the MIMO channel of Fig. 1 equates

$$I(\vec{\underline{\mathbf{y}}};\vec{\underline{\boldsymbol{\phi}}}|\widehat{\mathbf{H}}) \le T_{\rm pay}\lg\det\left[\mathbf{I}_r + \frac{\gamma T_{\rm pay}}{t(1+\gamma\sigma_\varepsilon^2)}\widehat{\mathbf{H}}^T\mathbf{R}_{\boldsymbol{\phi}}\widehat{\mathbf{H}}^*\right] - r\lg\frac{\det\left[\mathbf{I}_t + \frac{\sigma_\varepsilon^2\gamma}{t}T_{\rm pay}^2\mathbf{R}_{\boldsymbol{\phi}}\right]}{(1+\gamma\sigma_\varepsilon^2)^{T_{\rm pay}}} \quad (9)$$

with equality if (but not iff) one of the following conditions is met:

a) both $T_{\rm pay}$ and t are large (9.1)

b) σ_ε^2 vanishes (9.2)

c) SNR γ vanishes. $\diamond$ (9.3)

Before proceeding, it may be of interest to shortly point out the relative roles played by two terms present at the r.h.s of (9).

Specifically, the first one explicitly accounts for the available channel estimates $\hat{\mathbf{H}}$ so that this term is the dominating one[5] for $\sigma_\varepsilon^2 \to 0$, while it vanishes.[6] On the contrary, the second term at the r.h.s. of (9) dictates the information throughput sustained by the considered system in noncoherent applications, and then, it dominates for $\sigma_\varepsilon^2 \to 1$ (e.g., for $\hat{\mathbf{H}} \to \mathbf{0}$), while it vanishes for $\sigma_\varepsilon^2 \to 0$ (e.g., for $\hat{\mathbf{H}} \to \mathbf{H}$). Finally, about the asymptotic nature of the condition in (9.1), Appendix A points out that it arises from a combined application of the law of large numbers and the central limit theorem. However, it may be numerically ascertained that the condition in (9.1) is, by fact, virtually met even for values of t and $T_{\rm pay}$ limited and as low as $t = 5$ and $T_{\rm pay} = 7t$. Therefore, since in our framework the length $T = T_{\rm pay} + T_{\rm tr}$ of the transmitted packets equates the coherence time $T_{\rm coh}$ of the underlying MIMO channel (see Note 1), we argue that in practice, channel coherence times of the order of $T_{\rm tr} + 35$ suffice to apply (9) even for nonvanishing values of σ_ε^2 and γ.

Optimized Powers Allocation Over the Transmit Antennas: Therefore, according to (8), the next step concerns the maximization of the throughput in (9) under the power constraint (4.1). Toward this end, let us introduce the Singular Value Decomposition (SVD)

$$\widehat{\mathbf{H}} = \widehat{\mathbf{U}}\hat{\mathbf{\Lambda}}\widehat{\mathbf{V}}^{\dagger} \tag{10}$$

of the available MIMO channel estimates $\widehat{\mathbf{H}}$, where $\widehat{\mathbf{U}}$ and $\widehat{\mathbf{V}}$ are the unitary matrices with columns filled with the right and left eigenvectors of $\widehat{\mathbf{H}}\widehat{\mathbf{H}}^{\dagger}$ and $\widehat{\mathbf{H}}^{\dagger}\widehat{\mathbf{H}}$ respectively, while

$$\hat{\mathbf{\Lambda}} \triangleq \text{diag}\left\{\widehat{\lambda}_1, \ldots, \widehat{\lambda}_s, \underline{\mathbf{0}}_{t-s}\right\} \tag{10.1}$$

is the corresponding $\mathbb{R}^{t\times r}$ diagonal matrix of the resulting $s \triangleq \min\{r, t\}$ *magnitude-ordered* singular values $\widehat{\lambda}_1 \geq \widehat{\lambda}_2 \geq \cdots \geq \widehat{\lambda}_s$. Thus, after introducing the dummy positions

$$\alpha_i \triangleq \frac{\gamma T_{\rm pay}\widehat{\lambda}_i^2}{t(1+\gamma\sigma_\varepsilon^2)}, \quad 1 \leq i \leq s, \;\; \beta \triangleq \frac{\sigma_\varepsilon^2\gamma}{t}T_{\rm pay}^2 \tag{11}$$

in Appendix A-1, we prove that a suitable application of the Kuhn–Tucker conditions [27] to the objective function in (9) allows us to compute the optimal transmit powers $\{P^\star(i), 1 \leq i \leq t\}$ achieving the constrained supremum in (8), as detailed by the following proposition.

[5]The second term is zeroed when $\sigma_\varepsilon^2 = 0$ because $\lg\det \mathbf{I}_r$ equates to zero.

[6]The first term is equal to 0 when $\sigma_\varepsilon^2 = 1$ because in this case, $\hat{\mathbf{H}}$ vanishes (see footnote 3).

Proposition 2 (Optimized Power Allocation for Gaussian Input Signals): Let us assume that at least one of the above operating conditions (9.1)–(9.3) is met. Thus, for $i = s+1, \ldots, t$, the optimized powers achieving the supremum in (8) *vanish*, whereas for $i = 1, \ldots, s$, they must satisfy the two relationships in (12) and (13), shown at the bottom of the page. Furthermore, the non-negative parameter ρ in (12) and (13) is chosen to satisfy the following power constraint [see (4.1)]:

$$\sum_{i\in\mathfrak{I}(\rho)} P^\star(i) = \frac{t}{T_{\rm pay}} \tag{14}$$

where

$$\mathfrak{I}(\rho) \triangleq \left\{i = 1, \ldots, s : \widehat{\lambda}_i^2 > \left(1+\gamma\sigma_\varepsilon^2\right)\left(r\sigma_\varepsilon^2 + \frac{t}{\gamma T_{\rm pay}\rho}\right)\right\} \tag{14.1}$$

is the ρ–depending set of i–indexes meeting the inequality in (13). Finally, the corresponding optimized correlation matrix[7] for the Gaussian-distributed input signals equates to

$$\mathbf{R}_{\boldsymbol{\phi}}^{(opt.)} = \widehat{\mathbf{U}}^{*}\text{diag}\left\{P^\star(1), \ldots, P^\star(s), \underline{\mathbf{0}}_{t-s}\right\}\widehat{\mathbf{U}}^{T} \tag{15}$$

whereas the resulting maximal throughput $\text{TR}_G(\widehat{\mathbf{H}})$ in (8) is directly computable as in

$$\text{TR}_G(\widehat{\mathbf{H}}) = r\lg\left(1+\gamma\sigma_\varepsilon^2\right) + \sum_{i=1}^{s}\lg\left[\frac{1+\alpha_i P^\star(i)}{\left(1+\beta P^\star(i)\right)^{r/T_{\rm pay}}}\right] \left(\frac{\text{nats}}{\text{payload slot}}\right). \quad \diamond \tag{16}$$

In the limits as $\sigma_\varepsilon^2 \to 0$ and $\sigma_\varepsilon^2 \to 1$, the relationships presented in (12) and (13) for the optimized power allocation assume interesting forms: We examine these briefly in the following Remarks.

Remark 1 (Case of PCSI): For $\sigma_\varepsilon^2 \to 0$, it can be easily viewed that the corresponding limit forms assumed by (12) and (13) may be collected in the following simple water-filling like expression [see (A.13) of Appendix A]:

$$P^\star(i) = \max\left\{0, \rho - \frac{t}{\gamma T_{\rm pay}\widehat{\lambda}_i^2}\right\}, \quad 1 \leq i \leq s. \tag{17}$$

Thus, we conclude that the relationships (12) and (13) *generalize* the standard water-filling ones, and they *reduce* to them when the available channel estimates $\widehat{\mathbf{H}}$ are exact. $\diamond$

[7]We stress that the optimum $\mathbf{R}_{\boldsymbol{\phi}}$ presents the shape in (14) because, being a unitary matrix, the second term in (9) does not change.

$$P^\star(i) = 0, \text{ for } \widehat{\lambda}_i^2 \leq \left(1+\gamma\sigma_\varepsilon^2\right)\left(r\sigma_\varepsilon^2 + \frac{t}{\gamma T_{\rm pay}\rho}\right) \tag{12}$$

$$P^\star(i) = \frac{1}{2\beta}\left\{\beta\left[\rho\left(1-\frac{r}{T_{\rm pay}}\right) - \frac{1}{\alpha_i}\right] - 1 + \sqrt{\left\{\beta\left[\rho\left(1-\frac{r}{T_{\rm pay}}\right) - \frac{1}{\alpha_i}\right] - 1\right\}^2 + 4\beta\left(\rho - \frac{1}{\alpha_i} - \frac{\rho\beta r}{\alpha_i T_{\rm pay}}\right)}\right\}$$

$$\text{for } \widehat{\lambda}_i^2 > \left(1+\gamma\sigma_\varepsilon^2\right)\left(r\sigma_\varepsilon^2 + \frac{t}{\gamma T_{\rm pay}\rho}\right). \tag{13}$$

Remark 2 (Case of NCSI): Pausing now to consider the opposite limit case of NSCI, we observe that for σ_ε^2 approaching unity, the relationship in (9) becomes

$$\lim_{\sigma_\varepsilon^2 \to 1} I(\underline{\vec{y}}; \underline{\vec{\phi}} | \widehat{\mathbf{H}}) = r T_{\rm pay} \lg(1+\gamma) - r \lg \det \left[\mathbf{I}_t + \frac{\gamma}{t} T_{\rm pay}^2 \mathbf{R}_{\boldsymbol{\phi}} \right].$$

Therefore, since the corresponding maximizing $\mathbf{R}_{\boldsymbol{\phi}}$ exhibits only one nonzero eigenvalue equating $t/T_{\rm pay}$, we obtain for the resulting maximal throughput $\mathrm{TR}_G(\widehat{\mathbf{H}})$ in (8) the simple expression

$$\lim_{\sigma_\varepsilon^2 \to 1} \mathrm{TR}_G(\widehat{\mathbf{H}}) = r \lg \left[\frac{1+\gamma}{(1+\gamma T_{\rm pay})^{1/T_{\rm pay}}} \right], \left(\frac{\text{nats}}{\text{payload slot}} \right). \quad (18)$$

Since (18) holds for large t and $T_{\rm pay}$ (see (9.1)), we conclude that the relationship in (18) *agrees* with the conclusion originally reported in [3] about the capacity-achieving property retained for large $T_{\rm pay}$ by the input Gaussian pdf, *even* in application scenarios with no CSI at the receiver.[8] Interestingly enough, this property exhibited by (18) is *not retained* by the lower bound on $\mathrm{TR}_G(\widehat{\mathbf{H}})$ that recently appeared in [9]–[11]. Specifically, this lower bound assumes the following form [9]–[11]:

$$\mathrm{TR}_G(\hat{\mathbf{H}}) \geq \sup_{\mathrm{TR}[\mathbf{R}_{\boldsymbol{\phi}}] \leq (t/T_{\rm pay})} \lg \det \left[\mathbf{I}_r + \frac{\gamma T_{\rm pay}}{t} \hat{\mathbf{H}}^T \mathbf{R}_{\boldsymbol{\Phi}} \hat{\mathbf{H}}^* \right] \quad (18.1)$$

where the maximizing $\mathbf{R}_{\boldsymbol{\phi}}$ is evaluated via the application of the (usual) water-filling algorithm to the channel estimates matrix $\widehat{\mathbf{H}}$ [9]–[11]. Therefore, since, for $\sigma_\varepsilon^2 \to 1$ (e.g., for $\widehat{\mathbf{H}} \to 0$) the lower bound (18.1) vanishes, *regardless* of the values assumed by $T_{\rm pay}$, this lower bound is *not able* to reflect the above-mentioned capacity-achieving asymptotic property retained by the input Gaussian pdf. ◇

B. Algorithm for Computing the Optimized Power Allocation in (12) and (13)

The first step for computing the optimized powers of (12) and (13) consists of solving (nonlinear) (14) with respect to the (*a priori* unknown) ρ parameter. Although the computation of ρ via (14) resists a closed-form formulation, nevertheless, we observe that the cardinality $|\Im(\rho)|$ of the set (14.1) vanishes at $\rho = 0$, and then, it increases for growing values of ρ. In turn, this means that the corresponding power summation present on the left-hand-side (l.h.s.) of (14) vanishes at $\rho = 0$, and then, it increases for large values of ρ. Therefore, the solution of (14) can be found via a numerical iterative procedure that starts with $\rho = 0$ and then, at each cycle, increases the current value of ρ by an assigned step-size Δ.[9] The procedure stops when the corresponding summation in (14) equates the requested value $t/T_{\rm pay}$.

[8]According to the rationales reported in [3, p. 145], for large $T_{\rm pay}$, we may assume that a small portion of the payload stream is employed for sending training data from which the receiver can estimate the channel coefficients $\{h_{ji}\}$. Therefore, for $T_{\rm pay} \to \infty$, $\mathbf{H}$ may be assumed perfectly known to the receiver, and then, the capacity achieving pdf becomes the Gaussian one. This conjecture done in [3, Sec. V-A] is explicitly supported by (18).

[9]We have ascertained via numerical tests that $\Delta = 0.1t/T_{\rm pay}$ suffices for computing the values of ρ with suitable accuracy.

TABLE I
ALGORITHM FOR THE OPTIMIZED POWER ALLOCATION FOR GAUSSIAN INPUT SIGNALS AND IDEAL FEEDBACK LINK

1. Compute the SVD in Eq. (10) of $\widehat{\mathbf{H}}$;
2. Perform the magnitude-ordering of the singular values of $\widehat{\mathbf{H}}$;
3. Set $P^\star(i) = 0,\ 1 \leq i \leq t$;
4. Set $\rho = 0$ and $\Im(\rho) = \emptyset$;
5. Set the step-size Δ;
6 While $\left(\sum_{i \in \Im(\rho)} P^\star(i) < t/T_{pay} \right)$ do
7. Update $\rho = \rho + \Delta$;
8. Update $\Im(\rho)$ via Eq. (14.1);
9. Compute the powers $\{P^\star(i),\ i \in \Im(\rho)\}$ via Eq. (13);
10. end;
11. Compute the optimized powers $\{P^\star(i),\ 1 \leq i \leq s\}$ via Eqs. (12), (13);
12. Compute $\mathbf{R}_{\boldsymbol{\phi}}(opt.)$ via Eq. (15);
13. Compute the conveyed throughput $TR_G(\widehat{\mathbf{H}})$ via Eq. (16).

The resulting algorithm for computing the optimized power allocation of (12) and (13) is reported in Table I. Section IV reports some numerical performance plots obtained via computer implementation of the algorithm of Table I.

IV. Numerical Results and Comparisons About the Ergodic Throughput

Although the pdf of the channel estimates $\widehat{\mathbf{H}}$ assumes the simple form in (6.1), the corresponding expectation

$$\mathrm{TR}_G \triangleq E\left\{ \mathrm{TR}_G(\widehat{\mathbf{H}}) \right\} \quad (19)$$

of the conditional throughput in (8) resists closed-form analytical computations *even* in the limit cases of $\sigma_\varepsilon^2 \to 0$ and $\sigma_\varepsilon^2 \to 1$ (see [18] for the state-of-art about this topic). Therefore, as in [7], we proceed to evaluate the average throughput of (19) via a Monte Carlo approach based on the sample-averages of 10 000 independent realizations of $\mathrm{TR}_G(\widehat{\mathbf{H}})$. The resulting values obtained for TR_G in (19) are drawn in Figs. 2 and 3 for some test scenarios. Both figures allow us to appreciate the combined effect of σ_ε^2 and γ on TR_G. Specifically, Fig. 2 points out that the loss induced on TR_G by channel estimation errors is limited up to few nats at medium SNRs for σ_ε^2 below 0.01, whereas this loss increases quickly beyond 15 nats when σ_ε^2 exceeds 0.1. Furthermore, Fig. 3 shows that the sensitivity of TR_G on σ_ε^2 is high for low γ values, whereas it decreases for growing SNRs.

A. Advantages of Power Allocation Algorithm With Respect to Uniform Allocation

Therefore, since the reported plots lead to the conclusion that the actual effectiveness of the power allocation algorithm decreases quickly for increasing values of σ_ε^2, we conclude that, in practice, the introduction of an (ideal) feedback link in the communication system of Fig. 1 may be *justified only when* the average throughput TR_G in (19) achieved via the power allocation algorithm of Table I outperforms the corresponding one

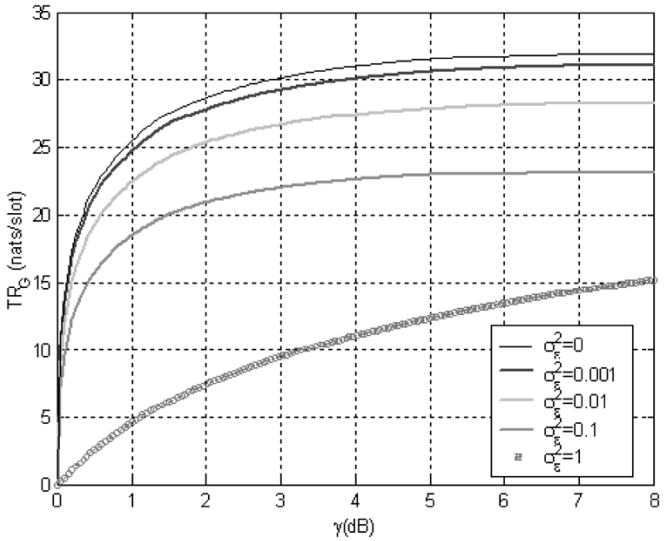

Fig. 2. Average throughput (TR_G) for different values of σ_ε^2 ($T_{\mathrm{pay}} = 80$, $t = 12, r = 6$).

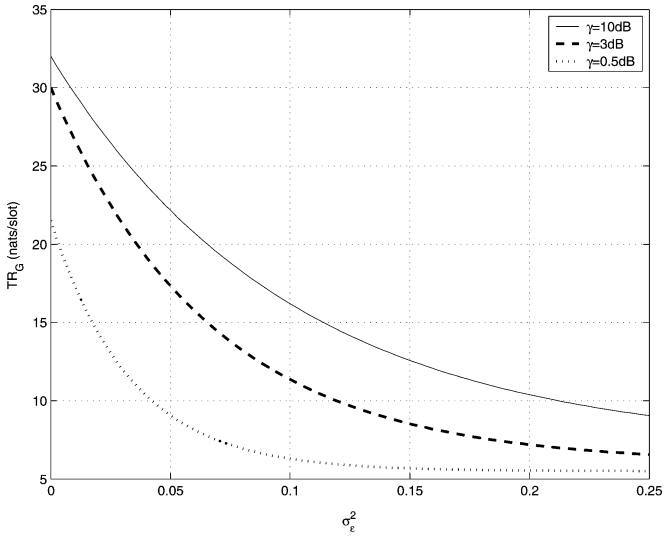

Fig. 3. Average throughput (TR_G) for various values of γ ($T_{\mathrm{pay}} = 80, t = 12, r = 6$).

$\mathrm{TR}^{(0)} \triangleq E\{\mathrm{TR}^{(0)}(\widehat{\mathbf{H}})\}$ attained by *evenly* splitting the available total power over all transmit antennas. Therefore, values beyond unity for the resulting *Power Allocation Gain* (PAG)

$$\mathrm{PAG} \triangleq \frac{\mathrm{TR}_G}{\mathrm{TR}^{(0)}} \equiv \frac{E\left\{\mathrm{TR}_G(\widehat{\mathbf{H}})\right\}}{E\left\{\mathrm{TR}^{(0)}(\widehat{\mathbf{H}})\right\}} \tag{20}$$

support the utilization of an (ideal) feedback link, whereas PAG values approaching unit indicate uselessness of any feedback information. A three-dimensional (3-D) plot of the PAG of (20) is reported in Fig. 4 for a system equipped with 12 transmit and six receive antennas. This plot allows us to appreciate the dependence of the PAG on both σ_ε^2 and operating SNR γ, and in fact, it points out that the PAG values approach unit for σ_ε^2 exceeding 0.25. Finally, the joined effects of σ_ε^2 and T_{pay} on TR_G may be understood by the 3-D plot of Fig. 5. Interestingly enough, the 3-D plot of Fig. 5 stresses that TR_G quickly approaches its ultimate limit attained at $\sigma_\varepsilon^2 = 0$ also in application scenarios

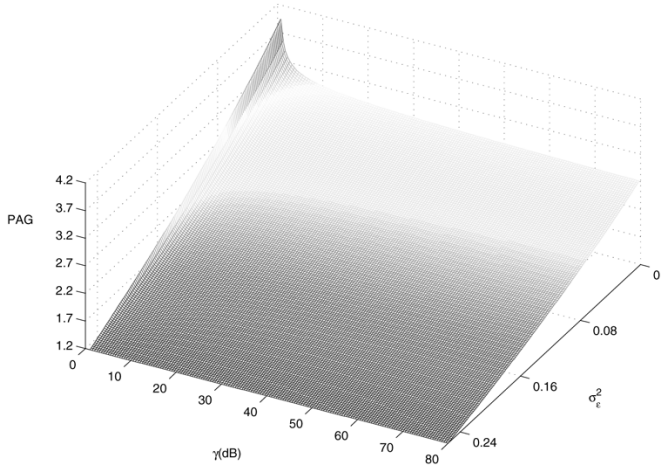

Fig. 4. Power Allocation Gain (PAG) for $T_{\mathrm{pay}} = 80, t = 12, r = 6$.

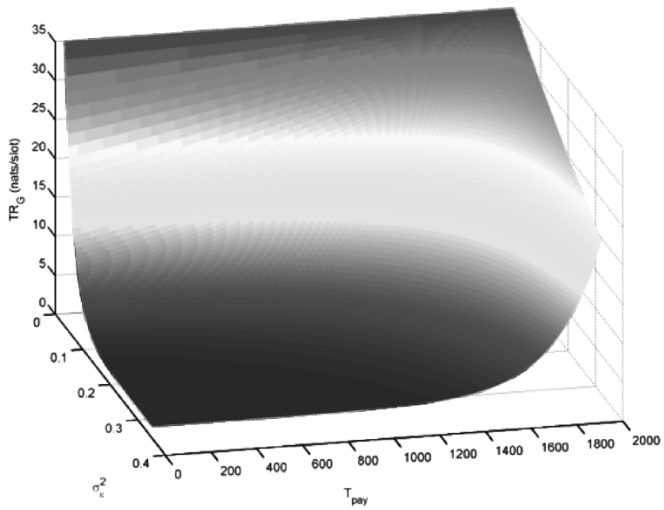

Fig. 5. Average Throughput for various values of σ_ε^2 and T_{pay}.

impaired by channel estimation errors, *provided that* T_{pay} is sufficiently large and on the order of about 1000 slots. Therefore, the plot in Fig. 5 directly supports the conjecture of [3] about the asymptotic optimality of the Gaussian pdf for the input signals of MIMO systems in Fig. 2, even in the presence of *substantial* channel estimation errors at the receiver.

B. Sensitivity of the Power Allocation Algorithm on Error Introduced by the Feedback Link

To test the sensitivity and robustness of the power allocation algorithm of Table I on errors (due to noise, feedback delay, and quantization effects) possibly introduced by the feedback link of Fig. 1, we have perturbed the channel-estimate matrix $\widehat{\mathbf{H}}$ that is available at the receiver via a randomly generated $\mathbb{C}^{t\times r}$ matrix $\mathbf{N}$ composed by zero-mean proper complex unit-variance independent Gaussian samples. We have generated the corresponding channel estimates $\widehat{\mathbf{H}}'$ available at the output of feedback link (see Fig. 1) according to the following relationship:

$$\widehat{\mathbf{H}}' = \widehat{\mathbf{H}} + \sqrt{1-\sigma_\varepsilon^2}\sqrt{\delta}\mathbf{N} \tag{21}$$

where $\delta \triangleq E\left\{\|\widehat{\mathbf{H}}' - \widehat{\mathbf{H}}\|_E^2\right\} / E\left\{\|\widehat{\mathbf{H}}\|_E^2\right\}$ is a deterministic parameter set according to the desired average squared error

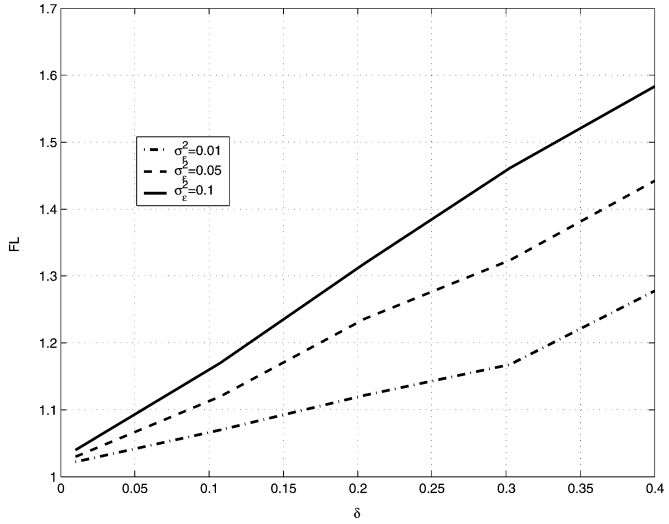

Fig. 6. Feedback loss for $T_{\rm pay} = 80, t = 12, r = 6$.

between $\widehat{\mathbf{H}}'$ and $\widehat{\mathbf{H}}$. Afterwards, *using* the perturbed $\widehat{\mathbf{H}}'$ *in place* of $\widehat{\mathbf{H}}$, we have performed the power allocation dictated by the relationships in (12) and (13), and then, we have computed the resulting conditional throughput $\mathrm{TR}_G^{(f)}(\widehat{\mathbf{H}}')$ via the expression at the r.h.s. of (16). Finally, we have numerically evaluated the resulting average Feedback Loss (FL)

$$\mathrm{FL} \triangleq \frac{E\{\mathrm{TR}_G(\widehat{\mathbf{H}})\}}{E\left\{\mathrm{TR}_G^{(f)}(\widehat{\mathbf{H}}')\right\}} \equiv \frac{\mathrm{TR}_G}{\mathrm{TR}_G^{(f)}} \tag{22}$$

for some test cases. The plots reported in Fig. 6 support the main conclusion that the FL is negligible for values of δ limited up to 0.1, but it increases quickly for δ over 0.3, especially when σ_ε^2 is below 0.1 (e.g., the channel estimates $\widehat{\mathbf{H}}$ available at the receiver become reliable).

C. Outage Analysis and Related System Performance

Since the ergodic capacity in (6) dictates the maximum conveyed information rate averaged over *all* fading states of the underlying MIMO channel, it may represent a meaningful system performance index for nonreal-time data services. However, the ergodic capacity may no longer be a relevant performance index for real-time applications (as, for example, voice and video) in (very) slow fading environments, where the maximum admissible delay T is (largely) below the coherence time $T_{\rm coh}$ of the underlying MIMO channel [18], [19]. In fact, when the coherence time $T_{\rm coh}$ of the MIMO channel of Fig. 1 largely exceeds the length T of the transmitted packet, the MIMO channel cannot be considered ergodic any longer, and,\ in this case, the information outage probability

$$P_{\rm out}(\eta) \triangleq P(C(\widehat{\mathbf{H}}) \leq \eta) \tag{23}$$

becomes a meaningful performance index [1], [7], [19]. Unfortunately, no analytical formulas are currently available for the computation of $C(\widehat{\mathbf{H}})$ in (7) for the general case of $0 < \sigma_\varepsilon^2 < 1$. However, the considerations reported in the first part of Section III lead to the conclusion that $\mathrm{TR}_G(\widehat{\mathbf{H}})$ in (8) approaches $C(\widehat{\mathbf{H}})$ for large $T_{\rm pay}$ or vanishing σ_ε^2 so that, *at least in these operating conditions*, we may write[10]

$$P_{\rm out}(\eta) = P(\mathrm{TR}_G(\widehat{\mathbf{H}}) \leq \eta), \quad \left(\text{large } T_{\rm pay} \text{ or small } \sigma_\varepsilon^2\right). \tag{24}$$

Unfortunately, due to the nonlinear nature of the relationship in (16) linking $\mathrm{TR}_G(\widehat{\mathbf{H}})$ to the optimized powers $\{P^\star(i), 1 \leq i \leq t\}$ output of the algorithm of Table I, the computation of $P(\mathrm{TR}_G(\widehat{\mathbf{H}}) \leq \eta)$ resists closed-form evaluation, *even* in the limit case when $\sigma_\varepsilon^2 \to 0$ [7]. However, in Appendix B, it is proved that for an ideal feedback link, the probability (24) may be upper bounded as in

$$P(\mathrm{TR}_G(\widehat{\mathbf{H}}) \leq \eta) \leq \left\{1 - \max_{\tau \in \{1 \ldots t\}} \left[e^{-Th_1(\eta,\tau)} \times \sum_{j=0}^{r\tau-1} \frac{1}{j!}\left(\frac{Th_1(\eta,\tau)}{2}\right)^j\right]\right\} \times u_{-1}(Th_1(\eta,\tau)) \tag{25}$$

and lower bounded as

$$P(\mathrm{TR}_G(\widehat{\mathbf{H}}) \leq \eta) \geq \left\{1 - e^{-Th_2(\eta)} \sum_{j=0}^{rt-1} \frac{1}{j!}\left(\frac{Th_2(\eta)}{2}\right)^j\right\} \times u_{-1}(Th_2(\eta)) \tag{26}$$

where $u_{-1}(\cdot)$ is the Heaviside-function (e.g., the step function), whereas the thresholds $Th_1(\cdot)$ and $Th_2(\cdot)$ are defined as

$$Th_1(\eta,\tau) \triangleq \frac{2\tau(1+\gamma\sigma_\varepsilon^2)}{\gamma(1-\sigma_\varepsilon^2)} \times \left[e^{\eta} \frac{\left(1+\frac{\gamma\sigma_\varepsilon^2 T_{\rm pay}}{\tau}\right)^{r\tau/T_{\rm pay}}}{(1+\gamma\sigma_\varepsilon^2)^r} - 1\right] \tag{25.1}$$

$$Th_2(\eta) \triangleq \frac{2t\left(1+\gamma\sigma_\varepsilon^2\right)}{\gamma\left(1-\sigma_\varepsilon^2\right)} \times \left[e^{\eta/t} \frac{\left(1+\gamma\sigma_\varepsilon^2 T_{\rm pay}\right)^{r/tT_{\rm pay}}}{(1+\gamma\sigma_\varepsilon^2)^{r/t}} - 1\right]. \tag{26.1}$$

As the tightness of the presented bounds, we observe that the thresholds (25.1) and (26.1) approach the following limit values:

$$\lim_{\gamma\to 0} Th_1(\eta,\tau) = \lim_{\gamma\to 0} Th_2(\eta) = +\infty$$
$$\lim_{\gamma\to+\infty} Th_1(\eta,\tau) \leq 0 \text{ and } \lim_{\gamma\to+\infty} Th_2(\eta) \leq 0 \tag{27}$$
$$\lim_{\eta\to\infty} Th_1(\eta,\tau) = \lim_{\eta\to\infty} Th_2(\eta) = +\infty$$
$$\lim_{\eta\to 0} Th_1(\eta,\tau) \leq 0 \text{ and } \lim_{\eta\to 0} Th_2(\eta) \leq 0 \tag{28}$$

so that the corresponding bounds in (25) and (26) are guaranteed to approach $P(\mathrm{TR}_G(\widehat{\mathbf{H}}) \leq \eta)$ for large and small SNRs γ, as well as for large and small values of the outage parameter η. The numerical plots of Fig. 7 agree with these analytical conclusions. In addition, they also show that the relative gap between our bounds and actual $P(\mathrm{TR}_G(\widehat{\mathbf{H}}) \leq \eta)$ is of the order of about

[10]For small $T_{\rm pay}$ and large σ_ε^2, the r.h.s. of (24) reduces to an upper bound on the actual outage probability $P_{\rm out}(\eta)$.

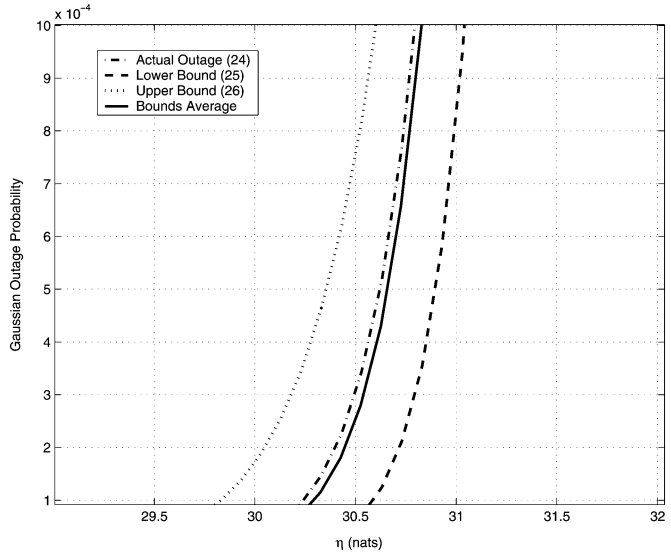

Fig. 7. Comparison between actual outage in (24) and corresponding lower and upper bounds of (25) and (26) for a system with $T_{\rm pay} = 80$, $t = 12$, $r = 6$, $\sigma_\varepsilon^2 = 10^{-3}$, $\gamma = 10$ dB.

$\pm 1\%$ over the overall range of considered outages. Interestingly enough, the arithmetical average of the bounds in (25) and (26) virtually overlaps the actual $P(\mathrm{TR}_G(\widehat{\mathbf{H}}) \leq \eta)$ so that we argue that *in practice*, this arithmetical average may give rise to reliable evaluations of the corresponding $P(\mathrm{TR}_G(\widehat{\mathbf{H}}) \leq \eta)$.

V. Conclusions

From the main results presented in this paper, two basic guidelines emerge about the optimized design of the MIMO system of Fig. 1.

First, the numerical results previously detailed in the first part of Section IV point out that the power allocation algorithm gives rise to noticeable improvements in the conveyed ergodic throughput when the estimates $\widehat{\mathbf{H}}$ available at the receiver are sufficiently reliable (we say for values of σ_ε^2 in (2.1) below about 0.25). Second, the numerical plots detailed in Section IV-B lead to the conclusion that the throughput loss induced in the proposed power allocation algorithm by noisy feedback channels is limited for δ in (21) below 0.2, but it increases quickly for δ exceeding 0.3.

Overall, an interesting problem arising from the above conclusions concerns the actual design of practical space-time codes that approach the performance of Gaussian input signals at low σ_ε^2. Some results along this direction have been recently reported in [22], where a new class of space-time codes "self-matching" to the channel estimation errors has been presented.

Appendix A
Derivation of the Throughput Formula of Section III-A

To begin with, we develop the mutual information $I(\underline{\vec{\mathbf{y}}}; \underline{\vec{\boldsymbol{\phi}}} | \widehat{\mathbf{H}})$ conditioned on the available realization $\widehat{\mathbf{H}}$ of the channel estimates as

$$I(\underline{\vec{\mathbf{y}}}; \underline{\vec{\boldsymbol{\phi}}} | \widehat{\mathbf{H}}) = h(\underline{\vec{\mathbf{y}}} | \widehat{\mathbf{H}}) - h(\underline{\vec{\mathbf{y}}} | \underline{\vec{\boldsymbol{\phi}}}, \widehat{\mathbf{H}}) \tag{A.1}$$

where $h(\cdot)$ denotes the differential entropy operator. Since, from the assumed channel model in (5), it follows that the conditional r.v. $\underline{\vec{\mathbf{y}}} | \underline{\vec{\boldsymbol{\phi}}}, \widehat{\mathbf{H}}$ is proper, Gaussian, and with covariance matrix given by

$$\mathbf{Cov}(\underline{\vec{\mathbf{y}}} | \underline{\vec{\boldsymbol{\phi}}}, \widehat{\mathbf{H}}) = \left\{ \mathbf{I}_{T_{\rm pay}} + \frac{\gamma \sigma_\varepsilon^2 T_{\rm pay}}{t} \begin{bmatrix} \underline{\boldsymbol{\phi}}^T(T_{\rm tr}+1) \\ \vdots \\ \underline{\boldsymbol{\phi}}^T(T) \end{bmatrix} \times \left[\underline{\boldsymbol{\phi}}^*(T_{\rm tr}+1) \dots \underline{\boldsymbol{\phi}}^*(T) \right] \right\} \otimes \mathbf{I}_r \tag{A.2}$$

we may develop $h(\underline{\vec{\mathbf{y}}} | \underline{\vec{\boldsymbol{\phi}}}, \widehat{\mathbf{H}})$in (A.1) as in

$$\begin{aligned} h(\vec{\mathbf{y}} | \underline{\vec{\boldsymbol{\phi}}}, \widehat{\mathbf{H}}) &\overset{(a)}{=} E_{\underline{\vec{\boldsymbol{\phi}}}} \left\{ \lg \det \left[(\pi e)^{r T_{\rm pay}} \mathbf{Cov}(\underline{\vec{\mathbf{y}}} | \underline{\vec{\boldsymbol{\phi}}}, \widehat{\mathbf{H}}) \right] \right\} \\ &\overset{(b)}{=} r T_{\rm pay} \lg(\pi e) \\ &\quad + r E_{\underline{\vec{\boldsymbol{\phi}}}} \left\{ \lg \det \left[\mathbf{I}_t + \frac{\gamma \sigma_\varepsilon^2 T_{\rm pay}^2}{t} \frac{1}{T_{\rm pay}} \times \sum_{n=T_{\rm tr}+1}^{T} \underline{\boldsymbol{\phi}}(n) \underline{\boldsymbol{\phi}}(n)^\dagger \right] \right\} \end{aligned} \tag{A.3}$$

where (a) follows from the expression for the differential entropy of a proper complex Gaussian r.v.s, whereas for deriving (b), we have exploited the properties $\det[\mathbf{I} + \mathbf{AB}] = \det[\mathbf{I} + \mathbf{BA}]$ and also $\det[\mathbf{A}] = \det[\mathbf{A}^T]$. Now, although the transmitted signal vector $\underline{\vec{\boldsymbol{\phi}}}$ is assumed proper and Gaussian, nevertheless, the corresponding expectation in (A.3) resists a closed-form analytical evaluation [7], and in any case, its computation requires multiple nested numerical integrations [7]. However, since the random vectors $\{\underline{\boldsymbol{\phi}}(n)\}$ present in the summation (A.3) are uncorrelated and equally distributed (see Section III), the law of large numbers [13, eq. (8.96), p. 265] guarantees that for large $T_{\rm pay} = T - T_{\rm tr}$, the sample average $(1/T_{\rm pay}) \sum_{n=T_{\rm tr}+1}^{T} \underline{\boldsymbol{\phi}}(n) \underline{\boldsymbol{\phi}}(n)^\dagger$ in (A.3) converges (in the mean square sense) to the corresponding expectation $\mathbf{R}_{\boldsymbol{\phi}}$. Hence, for large $T_{\rm pay}$, (A.3) assumes the following limit form:

$$\begin{aligned} h(\vec{\mathbf{y}} | \underline{\vec{\boldsymbol{\phi}}}, \widehat{\mathbf{H}}) &\equiv r T_{\rm pay} \lg(\pi e) \\ &\quad + r \lg \det \left[\mathbf{I}_t + \frac{\gamma \sigma_\varepsilon^2 T_{\rm pay}^2}{t} \mathbf{R}_{\boldsymbol{\phi}} \right], \quad (\text{large } T_{\rm pay}). \end{aligned} \tag{A.4}$$

Considering now the first differential entropy $h(\underline{\vec{\mathbf{y}}} | \widehat{\mathbf{H}}) \triangleq -E_{\underline{\vec{\mathbf{y}}} | \widehat{\mathbf{H}}} \{ \lg p(\underline{\vec{\mathbf{y}}} | \widehat{\mathbf{H}}) \}$ reported at the r.h.s. of (A.1), we stress that the conditional pdf $p(\underline{\vec{\mathbf{y}}} | \widehat{\mathbf{H}})$ therein *does not admit* a closed-form representation, even in the limit case of $r = t = 1$ (see [18] and references therein). However, directly from the channel model in (5), we deduce that the conditional random variable $\underline{\vec{\mathbf{y}}} | \widehat{\mathbf{H}}$ is zero-mean, and its covariance equates to

$$\begin{aligned} \mathbf{Cov}(\underline{\vec{\mathbf{y}}} | \widehat{\mathbf{H}}) &= (1 + \gamma \sigma_\varepsilon^2) \\ &\quad \times \left\{ \mathbf{I}_{T_{\rm pay}} \otimes \left[\mathbf{I}_r + \frac{\gamma T_{\rm pay}}{t(1 + \gamma \sigma_\varepsilon^2)} \widehat{\mathbf{H}}^T \mathbf{R}_{\boldsymbol{\phi}} \widehat{\mathbf{H}}^* \right] \right\}. \end{aligned} \tag{A.5}$$

Therefore, arguments based on the central limit theorem assure that the resulting bound

$$h\left(\underline{\vec{\mathbf{y}}} | \widehat{\mathbf{H}} \right) \leqslant \lg \left\{ (\pi e)^{r T_{\rm pay}} \det \left[\mathbf{Cov}\left(\underline{\vec{\mathbf{y}}} | \widehat{\mathbf{H}} \right) \right] \right\} \tag{A.6}$$

is satisfied as equality for large values of the number t of transmit antennas. Hence, the insertion of (A.4) and (A.5) into

(A.1) directly leads to the relationship in (9) that is valid for large t and $T_{\rm pay}$.

Finally, we observe that for vanishing values of the product $\gamma\sigma_\varepsilon^2$, the relationships in (A.3) and (A.4) converge to the same limit $rT_{\rm pay}\lg(\pi e)$, regardless of $T_{\rm pay}$. In addition, for $\gamma \to 0$, the observation $\vec{\underline{\mathbf{y}}}$ converges (with Probability 1) to the Gaussian random variable $\vec{\underline{\mathbf{w}}}$ [see (5)], whereas for $\sigma_\varepsilon^2 \to 0$, the channel estimates $\widehat{\mathbf{H}}$ approach (in the mean-square sense) the actual ones $\mathbf{H}$ (see (2.1)) so that the r.v. $\vec{\underline{\mathbf{y}}}|\widehat{\mathbf{H}}$ converges to the corresponding Gaussian random variable $\vec{\underline{\mathbf{y}}}|\mathbf{H}$. These last considerations explicitly prove the validity of (9) even under the operating conditions reported in (9.2) and (9.3).

1. Proof of the Optimized Power Allocation of (12) and (13)

After introducing into (9) the SVD in (10) for $\widehat{\mathbf{H}}$, quite standard (but tedious) algebraic manipulations allow us to put the relationship of (9) in the equivalent form of (A.7), shown at the bottom of the page, where

$$f(P^\star(1),\ldots,P^\star(s)) \triangleq \sum_{i=1}^{s} f_i\left(P^\star(i)\right) \equiv \sum_{i=1}^{s}\left[\lg\left(1+\alpha_i P^\star(i)\right) - \frac{r}{T_{\rm pay}}\lg\left(1+\beta P^\star(i)\right)\right] \tag{A.8}$$

is an additive objective function depending *only* on the powers $\{P^\star(i), 1 \le i \le s\}$ to be allotted to the first s transmit antennas. Thus, since the supremum in (A.7) may be achieved only for $P^\star(i) = 0$, $s+1 \le i \le t$, we can rewrite (A.7) as

$$\mathrm{TR}_G\left(\hat{\mathbf{H}}\right) = r\lg\left(1+\gamma\sigma_\varepsilon^2\right) + \sup_{\sum_{i=1}^{s} P^\star(i) \leqslant (t/T_{\rm pay})} \left\{f\left(P^\star(1),\ldots,P^\star(s)\right)\right\}. \tag{A.9}$$

Now, an examination of the sign of the second derivatives $\partial^2 f_i\left(P^\star(i)\right)/\partial P^\star(i)^2$ of the logarithmic functions $f_i\left(P^\star(i)\right)$ in (A.8) leads to the conclusions that the resulting sum-function $f\left(P^\star(1),\ldots,P^\star(s)\right)$ is guaranteed to be $\cap$-convex (at-least) over the region $\mathfrak{D}$ of $\mathbb{R}^s$ given by

$$\mathfrak{D} \triangleq \left\{(P^\star(1),\ldots,P^\star(s)) : P^\star(i) \ge \max\left\{0, \frac{\beta - \alpha_i\sqrt{\frac{T_{\rm pay}}{r}}}{\beta\alpha_i\left(\sqrt{\frac{T_{\rm pay}}{r}}-1\right)}\right\},\ i=1,\ldots,s\right\}. \tag{A.10}$$

This last region approaches the overall positive orthant $\mathbb{R}_+^s$ of $\mathbb{R}^s$ when σ_ε^2 vanishes or when $T_{\rm pay}$ grows unbounded. Therefore, after assuming having met at least one of the above operating conditions,[11] we may apply the Kuhn–Tucker conditions [27], [14, eqs. (4.4.10), (4.4.11)] to search for the constrained maximum of the objective function in (A.8). In so doing, we arrive at the following relationships:

$$\alpha_i\left(1+\alpha_i P^\star(i)\right)^{-1} - \frac{r}{T_{\rm pay}}\beta\left(1+\beta P^\star(i)\right)^{-1} \le \frac{1}{\rho} \quad \text{for all } i \text{ such that } : P^\star(i) = 0 \tag{A.11}$$

$$\alpha_i\left(1+\alpha_i P^\star(i)\right)^{-1} - \frac{r}{T_{\rm pay}}\beta\left(1+\beta P^\star(i)\right)^{-1} \equiv \frac{1}{\rho} \quad \text{for all } i \text{ such that } : P^\star(i) > 0 \tag{A.12}$$

where the expression on the r.h.s. of (A.11) and (A.12) is the derivative of the objective function in (A.8) with respect to the ith variable $P^\star(i)$. Now, (A.11) directly leads to (12), whereas (A.12) may be rewritten in the following equivalent form:

$$\beta P^*(i)^2 + \left\{1 - \beta\left[\left(1-\frac{r}{T_{\rm pay}}\right)\rho - \frac{1}{\alpha_i}\right]\right\}P^*(i) - \frac{\rho}{\alpha_i}\left[\alpha_i - \frac{1}{\rho} - \frac{\beta r}{T_{\rm pay}}\right] = 0, \text{ for } P^*(i) > 0. \tag{A.13}$$

Therefore, after resolving (A.13) with respect to $P^\star(i)$ and, then, retaining only the corresponding non-negative root, we directly obtain (13).

Appendix B
Derivation of the Outage Bounds Given by (25) and (26)

After assuming[12] $t \le r$, let us indicate as $\mu_1 \ge \mu_2 \ge \cdots \ge \mu_t$ the magnitude-ordered eigenvalues of the $\mathbb{C}^{t\times t}$ Wishart matrix

$$\mathfrak{W} \triangleq \frac{2}{1-\sigma_\varepsilon^2}\widehat{\mathbf{H}}\widehat{\mathbf{H}}^\dagger. \tag{B.1}$$

Therefore, since α_i in (11) may be rewritten as $\alpha_i = \gamma\mu_i T_{\rm pay}(1-\sigma_\varepsilon^2)/(2t(1+\gamma\sigma_\varepsilon^2))$, a direct exploitation of the relationship (16) allows us to express the probability $P(\mathrm{TR}_G(\widehat{\mathbf{H}}) \le \eta)$ as in

$$P(\mathrm{TR}_G(\widehat{\mathbf{H}}) \le \eta) = P\left(\chi \le \eta - r\lg\left(1+\gamma\sigma_\varepsilon^2\right)\right) \tag{B.2}$$

[11] From a practical point of view, simple-to-test sufficient conditions guaranteeing $\cap$-convexity of the objective function in (A.8) over overall positive orthant $\mathbb{R}_+^s$ for $\sigma_\varepsilon^2 \ge 0$ for any *finite* value of $T_{\rm pay} \ge r$ are the following ones: $\hat{\lambda}_i^2 \ge (1+\sigma_\varepsilon^2)\sqrt{rT_{\rm pay}}$, $i = 1,..,s$. In fact, when the last of these are met, all max's in (A.10) vanish, and then, $f(\cdot)$ in (A.8) is $\cap$-convex over overall $\mathbb{R}_+^s$. In all the carried out numerical tests, we have ascertained the satisfaction of the above conditions.

[12] After replacing $\widehat{\mathbf{H}}\widehat{\mathbf{H}}^\dagger$ in (B.1) by $\widehat{\mathbf{H}}^\dagger\widehat{\mathbf{H}}$, the same developments and conclusions also hold for the case of $t \ge r$.

$$\mathrm{TR}_G\left(\hat{\mathbf{H}}\right) = r\lg\left(1+\gamma\sigma_\varepsilon^2\right) + \sup_{\sum_{i=1}^{t} P^\star(i) \leqslant (t/T_{\rm pay})} \left\{f\left(P^\star(1),\ldots,P^\star(s)\right) - \frac{r}{T_{\rm pay}}\sum_{i=s+1}^{t}\lg\left(1+\frac{\gamma\sigma_\varepsilon^2 T_{\rm pay}^2}{t}P^\star(i)\right)\right\} \tag{A.7}$$

where the real-valued r.v. χ is defined as

$$\chi \triangleq \sup_{\sum_{i=1}^{t} P(i) \leqslant t/T_{\rm pay}} \left\{ \sum_{i=1}^{t} \lg\left(1 + \frac{\left(1-\sigma_\varepsilon^2\right)\gamma T_{\rm pay}}{2t\left(1+\gamma\sigma_\varepsilon^2\right)} P(i)\mu_i\right) - \left(\frac{r}{T_{\rm pay}}\right) \lg\left(1 + \frac{\gamma\sigma_\varepsilon^2}{t} T_{\rm pay}^2 P(i)\right)\right\}. \quad \text{(B.3)}$$

Although the joint pdf of the eigenvalues $\{\mu(i),\ 1 \leq i \leq 1\}$ of the Wishart matrix in (B.1) is well-known [16], nevertheless, the computation of the pdf of the r.v. of (B.3) resists closed-form expression, even in the limit cases of $\sigma_\varepsilon^2 \to 0$ and $\sigma_\varepsilon^2 \to 1$ (see [7] and references therein). Therefore, in the remaining part of this Appendix, we proceed to upper and lower bound the r.v. χ in (B.3).

1. Upper Bound in (25)

By generalizing a simple power allocation strategy originally pursued in [26, p. 213] for asymmetric digital subscriber line applications, it has been recently remarked in [17] that for large t and at medium-high SNRs, an effective (although *sub*optimal) power allocation policy consists of assigning equal powers to the first $\tau \geq 1$ transmit antennas and, thus, turn off the last $(t-\tau)$ ones, according to the following ON–OFF like law:

$$\widetilde{P}(i) \triangleq \frac{t}{\tau T_{\rm pay}},\ 1 \leq i \leq \tau \quad \text{and } \widetilde{P}(i) \triangleq 0,\quad \tau+1 \leq i \leq t. \quad \text{(B.4)}$$

Therefore, after replacing the above ON–OFF power allocation of (B.4) in (B.3), the resulting r.v.

$$\begin{aligned}\Theta &\triangleq \sum_{i=1}^{t} \lg\left(1 + \frac{\left(1-\sigma_\varepsilon^2\right)\gamma T_{\rm pay}}{2t\left(1+\gamma\sigma_\varepsilon^2\right)} \mu_i \tilde{P}(i)\right) - \frac{r}{T_{\rm pay}} \lg\left(1 + \frac{\gamma\sigma_\varepsilon^2}{t} T_{\rm pay}^2 \tilde{P}(i)\right) \\ &\equiv -\left(\frac{r\tau}{T_{\rm pay}}\right) \lg\left(1 + \frac{\gamma\sigma_\varepsilon^2 T_{\rm pay}}{\tau}\right) + \sum_{i=1}^{\tau} \lg\left(1 + \frac{\left(1-\sigma_\varepsilon^2\right)\gamma}{2\tau\left(1+\gamma\sigma_\varepsilon^2\right)} \mu_i\right) \end{aligned} \quad \text{(B.5)}$$

obviously lower bounds the χ one in (B.3). Thus, we can upper bound the probability in (B.2) as in

$$\begin{aligned} P\left(\mathrm{TR}_G\left(\hat{\mathbf{H}}\right) \leqslant \eta\right) &\leqslant P\left(\Theta \leqslant \eta - r\lg\left(1+\gamma\sigma_\varepsilon^2\right)\right) \\ &\equiv P\left(\lg\det\left[\mathbf{I}_\tau + \frac{\left(1-\sigma_\varepsilon^2\right)\gamma}{2\tau\left(1+\gamma\sigma_\varepsilon^2\right)} \times \mathrm{diag}\left\{\mu_1, \ldots, \mu_\tau\right\}\right] \leqslant k\right) \end{aligned} \quad \text{(B.6)}$$

where we have also introduced the dummy position

$$k \triangleq \eta + r\lg\left[\frac{\left(1 + \frac{\gamma\sigma_\varepsilon^2 T_{\rm pay}}{\tau}\right)^{\tau/T_{\rm pay}}}{1+\gamma\sigma_\varepsilon^2}\right]. \quad \text{(B.7)}$$

Hence, after exploiting the inequality $\det[\mathbf{I}+\mathbf{A}] \geq 1 + \mathrm{TR}[\mathbf{A}]$ holding for any covariance matrix, we may upper bound (B.6) as in

$$P\left(\mathrm{TR}_G\left(\hat{\mathbf{H}}\right) \leqslant \eta\right) \leqslant P\left(v \leqslant Th_1\left(\eta, \tau\right)\right) \quad \text{(B.8)}$$

where $Th_1(\eta,\tau)$ is given by (25.1), and $v \triangleq \sum_{i=1}^{\tau} \mu_i$ is the summation of the first τ eigenvalues of the Wishart matrix in (B.1). Since v is a centered chi-square r.v. with $2r\tau$ degrees of freedom, for any assigned τ, the probability at the r.h.s. of (B.8) may be directly computed by resorting to the cumulative distribution function of v. The resulting limit upper bounds the outage $P(\mathrm{TR}_G(\hat{\mathbf{H}}) \leq \eta)$ for any τ, and then, we may choose τ to minimize this bound. In so doing, we directly obtain the relationship in (25).

2. Lower Bound in (26)

For lower bounding $P(\mathrm{TR}_G(\hat{\mathbf{H}}) \leq \eta)$ in (B.2), we proceed to upper bound the r.v. χ defined in (B.3). For this purpose, we note that under the constraint $\sum_{i=1}^{t} P(i) = t/T_{\rm pay}$, the second summation in (B.3) achieves its minimum when only one $P(i)$ is strictly positive so that the following lower bound takes place:

$$\min_{\sum_{i=1}^{t} P(i) = t/T_{\rm pay}} \left\{ \sum_{i=1}^{t} \lg\left(1 + \frac{\gamma\sigma_\varepsilon^2}{t} T_{\rm pay}^2 P(i)\right)\right\} \equiv \lg\left(1 + \gamma\sigma_\varepsilon^2 T_{\rm pay}\right). \quad \text{(B.9)}$$

Therefore, after indicating as $\mu_{\max}$ the largest eigenvalue of the Wishart matrix $\mathfrak{W}$ in (B.1), we proceed to upper bound χ in (B.3) according to (B.10), shown at the top of the next page, where (a) follows from (B.9) and the definition of $\mu_{\max}$, whereas (b) arises from an application of Jensen's inequality. Although no closed-form expression is available for the pdf of $\mu_{\max}$ [16], nevertheless, we recall that $\mu_{\max}$ is upper bounded by the Euclidean norm $\|\mathfrak{W}\|_E$ of the corresponding Wishart matrix $\mathfrak{W}$ [15, p. 359] so that directly from (B.1), we obtain the following chain of inequalities:

$$\begin{aligned} \mu_{\max} \leqslant \|\mathfrak{W}\|_E &\leqslant \frac{2}{1-\sigma_\varepsilon^2} \left\|\hat{\mathbf{H}}\right\|_E^2 \\ &\equiv \frac{2}{1-\sigma_\varepsilon^2} \sum_{j=1}^{r} \sum_{i=1}^{t} \left\|\hat{h}_{ji}\right\|^2 \triangleq \Psi. \end{aligned} \quad \text{(B.11)}$$

Hence, after introducing (B.11) to (B.10), we are able to lower bound (B.2) as in

$$P\left(\mathrm{TR}_G\left(\hat{\mathbf{H}}\right) \leqslant \eta\right) \geqslant P\left(\Psi \leqslant Th_2\left(\eta\right)\right) \quad \text{(B.12)}$$

$$\chi \overset{(a)}{\leqslant} \sup_{\sum_{i=1}^{t} P(i)=t/T_{\mathrm{pay}}} \left\{ \sum_{i=1}^{t} \lg\left(1+\frac{\left(1-\sigma_{\varepsilon}^{2}\right)}{2t\left(1+\gamma\sigma_{\varepsilon}^{2}\right)}\mu_{\max}P(i)\right)\right\} - \frac{r}{T_{\mathrm{pay}}}\lg\left(1+\gamma\sigma_{\varepsilon}^{2}T_{\mathrm{pay}}\right)$$
$$\overset{(b)}{\leqslant} t\lg\left(1+\frac{\left(1-\sigma_{\varepsilon}^{2}\right)\gamma}{2t\left(1+\gamma\sigma_{\varepsilon}^{2}\right)}\mu_{\max}\right) - \frac{r}{T_{\mathrm{pay}}}\lg\left(1+\gamma\sigma_{\varepsilon}^{2}T_{\mathrm{pay}}\right) \tag{B.10}$$

where $Th_2(\cdot)$ is given by (26.1). Finally, after noting that the r.v. Ψ in (B.11) is chi-square distributed with $2rt$ degrees of freedom, the lower bound in (26) directly arises from (B.12).

REFERENCES

[1] G. J. Foschini and M. J. Gans, "On the limit of wireless communications in fading environments when using multiple antennas," *Wireless Pers. Commun.*, vol. 6, no. 3, pp. 311–325, Jun. 1998.

[2] E. Baccarelli and M. Biagi, "Throughput Analysis of MIMO Systems With Imperfect Channel Estimates,", Rome, Italy, INFO-COM Internal Tech. Rep. 04II03-2004.

[3] T. L. Marzetta and B. M. Hochwald, "Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading," *IEEE Trans. Inf. Theory*, vol. 45, no. 1, pp. 139–157, Jan. 1999.

[4] B. M. Hochwald and T. L. Marzetta, "Unitary space-time modulation for multiple-antenna communications in Rayleigh flat fading," *IEEE Trans. Inf. Theory*, vol. 16, no. 2, pp. 543–564, Mar. 2000.

[5] E. Vitotsky and U. Madhow, "Space-time transmit precoding with imperfect feedback," *IEEE Trans. Inf. Theory*, vol. 47, no. 6, pp. 2632–2639, Sep. 2001.

[6] S. Zhou and G. B. Giannakis, "Optimal transmitter eigen-beamforming and space-time block coding-based on channel-mean," in *Proc. ICASSP*, Orlando, FL, May 2002, pp. 2853–2856.

[7] C.-N. Chuan, N. D. S. Tse, J. M. Kahn, and R. A. Valenzuela, "Capacity-scaling in MIMO wireless system under correlated fading," *IEEE Trans. Inf. Theory*, vol. 48, no. 3, pp. 637–650, Mar. 2002.

[8] S. Zeadally and L. Zhang, "Enabling Gigabit network access to end users," *Proc. IEEE*, vol. 92, no. 2, pp. 341–353, Feb. 2004.

[9] J. Baltarsee, G. Fock, and H. Meyr, "Achievable rate of MIMO channels with data-aided channel-estimation and perfect interleaving," *IEEE J. Sel. Areas Commun.*, vol. 19, no. 12, pp. 2358–2368, Dec. 2001.

[10] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?," *IEEE Trans. Inf. Theory*, vol. 49, no. 4, pp. 951–963, Apr. 2003.

[11] L. Zheng and D. N. C. Tse, "Communication on the Grassman manifold: A geometric approach to the noncoherent multiple-antenna channel," *IEEE Trans. Inf. Theory*, vol. 48, no. 2, pp. 359–383, Feb. 2002.

[12] A. Narula, M. J. Lopez, M. D. Trott, and G. W. Wornell, "Efficient use of side information in multiple-antenna data transmission over fading channles," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1423–1436, Oct. 1998.

[13] A. Papoulis, *Probability, Random Variables and Stochastic Proceses*. New York: McGraw-Hill, 1965.

[14] R. G. Gallagher, *Information Theory and Reliable communication*. New York: Wiley, 1968.

[15] P. Lancaster and M. Tismenetsky, *The Theory of Matrices*, Second ed. New York: Academic, 1985.

[16] A. Edelman, "Eigenvalues and Condition Numbers of Random Matrices," Ph.D. dissertation, Math. Dept., Mass. Inst. Technol., 1989.

[17] R. S. Blum and J. H. Winters, "On optimum MIMO with antenna selection," *IEEE Commun. Lett.*, vol. 6, no. 8, pp. 322–324, Aug. 2002.

[18] S. N. Diggavi, N. A. Dhair, A. Stamoulis, and A. R. Claderbank, "Great expectations: The value of spatial diversity in wireless networks," *Proc. IEEE*, vol. 92, no. 2, pp. 219–270, Feb. 2004.

[19] E. Baccarelli, "Evaluation of the reliable data rates supported by multiple-antenna coded wireless links for QAM transmission," *IEEE J. Sel. Areas Commun.*, vol. 19, no. 2, pp. 295–304, Feb. 2001.

[20] E. Baccarelli and M. Biagi, "Error resistant space-time coding for emerging 4G-WLANs," in *Proc IEEE Wireless Commun. Networking Conf.*, vol. 1, New Orleans, LA, Mar. 16–20, 2003, pp. 72–77.

[21] J. -C. Guey, M. P. Fitz, M. R. Bell, and W. -Y. Kuo, "Signal design for transmitter diversity wireless communication systems over Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 47, no. 4, pp. 527–537, Apr. 1999.

[22] E. Baccarelli and M. Biagi, "Performance and optimized design of space-time codes for MIMO wireless systems with imperfect channel-estimates," IEEE Trans. Signal Process., vol. 52, no. 10, pp. 2911–2923, Oct. 2004, to be published.

[23] S. Verdú and T. S. Han, "A general formula for the channel capacity," *IEEE Trans. Inf. Theory*, vol. 40, no. 6, pp. 1147–1157, Jul. 1994.

[24] D. W. Bliss, K. W. Forsythe, A. O. Hero, and A. F. Yegulalp, "Environmental issues for MIMO capacity," *IEEE Trans. Signal Process.*, vol. 50, no. 9, pp. 2128–2142, Sep. 2002.

[25] B. M. Hochwald, T. L. Marzetta, and B. Hassibi, "Space-time autocoding," *IEEE Trans. Inf. Theory*, vol. 47, no. 7, pp. 2761–2781, Nov. 2001.

[26] T. Starr, J. M. Cioffi, and P. J. Silverman, *Understanding Digital Subscriber Line Technology*. Englewood Cliffs, NJ: Prentice-Hall, 1999.

[27] D. G. Luenberger, *Linear and Nonlinear Programming*, Second ed. Reading, MA: Addison-Wesley, 1984.

[28] M. Skoglund and G. Jöngren, "On the capacity of a multiple-antenna communication link with channel side information," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 3, pp. 395–405, Apr. 2003.

[29] E. Baccarelli and M. Biagi, "Power-allocation policy and optimization design of multiple-antenna systems with imperfect channel estimation," *IEEE Trans. Veh. Technol.*, vol. 53, no. 1, pp. 136–145, Jan. 2004.

[30] E. Baccarelli, M. Biagi, C. Pelizzoni, and P. Bellotti, "A novel multi-antenna impulse radio UWB transceiver for broadband high-throughput 4G WLANs," IEEE Commun. Lett., vol. 8, no. 7, pp. 419–421, Jul. 2004, to be published.

Enzo Baccarelli received the Laurea degree (summa cum laude) in electronic engineering and Ph.D. degree in communication theory and systems, both from the University "La Sapienza," Rome, Italy, in 1989 and 1992, respectively. In 1995, he received the Post-Doctorate degree in information theory and applications from INFOCOM Department, University "La Sapienza," where he also served as Research Scientist from 1996 to 1998.

Since 1998 he has been an Associate Professor in signal processing and radio communications at the University "La Sapienza." From 1990 to 1995, he served as Project Manager at SELTI ELETTRONICA Corporation, Rome, where he worked on design of modems for high-speed data transmissions. From 1996 to 1998, he attended the international project: #AC-104: "Mobile Communication Services for high-speed trains—MONSTRAIN," where he worked on equalization and coding for fast time-varying radio-mobile links. Currently, he has been attending the international project SATURN, where he is involved in development of multi-antenna-based high-bit-rate HIPERLAN-type wireless LANs. His research activities focus on the areas of systems design and coding, with specific interest in development and optimization of high-performance transmission/access systems for coded wireless/wired multimedia applications. He is the author of more than 60 international IEEE publications in these areas, and he currently serves as reviewer of several IEEE TRANSACTIONS, Journals, and Conferences. He is also co-author of two international patents on adaptive equalization and turbo-decoding for high-speed wireless and wired data-transmission systems licensed by international corporations.

Mauro Biagi (S'01–M'05) was born in Rome, Italy, in 1974. He received the Laurea degree in telecommunication engineering in 2001 from the INFO-COM Department, University of Rome "La Sapienza," from which, in January 2005, he also received the Ph.D. degree in information and communication theory.

His current research is focused on wireless communications (multiple antenna systems and ultra wideband transmission technology), mainly dealing with space-time coding techniques and power allocation in MIMO *ad hoc* networks with special attention to game theory applications. Concerning UWB, his interests are focused on transceiver desing and UWB-MIMO applications. His research is also focused on wireline communications and, in particular, bit loading algorithms and channel equalization for xDSL systems.

Cristian Pelizzoni was born in Rome, Italy, in 1977. He received the Laurea degree in telecommunication engineering in 2003 from University of Rome "La Sapienza."

He is involved in research activity on wireless communications, in particular, multiantenna ultra wideband systems.