Multiuser Detection with an Unknown Number of Users in the Large System Regime

A. Tauste Campo, A. Guillén i Fàbregas
and E. Biglieri

CUED / F-INFENG / TR 601
October 2008

Large-System Analysis of Multiuser Detection with an Unknown Number of Users

Adrià Tauste Campo, Albert Guillén i Fàbregas and Ezio Biglieri

Abstract

We analyze multiuser detection under the assumption that the number of users accessing the channel is unknown by the receiver. In this environment, users' activity must be estimated along with any other parameters such as data, power, location, etc. Our main goal is to determine the performance loss caused by the need for estimating the identities of active users, which are not known a priori. To prevent a loss of optimality, we assume that identities and data are estimated jointly, rather than in two separate steps. We examine the performance of multiuser detectors when the number of potential users is large. Statistical-physics methodologies are used to determine two macroscopic performance parameters of the detector, viz., its multiuser efficiency and its spectral efficiency. A special attention is paid to the fixed-point equation whose solution yields the multiuser efficiency of the optimal (maximum a posteriori) detector in the large signal-to-noise ratio regime. Our analysis yields closed-form approximate bounds to the minimum mean-squared error in this regime. These illustrate the set of solutions of the fixed-point equation, and their relationship with the maximum system load. Next, we study the maximum load that the detector can support for a given quality of service (specified by error probability), as well as the system spectral efficiency and the capacity of the system.

A. Tauste Campo and A. Guillén i Fàbregas are with the Department of Engineering, University of Cambridge, Trumpington Street, Cambridge CB2 1PZ, UK, e-mail: at461@cam.ac.uk, albert.guillen@eng.cam.ac.uk

Ezio Biglieri is with the Universitat Pompeu Fabra, Passeig de Circumval·lació 8, 08003 Barcelona, Spain, e-mail: ezio.biglieri@upf.edu. His work was supported in part by the Spanish Ministery of Education and Science under Project TEC2006-01428/TCM, and by the STREP project No. IST-026905 (MASCOT) within the 6th framework program of the European Commission.

I. Introduction

In multiple-access communication, the evolution of user activity may play an important role. From one time instant to the next, some new users may become active and some old users inactive, while parameters of the persisting users, such as power or location, may vary. Now, most of the available multiuser detection (MUD) theory is based on the assumption that the number of active users is constant, known at the receiver, and equal to the maximum number of users entitled to access the system [1]. If this assumption is not verified, the receiver may exhibit a serious performance loss [2], [3]. In [4], the more realistic scenario in which the number of active users is unknown a priori, and varies with time with known statistics, is the basis of a new approach to detector design. The present work follows in the footsteps of [4], and is devoted to large-system analysis of this new type of detectors for Code Division Multiple Access (CDMA). Our main goal is to determine the performance loss caused by the need for estimating the identities of active users, which are not known a priori. To prevent a loss of optimality, we assume that identities and data are estimated jointly, rather than in two separate steps. Our interest is in randomly spread CDMA system in terms of multiuser and spectral efficiency, whose natural dimensions (number of users K, and spreading gain N) tend to infinity, while their ratio (the "system load") is kept fixed. In particular, we consider the optimal maximum a posteriori (MAP) multiuser detector, and use tools recently adopted from statistical physics [5], [6], [9], [7], [8]. Of special relevance in our analysis is the decoupling principle introduced in [6] for randomly spread CDMA. The general results derived from asymptotic analysis are validated by simulations run for a limited number of users [8].

In [10], the same analysis tools were applied to the static channel model described in [4], and used to derive asymptotic multiuser efficiency and bit error probability of a receiver performing joint estimation of user activity and data. The results of this report are an extension of those in [10], and focus on the degradation of multiuser efficiency and bit error probability when the uncertainty on the activity of the users grows. Asymptotic single-user capacity and spectral efficiency of the multiuser channel were derived in [11]. Here, we go one step beyond to provide an asymptotic analysis, in the large-system limit, of the effect of the activity factor on certain CDMA performance measures (minimum mean-square error, and multiuser efficiency).

This report is organized as follows. Section II introduces the system model and the main notations used throughout. Section III derives the large-system central fixed-point equation, and analytical bounds to the MMSE. Based on these results, Section IV discusses the interplay of maximum system load and multiuser efficiency. From the same asymptotic perspective, Section V analyzes the single-user capacity and spectral efficiency of the system. Finally, Section VI draws some concluding remarks. Proofs of some results can be found in the Appendix.

II. System Model

We consider static CDMA [4], and examine the optimum user-and-data detector. In particular, we study randomly spread direct-sequence (DS) CDMA with a maximum of K active users, where the received signal at time t is

$$\boldsymbol{y}_t = \boldsymbol{S}\boldsymbol{A}\,\boldsymbol{b}_t + \boldsymbol{z}_t \tag{1}$$

where $\boldsymbol{y}_t \in \mathbb{R}^N$, with N the length of the spreading sequences, $\boldsymbol{S} \in \mathbb{R}^{N\times K}$ is the matrix of the sequences, $\boldsymbol{A} = \mathrm{diag}(a_1, \ldots, a_K) \in \mathbb{R}^{K\times K}$ is the diagonal matrix of the users' signal amplitudes, $\boldsymbol{b}_t = (b_t^1, \ldots, b_t^K) \in \mathbb{R}^K$ is the users' data vector, and $\boldsymbol{z}_t$ is an additive white Gaussian noise vector with i.i.d. entries $\sim \mathcal{N}(0, \frac{1}{2})$. We assume that $a_k = \sqrt{\gamma}$, where γ is the average received signal-to-noise ratio (SNR)[1]. We define the system's activity rate as $\alpha \triangleq \Pr\{\text{user } k \text{ is active}\}$, $1 \le k \le K$. Active users employ binary phase-shift keying (BPSK) with equal probabilities. This scheme is equivalent to one where each user transmits

[1]The analysis presented in this report can be easily extended to different statistics of the a_k coefficients, like for example those induced by Rayleigh fading.

a ternary constellation $\mathcal{X} \triangleq \{-1, 0, +1\}$ with probabilities $\Pr\{b_t^k = -1\} = \Pr\{b_t^k = +1\} = \frac{\alpha}{2}$ and $\Pr\{b_t^k = 0\} = 1 - \alpha$. We define the maximum system load as $\beta \triangleq \frac{K}{N}$.

In a static channel model, the detector operation remains invariant along a data frame, indexed by t, but we often omit this time index for the sake of simplicity. Assuming that the receiver knows $\boldsymbol{S}$ and $\boldsymbol{A}$, the a posteriori probability (APP) of the transmitted data has the form

$$p(\boldsymbol{b}|\boldsymbol{y}, \boldsymbol{S}, \boldsymbol{A}) = \frac{1}{\sqrt{\pi}} e^{-\|\boldsymbol{y} - \boldsymbol{S}\boldsymbol{A}\boldsymbol{b}\|^2} \frac{p(\boldsymbol{b})}{p(\boldsymbol{y}|\boldsymbol{S}, \boldsymbol{A})}. \tag{2}$$

Hence, the maximum a posteriori (MAP) joint activity-and-data multiuser detector solves

$$\hat{\boldsymbol{b}} = \arg \max_{\boldsymbol{b} \in \mathcal{X}^K} p(\boldsymbol{b}|\boldsymbol{y}, \boldsymbol{S}, \boldsymbol{A}) \tag{3}$$

Similarly, optimum detection of single-user data and activity is obtained by marginalizing over the undesired users as follows:

$$\hat{b}^k = \arg \max_{b^k} \sum_{\boldsymbol{b} \in \{\mathcal{X}^K \setminus b^k\}} p(\boldsymbol{b}|\boldsymbol{y}, \boldsymbol{S}, \boldsymbol{A}). \tag{4}$$

A. The random-set perspective

The assumption of an unknown number of users in multiuser detection suggests Random-Set Theory (RST) as a tool for system analysis. From this perspective, RST, as used in [12], [4], yields a flexible and mathematically consistent framework for the problem under study. In fact, it naturally allows consideration of randomness not only in the data detected but also in the set of active users. This new approach leads to novel multiuser detectors (RST detectors), which improve on the bit error probability (BEP) performance with respect to detectors based on the assumption that K users are always simultaneously active [4].

A *random set* is a map from a sample space Ω to a family of subsets of a given space S. In the model we present, the space S is formed by the unknown data and parameters of the active interferers. Thus, the information carried by the users at time t can be modeled as a random set:

$$\underline{\boldsymbol{X}}_t = \{\boldsymbol{x}_t^{(1)}, \ldots, \boldsymbol{x}_t^{(k)}\} \tag{5}$$

whose elements (the parameters) are random vectors, and k (the number of active interferers) is itself a random integer taking values in the interval $[0, K-1]$. If everything about the interferers is known except their number and identity, then all random sets $\underline{\boldsymbol{X}}_t$ belong to the power set $2^{\mathbb{K}} \triangleq 2^{\{1,\ldots,K\}}$. Additionally, if we want to detect the (binary) data that the interferers transmit, $\underline{\boldsymbol{X}}_t$ takes values in a set with 3^K elements, denoted $3^{\mathbb{K}}$. Our objective here is to detect the number and identity of active interferers as well as the data they carry, assuming that the remaining parameters do not change during the tracking phase. Thus, time subscripts can be omitted, and the probability of interferer set $\underline{\boldsymbol{X}}$ depends only on its cardinality $|\underline{\boldsymbol{X}}|$. If we assume the transmission scheme introduced before, the prior random set distribution in this detection scenario turns out to be [4]

$$f_{\underline{\boldsymbol{X}}}(\boldsymbol{B}) = 2^{-\ell|\boldsymbol{B}|} \alpha^{|\boldsymbol{B}|} (1 - \alpha)^{K - |\boldsymbol{B}|} \tag{6}$$

where $\boldsymbol{B}$ is a realization of the random set $\underline{\boldsymbol{X}}$.

B. Statistical-physics approach to large-system analysis of classic MUD

Of late, the theory of multiuser detection has received important contributions stemming from a large-system approach. Some of them are due to Tanaka [5], who used statistical physics concepts and methodologies in multiuser detection to obtain large-system uncoded minimum BER and spectral efficiency with equal-power binary inputs. Tanaka's pioneering approach inspired additional work. Reference [14] studies the channel capacity, while [6] presents a unified treatment of Gaussian CDMA channels and multiuser detection in the large-system limit with arbitrary input distribution and flat fading.

The core of the application of statistical physics to communication problems lies in the concept of *free energy*. In statistical physics, the free energy $F(\boldsymbol{X})$ (where $\boldsymbol{X}$ is the state variable) relates the energy $\varepsilon(\boldsymbol{X})$ and the entropy $H(\boldsymbol{X})$ of a physical system of particles in the following way:

$$F(\boldsymbol{X}) = \varepsilon(\boldsymbol{X}) - TH(\boldsymbol{X}) \tag{7}$$

where T is the temperature of the system. At thermal equilibrium the free energy (7) is minimized, since the entropy is maximized, as time evolves, following the second law of thermodynamics. Under these conditions, the free energy can also be expressed as

$$F(\boldsymbol{X}) = -T \ln Z(\boldsymbol{X}) \tag{8}$$

where $Z(\boldsymbol{X}) = \sum_{\boldsymbol{x}} \exp\left(-\frac{\varepsilon(\boldsymbol{x})}{T}\right)$ is called the *partition function*, $\boldsymbol{x}$ denotes each microscopic value, $\varepsilon(.)$ is the energy operator, and the temperature $T \geq 0$ reflects the energy constraint. A key point here is that the free energy, as normalized to the dimension of the system, is a *self-averaging* quantity, which means that it tends to its expectation as the system dimension grows to infinity. The free energy is the starting point for calculating macroscopic properties of a thermodynamic system.

In communications systems, the role of the energy function can be played by the metric of a detector in a multiuser channel. Therefore, by assuming for the system under study the self-averaging property, any detector parametrized by a certain metric can be analyzed with the tools of statistical mechanics to derive results in the large-system limit. Hence, the calculation of the limiting free energy at equilibrium, i.e., the logarithm of the inverse of the partition function (8), leads to the characterization of the large-system detector performance in terms of its *macroscopic* parameters such as multiuser and spectral efficiency.

In a communications scheme such as the one modeled by (2), the goal of the multiuser detector is to infer the information-bearing symbols given the received signal $\boldsymbol{y}$ and the knowledge about the channel state. This leads naturally to the choice of the partition function $Z(\boldsymbol{y}, \boldsymbol{S}) = p(\boldsymbol{y} \mid \boldsymbol{S})$. The corresponding free energy, normalized by the number of users and with the choice $T = 1$, becomes

$$\mathcal{F}_K \triangleq -\frac{1}{K} \ln p(\boldsymbol{y} \mid \boldsymbol{S}) \tag{9}$$

To calculate this expression we make the self-averaging assumption, which states that the randomness of (9) vanishes as $K \to \infty$. This is tantamount to saying that the free energy per user $\mathcal{F}_K$ converges in probability to its expected value over the distribution of the random variables $\boldsymbol{y}$ and $\boldsymbol{S}$, denoted by

$$\mathcal{F} \triangleq \lim_{K\to\infty} \mathbb{E}\left\{-\frac{1}{K} \ln p(\boldsymbol{y} \mid \boldsymbol{S})\right\}. \tag{10}$$

Evaluation of (10) is made possible by the *replica method*[7], [9], which consists of introducing n independent replicas of the input variables, with corresponding density $p^n(\boldsymbol{y}|\boldsymbol{S})$, and computing $\mathcal{F}$ as follows:

$$\mathcal{F} = -\lim_{n\to 0} \frac{\partial}{\partial n} \left(\lim_{K\to\infty} \frac{1}{K} \ln \mathbb{E}\{p^n(\boldsymbol{y} \mid \boldsymbol{S})\}\right) \tag{11}$$

The equivalence between (10) and (11) uses the following equality:

$$\lim_{n\to 0} \frac{\partial}{\partial n} \ln \mathbb{E}\{Z^n\} = \mathbb{E}\{\ln Z\} \tag{12}$$

under the assumption that $Z^n = p^n(\boldsymbol{y} \mid \boldsymbol{S})$ is valid for all real valued n in the vicinity of $n = 0$, and that the limits in K and n can be interchanged.

C. The statistical-physics framework for MUD with an unknown number of users

Our large-system analysis is based on the assumption that the performance measures are self-averaging as K grows to infinity. Now, in our multiuser system, the number of active users is not known at the receiver, and activity is jointly detected with the data. Thus, previous results cannot be directly applied, because K, a random variable, cannot be used as the normalizing constant in the free energy function. A simple and sensible way to circumvent this problem consists of assuming a *ternary* constellation $\{\pm 1, 0\}$, where 1 and -1 account for BPSK transmission, whereas symbol 0 represents inactivity. This method allows us to keep (9) invariant.

An alternative solution might be the reformulation of the free energy approach by defining a new random variable, X_k, denoting the number of users. Such a formulation would imply the following:

1) Since the normalization quantity is a random variable, the expectation in the free energy must also be performed over the distribution of X_k (binomial, according to (6)).
2) The limit in the free energy must be performed over the mean of X_k, i.e., αK.
3) If we consider K to be the effective number of users, the detection only refers to data, since the active users are only transmitting BPSK.
4) Since K is placed outside the logarithm, a modified partition function should be defined as

$$\hat{Z}^n = p_{\boldsymbol{Y}|\boldsymbol{S}}^{n/K}(\boldsymbol{y}, \boldsymbol{S})$$

to verify (12).

Regardless of last consideration, which might lead to a more complex analysis, the most crucial point is that user identification is part of the estimation in our problem, and therefore, the alternative analysis can be used when only the estimation of the parameters of active users is of concern.

In conclusion, to the purpose of activity estimation in the large-system limit, we choose the approach based on the equivalent ternary representation of the random set distribution (6), with K the number of potential users accessing the multiuser channel.

D. The decoupling principle

To compute the free energy through replica analysis, one of the cornerstones in large deviation theorem, the Varadhan's theorem [15], is invoked to transform the calculation of the limiting free energy into a simplified optimization problem, whose solution is assumed to exhibit symmetry among its replicas. More specifically, in the case of a MAP individually optimum detector, the optimization yields a fixed-point equation, whose unknown is a single operational macroscopic parameter, which is claimed to be the multiuser efficiency (which reflects the degradation factor of SNR due to interference) of an equivalent Gaussian channel [6]. Due to the structure of the optimization problem, the multiuser efficiency must minimize the free energy. The above is tantamount to formulating the *decoupling principle*:

Claim 2.1: [6], [8] Given a multiuser channel, the distribution of the output $\hat{b}^k$ of the individually optimum (IO) detector, conditioned on $b^k = b$ being transmitted with amplitude a, converges to the distribution of the posterior mean estimate of the single-user Gaussian channel

$$y = \sqrt{\gamma} b^k + \frac{1}{\sqrt{\eta}} z \tag{13}$$

conditioned on $b^k = b$ being transmitted, where $z \sim \mathcal{N}(0, 1)$, and η, the multiuser efficiency, is the solution of the following fixed-point equation:

$$\eta^{-1} = 1 + \beta \mathbb{E}_\gamma \left[\gamma \mathrm{MMSE}\left(\gamma, \eta, \alpha\right) \right]. \tag{14}$$

If (14) admits more than one solution, we must choose the one minimizing the free energy function

$$\mathcal{F} = -\mathbb{E}\left[\int p(y \mid b^k) \ln p(y \mid b^k) \mathrm{d}y \right] - \frac{1}{2} \ln \frac{2\pi e}{\eta} + \frac{1}{2\beta} \left(\eta \ln e + \ln \frac{2\pi e}{\eta} \right). \tag{15}$$

In (14)-(15), $p(y \mid b^k)$ is the transition probability of the large-system equivalent single-user Gaussian channel described by (13), and

$$\text{MMSE}(\gamma, \eta, \alpha) \triangleq \mathbb{E}\left[\left(b^k - \hat{b}^k\right)^2\right] \tag{16}$$

denotes the minimum mean-square error in estimating b^k in Gaussian noise with amplitude equal to $\sqrt{\gamma}$, where $\hat{b}^k = \mathbb{E}\left[b^k|y\right]$ is the posterior mean estimate, known to minimize the MMSE [16].

E. A note on the validity of the replica method

We hasten to observe that, in spite of its approximating very well experimental data and being consistent with previous theoretical work [13], [17], [18], replica method analysis relies on some unproved assumptions: the self-averaging property of the free energy, the replica symmetry of the fixed-point solution, and the exchange of order of limits and analytic continuation of (12) to real n. Although the validation of the mathematical rigor of these assumptions is still an unsolved problem, there has been some recent progress in this direction [19], [20], [21].

III. Large-system multiuser efficiency

We illustrate here the behavior of multiuser efficiency and system load in the high-SNR region corresponding to detection with an unknown number of users. We start by shaping our problem into the statistical-physics framework [5], [6]. As mentioned earlier, the multiuser detector metric is regarded as the energy of a system of particles at state $\boldsymbol{X}$. Therefore, the role of the partition function $Z(\boldsymbol{X}) = \sum_{\boldsymbol{x}} \exp\left(-\varepsilon(\boldsymbol{x})/T\right)$ is played by the output density given the channel information, i.e., $p(\boldsymbol{y}|\boldsymbol{S}) = \pi^{-1/2} \sum_{\boldsymbol{b}} p(\boldsymbol{b}) \exp($

The energy operator $\varepsilon(.)$, as derived from the free energy, is related to the logarithm of the joint distribution $p(\boldsymbol{y} \mid \boldsymbol{b}, \boldsymbol{S})p(\boldsymbol{b})$:

$$\varepsilon(\boldsymbol{b}) = \|\boldsymbol{y} - \boldsymbol{S}\boldsymbol{A}\boldsymbol{b}\|^2 - \ln p(\boldsymbol{b}) \tag{17}$$

We can now invoke the decoupling principle (Claim 2.1) in the multiuser system (1), so as to use its single-user characterization. By doing this, the system is converted into the scalar Gaussian channel (13), where the input distribution for an arbitrary user k takes values $\mathcal{X} = \{-1, 0, +1\}$ with probabilities $\frac{\alpha}{2}, 1-\alpha$ and $\frac{\alpha}{2}$, respectively, the signal amplitudes from matrix $\boldsymbol{A}$ are $a = \sqrt{\gamma}$, where γ is the SNR per active user (we shall call it SNR), and the inverse noise variance is equal to the multiuser efficiency η. Hence, η is the solution of the fixed-point equation (14) that minimizes (15), where the MMSE is given by (16).

Our main results regarding joint multiuser identification and detection are stated as follows:

Proposition 3.1: Given a randomly spread DS-CDMA system with constant equal power per user, the large-system multiuser efficiency of an individually optimum detector that performs MAP estimation of users' identities and their data under BPSK transmission is the solution of the following fixed-point equation

$$\eta = \frac{1}{1 + \beta\left(\gamma\left[\alpha - \int \frac{1}{\sqrt{2\pi}} e^{\frac{-y^2}{2}} \frac{\alpha^2 \sinh(\eta\gamma - y\sqrt{\eta\gamma})}{\alpha\cosh(\eta\gamma - y\sqrt{\eta\gamma}) + (1-\alpha)e^{\eta\frac{\gamma}{2}}} \mathrm{d}y\right]\right)} \tag{18}$$

that minimizes the free energy (15).

Proof: See Appendix I. ∎

Our approach differs from that in [6], [14], [5], as the fixed-point equation (18) includes also for the prior distribution resulting from (6), which accounts for the users' activity in a static channel. Under MAP estimation, detection requires the knowledge not only of the prior information of the data, but also of the activity rate α. Thus, the fixed-point equations depend on MMSE, SNR, and system load. Numerical solutions vs. SNR at a load $\beta = 3/7$ are shown in Fig. 1. Plots like this one illustrate how the multiuser efficiency is affected by the level of noise and interference, and by the uncertainty in the users' activity

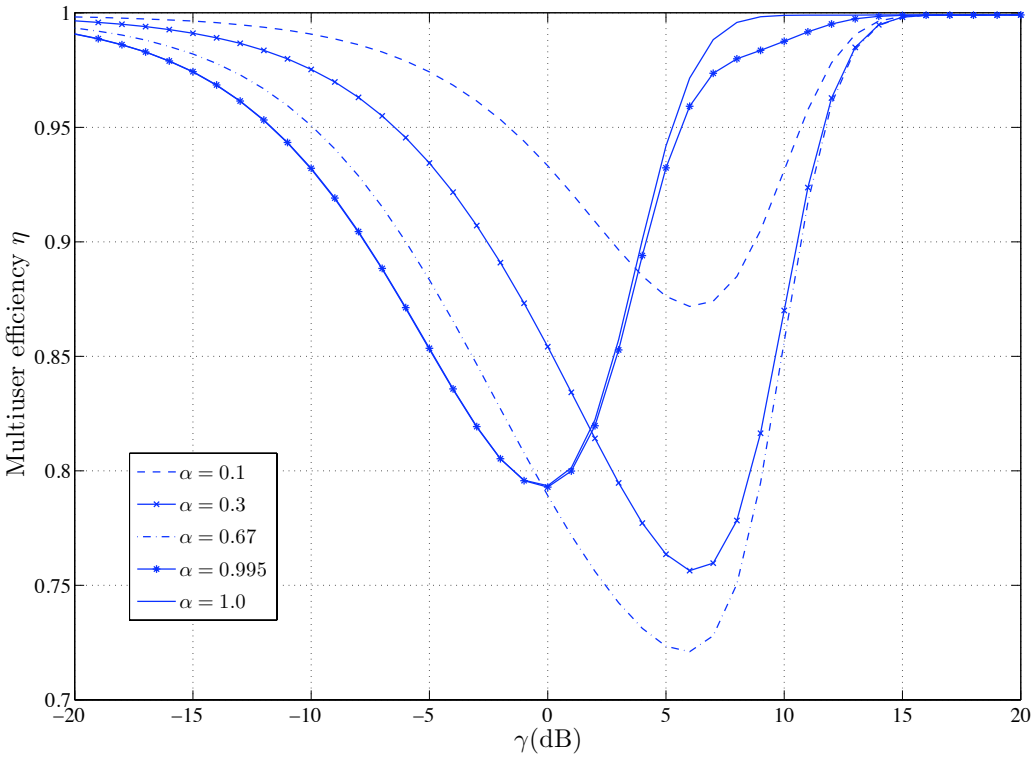

Fig. 1. Large-system multiuser efficiency of the user-and-data detector under MAP with prior knowledge of α and β=3/7.

rate. For low SNR, noise dominates, and the performance of the MMSE and the multiuser efficiency is degraded as α grows, since the presence of more active users adds more noise to the system. On the other hand, as we shall discuss later, for high SNR the MMSE strongly depends on the minimum distance between the transmitted symbols, and the activity rate here plays a secondary role. Hence, the gap between the multiuser efficiencies with $\alpha = 1$ and $\alpha \neq 1$ for larger SNR is due to the fact that the former constellation has twice the minimum distance of the latter. We can observe clearly the transition behavior from low to high SNR for values of α approaching 1. Moreover, when $\alpha = 1$, (18) reduces to the fixed-point equation for the classic assumption, in which all users are active and transmit a binary antipodal constellation [6]:

$$\eta^{-1} = 1 + \beta \left(\gamma \left[1 - \int \frac{1}{\sqrt{2\pi}} e^{-y^2/2} \tanh(\eta\gamma - y\sqrt{\eta\gamma}) \mathrm{d}y \right] \right) \tag{19}$$

In this case, it can be shown that, for high SNR, we have $\mathrm{MMSE}(\gamma, \eta, \alpha = 1) \approx \sqrt{\frac{2\pi}{\eta\gamma}} e^{-\eta\gamma/2}$. In fact, the following general result holds:

Lemma 3.2: [22] For large output SNR, the MMSE of a system transmitting an equiprobable M-ary normalized constellation with minimum Euclidean distance d in a Gaussian channel with noise variance $1/\eta$ is:

$$\mathrm{MMSE}(\gamma, \eta, \alpha = 1) = \kappa(\eta, \gamma) e^{-d^2\eta\gamma/8} \tag{20}$$

with $\kappa_1(\eta, \gamma) \leq \kappa(\eta, \gamma) \leq \kappa_2$, where $\kappa_1(\eta, \gamma) = \mathcal{O}(1/\sqrt{\eta\gamma})$ and κ_2 is a constant, given by the maximum distance between neighboring symbols.

For the whole range of activity rates, i.e., $\alpha \in [0, 1]$, we can derive lower and upper bounds illustrating analytically the transition between the classical assumption ($\alpha = 1$) and the introduction of activity

detection ($\alpha < 1$) for large SNR. Our calculations bring about a new analytical framework to deal with large-system analysis, as we will see in next section. Our bounds are consistent with lemma 3.2 and the lower bound includes the case $\alpha = 1$. The general result is stated as follows:

Theorem 3.3: The MMSE of joint user identification an data detection in a large system with an unknown number of users has the following behavior, valid for large values of the product $\eta\gamma$:

$$\kappa_1(\gamma, \eta, \alpha) \leq \text{MMSE}(\gamma, \eta, \alpha) \leq \kappa_2(\gamma, \eta, \alpha) \tag{21}$$

where

$$\kappa_1(\gamma, \eta, \alpha) = 2\sqrt{\frac{\alpha(1-\alpha)}{\pi\eta\gamma}}e^{-\eta\gamma/8} \tag{22}$$

and

$$\kappa_2(\gamma, \eta, \alpha) = 2\alpha e^{-\eta\gamma/2} + \sqrt{\frac{\pi\alpha(1-\alpha)}{\eta\gamma}}e^{-\eta\gamma/8} \tag{23}$$

Proof: See Appendix II. ■

Bounds (22) and (23) describe explicitly, in the high-SNR region, the relationship between the MMSE, the users' activity rate, and the SNR. In Fig. 2 these bounds are compared to the true MMSE values as a function of γ for fixed α. It can be seen that the uncertainty on the users' activity modifies substantially the exponential decay of the MMSE for high SNR. In fact, a value of α different from 1 causes the MMSE to decay as $\exp(-\eta\gamma/8)$, rather than as $\exp(-\eta\gamma/2)$, which would be the case when all users are active. Furthermore, we can observe that, for sufficiently large SNR, the behavior vs. α of the optimal detector is symmetric with respect to $\alpha = 1/2$, which corresponds to the maximum uncertainty on the activity rate. As plotted in Fig. 3, for large SNR, the MMSE essentially depends on the minimum distance between the inactivity symbol $\{0\}$ and the data symbols $\{-1, 1\}$, and thus users' identification prevails over data detection. Summarizing, the dependence of the MMSE must be symmetrical with respect to $\alpha = 1/2$, since it reflects the impact of prior knowledge on the user's activity into the estimation.

IV. System Load and Related Considerations

Recall the definition of maximum system load $\beta = \frac{K}{N}$, where K is the maximum number of users accessing the multiuser channel. When the number of active users is unknown, and there is a priori knowledge of the activity rate, the actual system load is $\beta' = \alpha\beta$. In this section, we focus on β and study some of its properties. Notice that, given an activity rate, results on the actual system load follow trivially.

A. Solutions of the large-system fixed-point equation

We characterize the behavior of the system load subject to quality-of-service constraints. Doing this helps shedding light into the nature of the solutions of the fixed-point equation (18). In particular, there might be cases where (18) has multiple solutions. These solutions correspond to those appearing in any simple mathematical model of magnetism based on the evaluation of the free energy with the fixed-point method [7]. They represent what in the statistical physics parlance is called *phase coexistence* (for example, this occurs in ice or liquid phase of water at 0^{o}C). In particular, at low temperature, the magnetic system might have 3 solutions $0 \leq \Psi_1 < \Psi_2 < \Psi_3 \leq 1$. Solutions Ψ_1 and Ψ_3 are stable: one of them is globally stable (it actually minimizes the free energy), whereas the other is metastable, and a local minimum. Solution Ψ_2 is always unstable, since it is is a local maximum. The "true" solution is therefore given by Ψ_1 and Ψ_3, for which the free energy is a minimum. The same consideration applies also to our multiuser detection problem where multiuser efficiencies for the IO detector might vary significantly depending on the value of the system load and SNR. More specifically, for large enough SNR, stable solutions may switch between a region that approaches the single-user performance ($\eta = 1 - \epsilon_1$) and a region approaching the worst performance ($\eta = \epsilon_0$), for $0 < \epsilon_1, \epsilon_0 \ll 1$. Following previous literature [5],

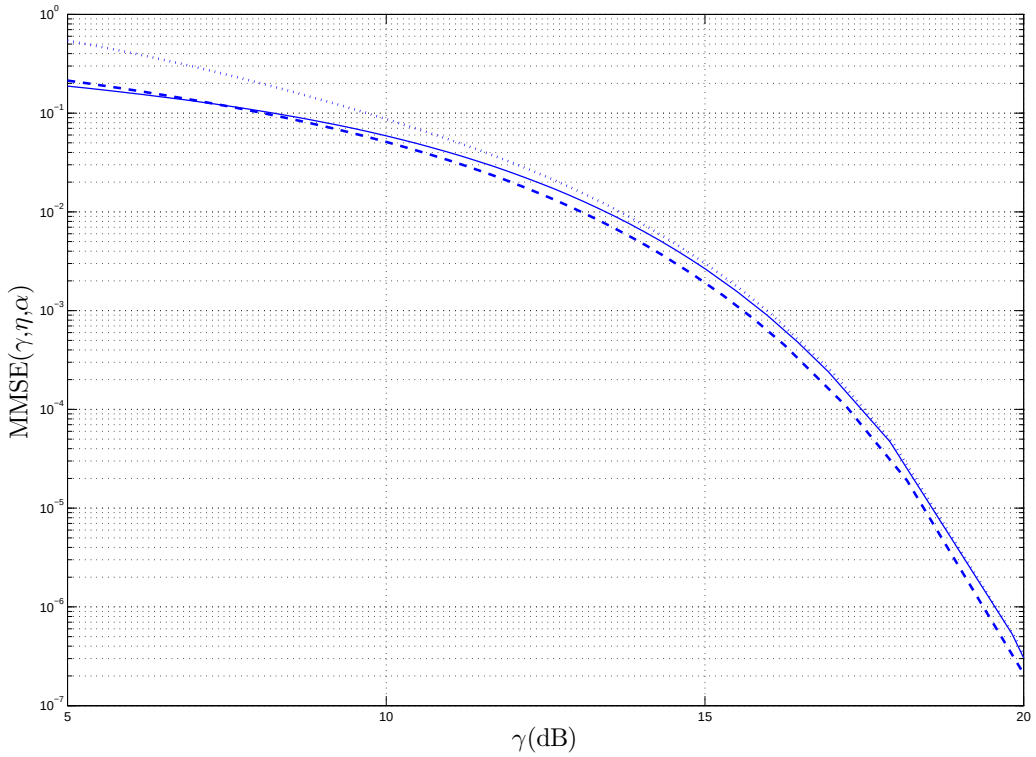

Fig. 2. Comparing exact MMSE with its upper and lower bounds for $\alpha = 0.5$.

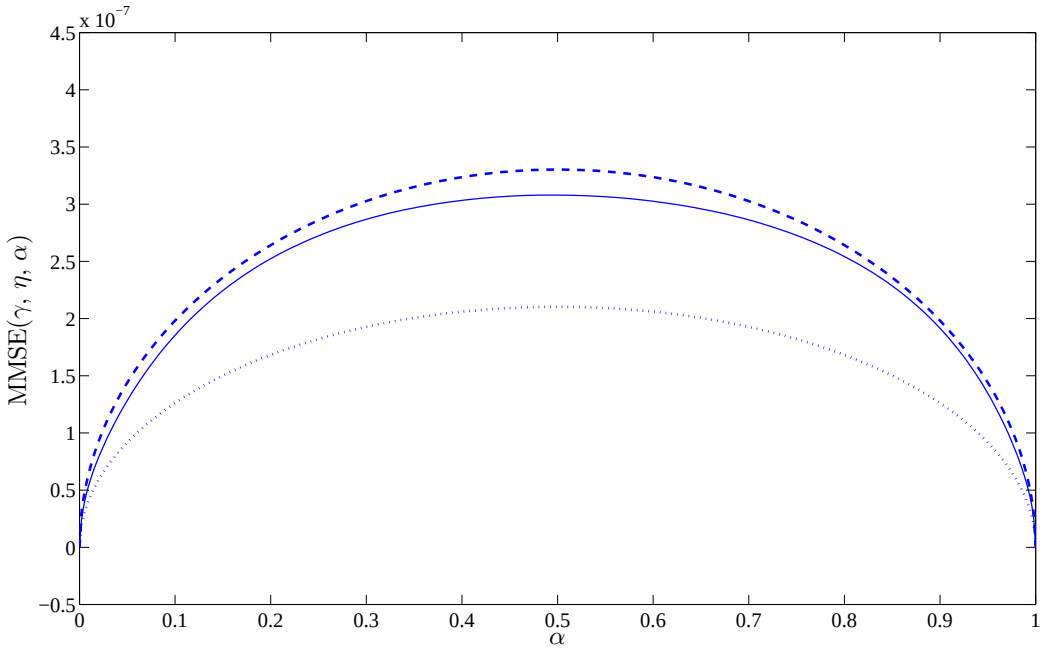

Fig. 3. Comparing exact MMSE with its upper and lower bounds for SNR$=$ 20 dB and $\alpha \in [0, 1]$.

we shall call the former solutions *good* and the latter *bad*. When the solution is unique, due to low or high system load, the multiuser efficiency is a globally stable solution that lies in either the good or the bad solution region. Then, for given system parameters, the set of *operational* (or globally stable) solutions is formed by solutions that lie on any of these sets and minimize the free energy.

It is interesting to discuss the relevance of good and bad solutions in our problem. From a computational perspective, we are particularly interested in single solutions, either bad or good, that surely avoid metastability and instability. These solutions belong to a specific subregion within the bad and good regions, and appear for low and high SNR, respectively. In an information-theoretic perspective, it might seem that the true solutions should capture all our interest. However, it has been shown that metastable solutions appear in suboptimal belief-propagation-based multiuser detectors, where the system is easily attracted into the bad solutions region (that corresponding to low multiuser efficiency), due to initial configurations far from the true solution [8]. Moreover, the region of good solutions has interest in the high-SNR analysis, because, for a given system load, it can be observed that the multiuser efficiency tends to 1, consistently with previous theoretical results [17]. In what follows, we provide an analysis of the boundaries of the stable solutions regions, as well as their computationally feasible subregions with practical interest in the low and the high SNR regime.

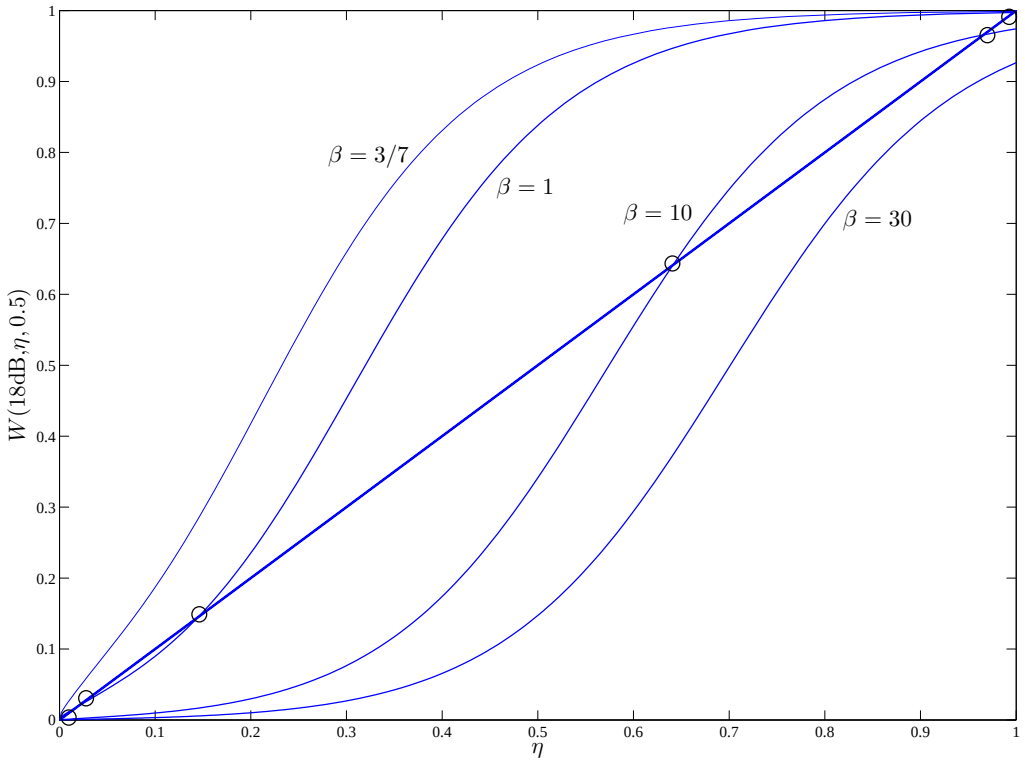

Fig. 4. Fixed-point solutions (marked by circles) for different values of β and fixed $\alpha = 0.5$, and $\gamma = 18$dB.

We can have a quantitative illustration of the above considerations by plotting the left- and right-hand sides of (18) to obtain fixed points for constant values of amplitude and activity rate, and as a function of the system load. Solutions of (18) are found at the intersection of the curve corresponding to the right-hand side with the $y = \eta$ line. Fig. 4 plots different solutions of the right-hand side of (18) for

increasing system load, $\alpha = 0.5$ and $\gamma = 18$ dB:

$$W(\gamma, \eta, \alpha) \triangleq \frac{1}{1 + \beta\gamma \text{MMSE}(\gamma, \eta, \alpha)}$$

Notice first that the structure of the fixed-point equation does not allow in general the solution $\eta = 0$, and for finite γ and β, $\eta = 1$ is also avoided. In fact, the latter is an asymptotic solution for large SNR and certain system loads, as the MMSE decays exponentially to 0. From Fig. 4, one can observe phase transitions and the coexistence of multiple solutions. In particular, we observe that for $\beta = 3/7$ the good solution is computationally feasible. On the other hand, for $\beta = 1$ and $\beta = 10$ the system has 3 solutions, where the true solution belongs to either the bad or the good solution region. When the system load achieves $\beta = 30$, the curve only intersects the identity near 0, and the operational solution is unique and lies in a subregion of the bad solutions.

B. System load and the space of fixed-point solutions

Even in the case of good solutions, the multiuser efficiency can be greatly degraded by the joint effect of the activity rate and the system load. In order to analyze the fixed-point equation (18) from a different perspective and shed light into the interplay between these parameters, we express the maximum system load as the following function, derived from (18):

$$\Upsilon_\beta(\gamma, \eta, \alpha) \triangleq \frac{(1 - \eta)}{\eta\gamma \text{MMSE}(\gamma, \eta, \alpha)} \tag{24}$$

Since MMSE is a continuous function for η, then $\Upsilon_\beta(., \eta, .)$ is a continuous function in any compact set over the domain $(0, 1]$ for given SNR and activity rate. It is also easy to observe that, for small values of η, $\Upsilon_\beta(\gamma, ., \alpha)$ tends to infinity, whereas in the high-η region, which is of interest here, it decays to 0. Before analyzing the behavior of (24), we introduce a few definitions that help describe the boundaries between the regions with and without coexistence (in the statistical-physics literature, these boundaries are called *spinodal lines* [5]). We also define appropriately the regions of potentially stable solutions as introduced before.

Definition 4.1: The *critical system load* $\beta^\star(\gamma, \alpha)$ is the maximum load up to which a stable good solution of (18) exists.

Definition 4.2: The *transition system load* $\beta_\star(\gamma, \alpha)$ is the load at which the true solution of (18), $\eta_\star$ starts to coexist with other solutions $\eta'_\star$.

Definition 4.3: The *good solution region* corresponds to the domain of (24) formed by the maximum η in every set of pre-images of Υ_β below the critical system load:

$$\mathcal{R}_g = \left\{\eta \in [0, 1], \eta = \max\{\Upsilon_\beta^{-1}(\beta)\}, \forall \beta \in [0, \beta^\star]\right\} \tag{25}$$

Similarly, the *bad solution region* corresponds to the domain of (24) formed by the minimum η in every set of pre-images of Υ_β above the transition system load:

$$\mathcal{R}_b = \left\{\eta \in [0, 1], \eta = \min\{\Upsilon_\beta^{-1}(\beta)\}, \forall \beta \in [\beta_\star, +\infty]\right\} \tag{26}$$

Fig. 5 illustrates Υ_β (for fixed SNR and activity rate) and shows the regions defined by the aforementioned parameters. It is important to remark that both system loads defined above delimit the regions from which there is phase coexistence ($\beta_\star \leq \beta \leq \beta^\star$) from the areas where there is one solution ($\beta > \beta^\star$ or $\beta < \beta_\star$). Additionally, Fig. 5 illustrates the set of solutions that satisfy conditions (25), (26).

As illustrated by the plots, it seems useful to define analytically the domain where stable solutions can be found. Before doing this, we differentiate for convenience the case with unknown number of users $\alpha \in (0, 1)$, from the case where all users are active ($\alpha = 1$). We do not consider the case $\alpha = 0$.

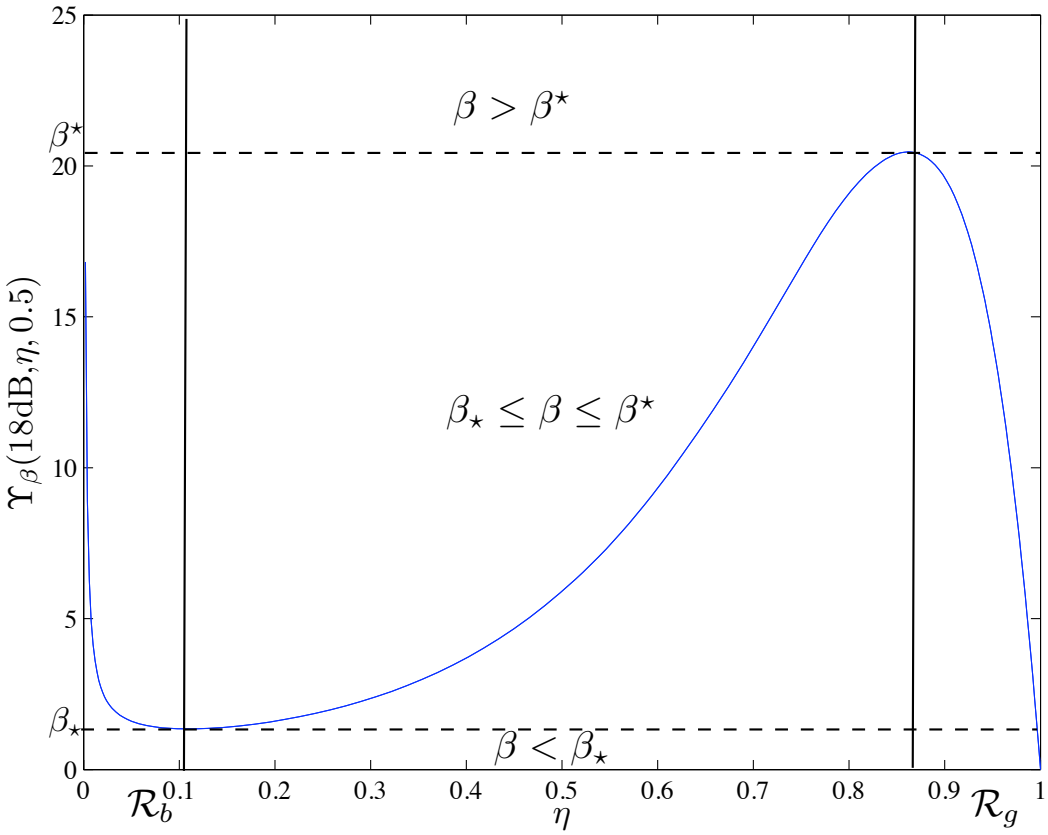

Fig. 5. System load function in the multiuser efficiency domain for $\alpha = 0.5$ and $\gamma = 18$dB.

1) Case $\alpha \in (0,1)$*:* In order to analyze the conditions on the system load, SNR, and activity rate, under which we can find a good solution, we use the asymptotic results on the MMSE (22) and (23), yielding lower and upper bounds $L(.,.,.) \leq \Upsilon_\beta(\gamma,\eta,\alpha) \leq U(.,.,.)$, where

$$L(\gamma,\eta,\alpha) \triangleq \frac{(1-\eta)}{\sqrt{\pi\eta\gamma\alpha(1-\alpha)}} e^{\eta\gamma/8}$$
$$U(\gamma,\eta,\alpha) \triangleq \frac{(1-\eta)}{2}\sqrt{\frac{\pi}{\eta\gamma\alpha(1-\alpha)}} e^{\eta\gamma/8}$$

Although not exact for low SNR, the dependence on η of the upper and lower bound provides a good approximation for the dependence of $\Upsilon_\beta(\gamma,.,\alpha)$ for large SNR and given α. Hence, by using $U(.,.,.)$ and $L(.,.,.)$, we obtain necessary and sufficient conditions that determine the regions of stable solutions and provide analytical expressions for the transition and critical system loads. The main result for $\alpha \in (0,1)$ follows:

Theorem 4.4: Given the range of activity rates $\alpha \in (0,1)$, a necessary condition for phase coexistence is

$$\gamma \geq 4(3+2\sqrt{2}), \quad (\gamma \geq 13.67 \text{ dB}) \tag{27}$$

Moreover, for high SNR, the condition is met and the transition system load is bounded by

$$L(\gamma,\eta_m,\alpha) < \beta_\star(\gamma,\alpha) < U(\gamma,\eta_m,\alpha) \tag{28}$$

while the critical system load is bounded by

$$L(\gamma, \eta_M, \alpha) < \beta^\star(\gamma, \alpha) < U(\gamma, \eta_M, \alpha) \tag{29}$$

and η_m, η_M are given by

$$\begin{aligned} \eta_m &\triangleq (\gamma/2 - 2 - 4\Delta(\gamma))/\gamma \\ \eta_M &\triangleq (\gamma/2 - 2 + 4\Delta(\gamma))/\gamma \\ \Delta(\gamma) &= \sqrt{(\gamma/8)^2 - 3\gamma/8 + 1/4} \end{aligned}$$

Hence, the bad-solution region is given by $\mathcal{R}_b = (0, \eta_m]$, whereas the good-solution region is $\mathcal{R}_g = [\eta_M, 1)$. Similarly, the subregions of single bad solutions, that we shall denote $R_{bc} = (0, \eta_{bc}] \subset \mathcal{R}_b$, and of single good solutions, denoted by $R_{gc} = [\eta_{gc}, 1) \subset \mathcal{R}_g$, satisfy:

$$\begin{aligned} \eta_{bc} &= \min\{\Upsilon_\beta^{-1}(\beta^\star)\} > \max\{U^{-1}(\beta^\star)\} \\ \eta_{gc} &= \max\{\Upsilon_\beta^{-1}(\beta_\star)\} < \max\{L^{-1}(\beta_\star)\} \end{aligned}$$

Proof: See Appendix III. ■

The above result provides the general boundaries of the space of solutions of our problem. It is important to note that η_m and η_M are very good approximations for high SNR of the positions of the minimum and maximum observed in Fig. 5, which determine transition and critical system loads. Remark also that the activity rate affects the boundaries in the same symmetrical manner as it does the MMSE (i.e., the worst case here also corresponds to $\alpha = 0.5$) but has no impact on the operational region, that is only reduced in size by increasing SNR. Since the different solution domains are unaltered by α, the overall performance maintains the same behavior as the MMSE. In particular, the region of good and bad solutions is characterized, in the limit of high SNR, as follows:

Corollary 4.5: In the limit of high SNR, $\mathcal{R}_g \to \{1\}$, and $\mathcal{R}_b \to \{0\}$.

Proof: It results from

$$\begin{aligned} \lim_{\gamma\to\infty} \eta_M &= \lim_{\gamma\to\infty} \frac{(\gamma/2 - 2 + 4\sqrt{(\gamma/8)^2 - 3\gamma/8 + 1/4})}{\gamma} = 1 \\ \lim_{\gamma\to\infty} \eta_m &= \lim_{\gamma\to\infty} \frac{(\gamma/2 - 2 - 4\sqrt{(\gamma/8)^2 - 3\gamma/8 + 1/4})}{\gamma} = 0 \end{aligned}$$

■

Note that, given a system load β with $\beta^\star > \beta$, for sufficiently large SNR the unique true (large-system) solution is $\eta = 1$, which corroborates the main result in [17]. Moreover, the description of the feasible good solutions by analytical means allows the computation of a sufficient condition on the system load to guarantee a given multiuser efficiency in practical implementations. More specifically, we use the aforementioned lower bound on Υ to state that any system load below $L(.,.,.)$ guarantees that the requirement is met. The result is stated as follows:

Corollary 4.6: The maximum system load, $\beta_{\alpha,\eta}$, for a given activity rate and multiuser efficiency requirement, $\eta = 1 - \epsilon$ $(0 < \epsilon \ll 1)$ that lies in R_{gc}, is lower-bounded in the high-SNR region by:

$$\beta_{\alpha,\eta} > \frac{\epsilon}{\sqrt{\pi\eta\gamma\alpha(1-\alpha)}} e^{(1-\epsilon)\gamma/8} \tag{30}$$

In Fig. 6 we show the numerical values of the transition and the critical system load as function of the SNR in a (γ, β) space. We also use the asymptotic expansion to derive upper and lower bounds, respectively. The curves plotted are the spinodal lines, and mark the boundary between the regions with and without solution coexistence. The $\beta_\star$ (lower branch) separates the region where the bad solution disappears, whereas $\beta^\star$ (upper branch) contains the bifurcation points at which the operational solution disappears. The intersection point between both branches corresponds to the SNR threshold (27), which provides the necessary condition for solution coexistence.

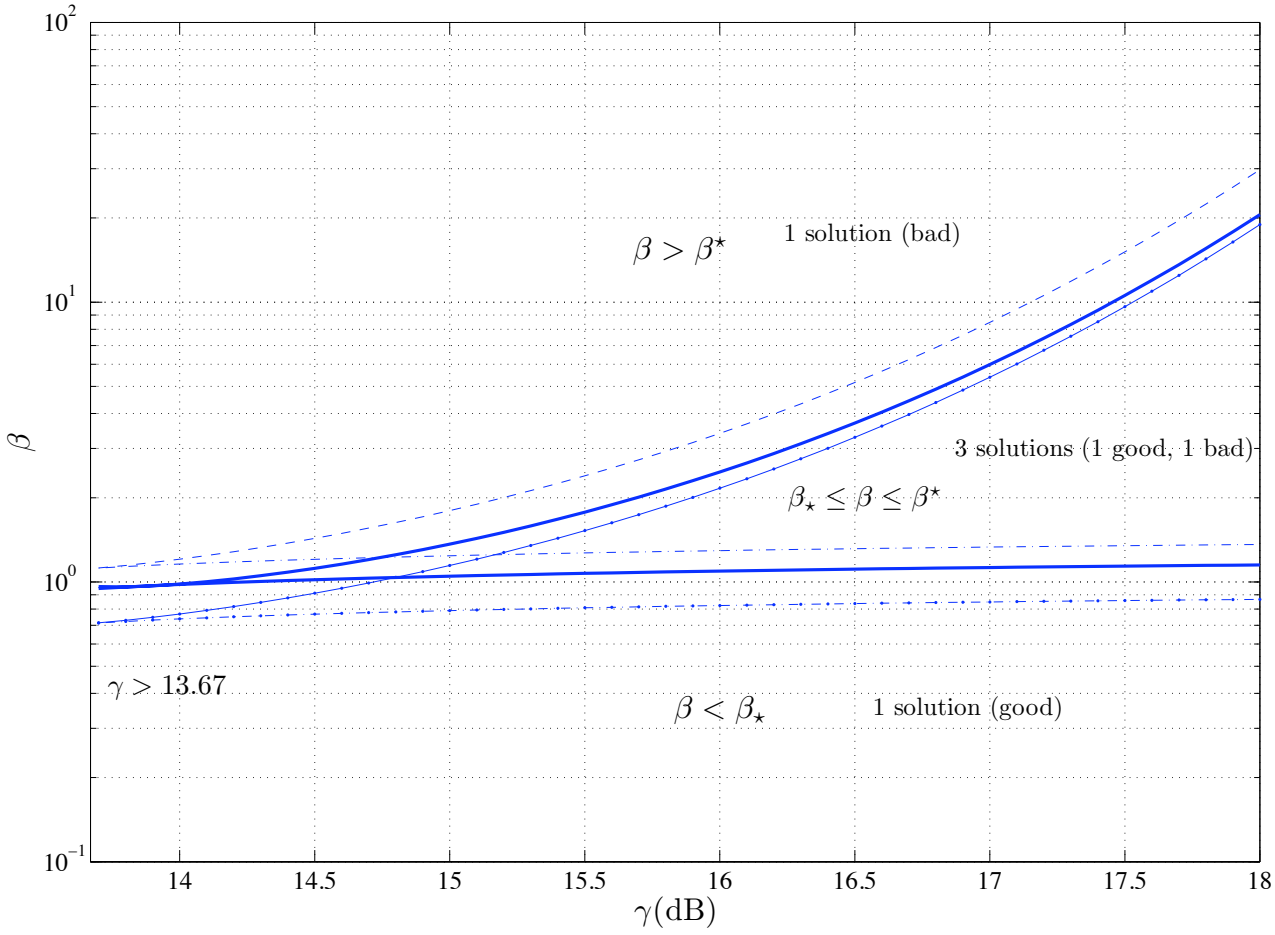

Fig. 6. Upper and lower bounds on the numerical spinodal lines (thicker line) for $\alpha = 0.5$.

2) Case $\alpha = 1.0$: We apply now the same reasoning to the "classical" approach to multiuser detection, corresponding to activity rate 1. In that case, using the approximation in [22], the system load function can be lower-bounded by:

$$\Upsilon_\beta(\gamma, \eta) = \frac{(1-\eta)}{\eta\gamma \text{MMSE}(\gamma\eta)} < \frac{(1-\eta)\, e^{\eta\gamma/2}}{\sqrt{\pi\eta\gamma}} \tag{31}$$

Hence, we can derive the following spinodal lines:

Corollary 4.7: Given $\alpha = 1$ a necessary condition for the phase coexistence is that:

$$\gamma \geq 3 + 2\sqrt{2}, \quad (\gamma \geq 7.65\text{dB})$$

Moreover, for high SNR, the condition is met and the transition system load is upper-bounded by

$$\beta_\star < \frac{(1-\eta_{m'})}{2}\sqrt{\frac{\pi}{\eta_{m'}\gamma}}e^{\frac{\eta_{m'}\gamma}{2}} \tag{32}$$

and the critical system load is upper-bounded by

$$\beta^\star < \frac{(1-\eta_{M'})}{2}\sqrt{\frac{\pi}{\eta_{M'}\gamma}}e^{\frac{\eta_{M'}\gamma}{2}} \tag{33}$$

where $\eta_{m'}$ and $\eta_{M'}$ are given by:

$$\begin{aligned}\eta_{m'} &\triangleq \frac{(\gamma/2 - 1/2 - \Lambda(\gamma))}{\gamma} \\ \eta_{M'} &\triangleq \frac{(\gamma/2 - 1/2 + \Lambda(\gamma))}{\gamma} \\ \Lambda(\gamma) &= \sqrt{(\gamma/2)^2 - 3\gamma/2 + 1/4},\end{aligned}$$

the bad solution region is given by $\mathcal{R}_b = (0, \eta_{m'}]$ whereas the good solution region is $\mathcal{R}_g = [\eta_{M'}, 1)$.

Proof: The proof is analogous to that of Theorem 4.4. ■

The same consequence for the asymptotic operational region holds here:

Corollary 4.8: In the limit of high SNR, $\mathcal{R}_g \to \{1\}$, and $\mathcal{R}_b \to \{0\}$.

Proof: It results from

$$\begin{aligned}\lim_{\gamma\to\infty} \eta_M &= \lim_{\gamma\to\infty} \frac{(\gamma/2 - 1/2 + \sqrt{(\gamma/2)^2 - 3\gamma/2 + 1/4})}{\gamma} = 1 \\ \lim_{\gamma\to\infty} \eta_m &= \lim_{\gamma\to\infty} \frac{(\gamma/2 - 1/2 - \sqrt{(\gamma/2)^2 - 3\gamma/2 + 1/4})}{\gamma} = 0\end{aligned}$$

■

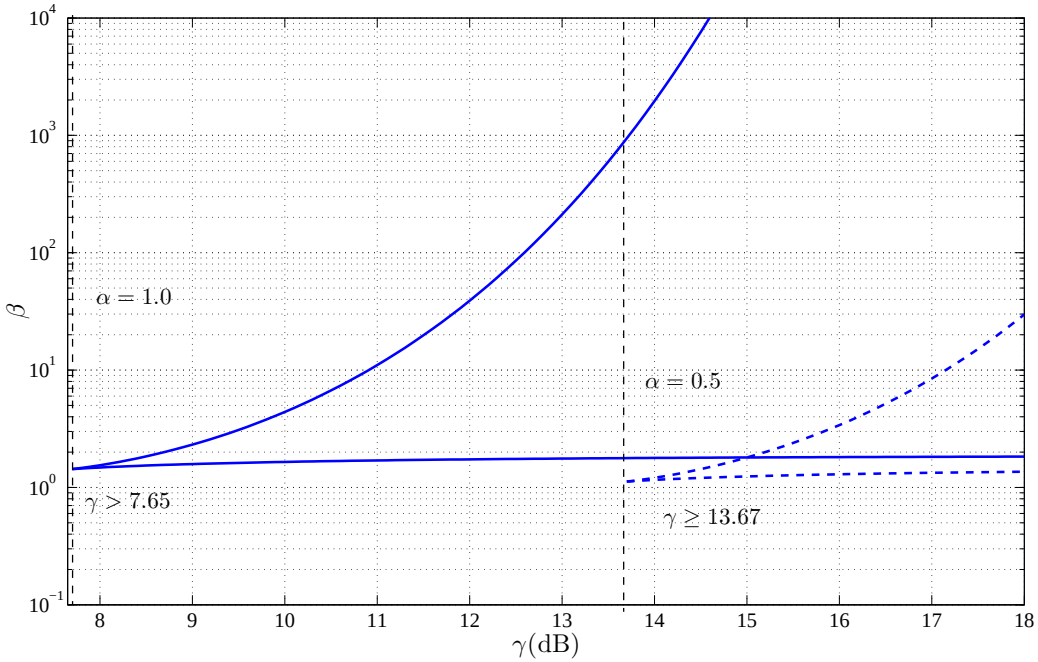

Fig. 7. Comparison of upper bounds on the spinodal lines for $\alpha = 1.0$ (left) and $\alpha = 0.5$ (right).

In Fig. 7, one can observe a 6 dB-difference between the spinodal lines corresponding to $\alpha = 0.5$ and to $\alpha = 1.0$. This is due to the minimum distance of the underlying constellations (which causes the MMSE to have different exponential decays). This can be interpreted by saying that the addition of activity detection to data detection is reflected in an increase of 6 dB of the SNR needed to achieve the same system load performance. Moreover, with $\alpha < 1$, the transition system load is lower than the case where all users are active, and, therefore, computationally good solutions admit lower values of the maximum system load.

C. Maximum system load with error probability constraints

A natural application of the above results to practical designs appears when the quality-of-service requirements of the system are specified in terms of uncoded error probability. Such an application can provide some extra insight into the plausible values of β with joint activity and data detection for efficient design of large CDMA systems. Once a multiuser-efficiency requirement is assigned, the corresponding probability of error follows naturally. Note first that, in order to detect the activity as well as the transmitted data, our model deals with a ternary constellation $\{-1, 0, 1\}$. When any of these symbols is transmitted by each user with constant SNR$= \gamma$ through a bank of large-system equivalent Gaussian channels with white noise variance $1/\eta$, the probability of error over $\mathcal{X}$ depends on the prior probabilities as well as the Euclidean distance between symbols. The actual error probability is

$$P_e = 2(1-\alpha)Q\left(\frac{\sqrt{\eta\gamma}}{2} + \frac{\lambda_\alpha}{\sqrt{\eta\gamma}}\right) + \alpha Q\left(\frac{\sqrt{\eta\gamma}}{2} - \frac{\lambda_\alpha}{\sqrt{\eta\gamma}}\right) \tag{34}$$

where $Q(x) \triangleq \frac{1}{\sqrt{2\pi}}\int_x^\infty e^{-\frac{t^2}{2}}dt$ is the Gaussian tail function, and $\lambda_\alpha \triangleq \ln\left(\frac{2(1-\alpha)}{\alpha}\right)$.

The relationship between η and P_e for our particular case can be used to reformulate the bounds on the function Υ_β in terms of error probability.

Corollary 4.9: The maximum system load, $\Upsilon_\beta(\eta, \gamma, \alpha)$, for a given error probability P_e, γ, and activity rate is bounded for high SNR by:

$$L(\gamma, \eta_{\max}, \alpha) < \Upsilon_\beta(\eta_{\max}, \gamma, \alpha) < U(\gamma, \eta_{\max}, \alpha) \tag{35}$$

where

$$\eta_{\max} \triangleq \max\{\eta_P, \eta_{gc}\}$$

$\eta_P = P_e^{-1}(\gamma, \alpha)$.

Proof: The result is obtained by noticing that the multiuser efficiency requirement extracted from P_e must lie on the subregion $[\eta_{gc}, 1)$. ■

Notice that, if the error probability satisfies $\eta_p \leq \eta_{gc} = \eta_{\max}$, then the constraint is given by the general transition load. However, if $\eta_{gc} < \eta_p = \eta_{\max}$, then, for Corollary 4.6, the maximum system load can be easily lower-bounded. Fig. 8 plots the critical system load for two different requirements in error probabilities and three different users' activity rate.

V. Spectral efficiency of a CDMA system with an unknown number of users

Statistical-physics tools can also be applied to analyze the information-theoretic capacity of CDMA channels. As shown in [5], the large-system capacity per user of CDMA can be written as

$$C = \lim_{K\to\infty} \frac{1}{K} I(b^1, \ldots, b^K; \boldsymbol{y}) = \mathcal{F} - \frac{1}{2\beta}\log(2\pi e) \tag{36}$$

where $I(\cdot\,;\cdot)$ denotes mutual information, $b^1, \ldots, b^K$ are the users' transmitted symbols, $\boldsymbol{y}$ is the received signal, and $\mathcal{F}$ is the free energy (15).

Under the decoupling principle, the mutual information between the input symbol and the output of any MAP estimator turns out to be equal to the input–output mutual information of the Gaussian equivalent channel, so that

$$I(b^k; y) = \mathbb{E}_{b^k, y}\left[\log\left(\frac{P(y|\gamma, \eta, b^k)}{P(y)}\right)\right] \tag{37}$$

where $P(b^k)$ is the input distribution and $P(y) = \sum_{b^k} P(b^k)P(y|\gamma, \eta, b^k)$ is the output density. Finally, the spectral efficiency under joint decoding is obtained by adding the single-user mutual information

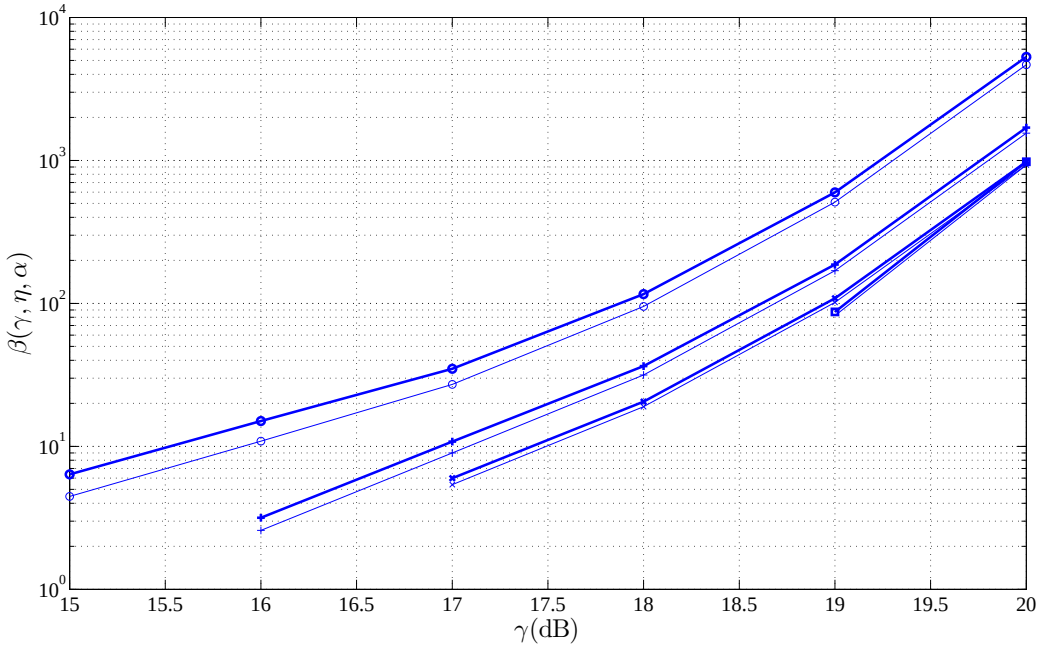

Fig. 8. Critical system load for different uncoded error probabilities and activity rates. Thicker lines represent numerical results, whereas regular lines show the corresponding lower bounds. Lines with circle markers: $P_e = 10^{-3}$ and $\alpha = 0.99$. Lines with cross markers: $P_e = 10^{-3}$ and $\alpha = 0.1$. Lines with star markers: $P_e = 10^{-3}$ and $\alpha = 0.5$. Lines with square markers: error probability 10^{-5} and $\alpha = 0.5$.

(normalized to the dimensionality of the multiuser channel—the spreading length) to a divergence factor depending on the multiuser efficiency η [8]:

$$\mathcal{C} = \beta\, \mathbb{E}_\gamma \left[I(b^k, y) \right] + \frac{1}{2} \left[(\eta - 1) \log e - \log \eta \right] \tag{38}$$

Based on the above, in this section we present some results concerning single-user and spectral capacity for the model introduced in this work.

Proposition 5.1: In randomly spread DS-CDMA with constant equal power per user, the large-system single-user capacity of a multiuser channel with MAP user identification and data detection and BPSK transmission is

$$\begin{aligned} C(\alpha) = &- \alpha \int \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} \log\left(\alpha \cosh(\eta\gamma - y\sqrt{\eta\gamma}) + (1-\alpha) e^{\eta\gamma/2} \right) \mathrm{d}y \\ &- (1-\alpha) \int \frac{e^{-\frac{(y-\sqrt{\eta\gamma})^2}{2}}}{\sqrt{2\pi}} \log\left(\alpha \cosh(\eta\gamma - y\sqrt{\eta\gamma}) + (1-\alpha) e^{\eta\gamma/2} \right) \mathrm{d}y \\ &+ \eta\gamma \frac{(\alpha+1)}{2} \log(e) \end{aligned} \tag{39}$$

where η, the multiuser efficiency, is the solution of (18) that minimizes the free energy (15), and the logarithm is in base-3 over $\alpha \in [0, 1)$, and in base-2 for $\alpha = 1.0$.

Proof: The single-user channel capacity and the joint spectral efficiency for each case are calculated from (37) and (38), by using the multiuser-efficiency fixed-point solution that yields the smallest free energy (or, equivalently, the smallest spectral efficiency). ■

Corollary 5.2: The optimal spectral efficiency of a detector performing user identification and data detection is

$$\mathcal{C}(\alpha) = \beta C(\alpha) + \frac{1}{2}\left[(\eta - 1)\log e - \log \eta\right] \tag{40}$$

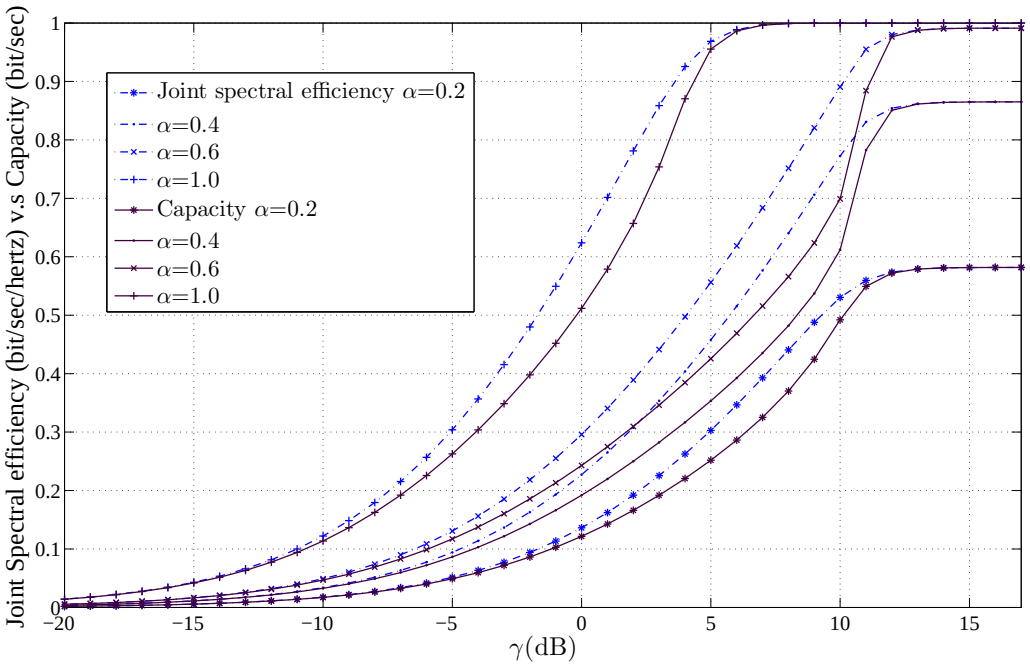

Fig. 9. Large-system joint spectral efficiency and channel capacity with MAP detection of user identities and their data with prior knowledge of α and β=1.

For sufficiently large SNR ($\gamma\eta \to \infty$) the limiting spectral efficiency is given by the following corollary:

Corollary 5.3: The optimal spectral efficiency of a detector performing user identification and data detection for constant and equally-powered users satisfies

$$\lim_{\eta\gamma\to\infty} \mathcal{C}(\alpha) = \beta\left(\alpha \log(2) + H_2(\alpha)\right) \tag{41}$$

where $H_2(\alpha) = -\alpha\log(\alpha) - (1-\alpha)\log(1-\alpha)$ is the binary entropy, and the $\log$ function is in base 3 for $\alpha \in (0,1)$, and in base 2 for $\alpha = 1$.

Proof: This result follows straightforwardly from the fact that the conditional entropy of the equivalent Gaussian channel satisfies $H(b^k|y) \to 0$ in the large SNR limit, and, consequently, $I(b^k; y) \to H(b^k)$, where b^k is the input user symbol. ■

Fig. 9 compares single-user capacity and spectral efficiency of the MAP detector of user identities and data for different values of α. The capacity curves show the additional cost incurred by the estimation of active-user identities with respect to the case where all users are known to be active ($\alpha = 1.0$). The maximum capacity is achieved for $\alpha = 2/3$, which corresponds to equal prior distribution of data and activity We observe from the figure that, in the low-SNR regime, the multiuser efficiency tends to compensate the differences in capacity due to the effect of users' activity in their a priori distribution. However, for high SNR, the multiuser efficiency approaches 1, and therefore the knowledge of the activity rate is the main factor that determines the difference in performance for each case. The limiting values as $\gamma \to \infty$ of each curve are given by Corollary 5.3.

VI. Conclusions

We have analyzed multiuser detectors for CDMA where the fraction of active users is unknown, and must be estimated in a tracking phase. Using a large-system approach and statistical-physics tools, we have derived a fixed-point equation for the optimal user-and-data detector, and provided asymptotic bounds for the corresponding MMSE. Further, we have described the space of of stable solutions of the fixed-point equation, and derived explicit bounds on the critical and transition system loads for all user's activity rates. These are consistent with the results obtained under the classic multiuser-detection assumption ($\alpha = 1.0$) made in previous literature. The study of the so-called spinodal lines allowed us to determine the regions of stable good and bad solutions, including subregions of single solutions (also called computationally feasible), in the system load–SNR parameter space of our model. Our results show that for a user-and-data detector, the boundaries of the space of solutions do depend on the activity rate, whereas the regions of stable solutions (good and bad) are only affected by the SNR. Hence, the overall system load performance keeps a symmetric behavior with respect to α. In practical implementations with high quality-of-service demands, we are interested in a critical system load as large as possible, keeping the optimal detector in the feasible subregion of good multiuser efficiencies, so that a a wider range of potential users can access successfully the channel with a given rate. By increasing the SNR, this goal can be achieved, but for limited SNR, the certainty on the users' activity allows allocation of more users for a given spreading length. A relevant example corresponds to a system requirement expressed in terms of error probability. For this case, we have shown that, for sufficiently large SNR, we can choose the minimum multiuser efficiency in the domain of feasible good solutions, and maximize the critical system load regardless of the error probability target.

We have also addressed the derivation of the asymptotic single-user capacity and the spectral efficiency under the decoupling principle. The results show the sensitivity of the system performance to the activity rate, and its impairment due to the uncertainty on the number of users accessing the system.

APPENDIX I
PROOF OF PROPOSITION 3.1

We first derive the MMSE for our particular ternary input distribution in the general fixed-point equation (14):

$$\begin{aligned}
\text{MMSE}(\gamma,\eta,\alpha) &= \mathbb{E}\left[\left(\boldsymbol{b}^k - \mathbb{E}\{\boldsymbol{b}^k|\boldsymbol{S},y\}\right)^2\right] \\
&= \alpha - \int \frac{\mathbb{E}^2_{\boldsymbol{b}^k}\{\boldsymbol{b}^k P(y|\eta,\boldsymbol{b}^k,S)\}}{\mathbb{E}_{\boldsymbol{b}^k}\{P(y|\eta,\boldsymbol{b}^k,S)\}}\mathrm{d}y \\
&= \alpha - \int \frac{\left(\sqrt{\frac{\eta}{2\pi}}\frac{\alpha}{2}\left[e^{-\frac{\eta}{2}(y-\sqrt{\gamma})^2} - e^{-\frac{\eta}{2}(y+\sqrt{\gamma})^2}\right]\right)^2}{\sqrt{\frac{\eta}{2\pi}}\left(\frac{\alpha}{2}\left[e^{-\frac{\eta}{2}(y-\sqrt{\gamma})^2} + e^{-\frac{\eta}{2}(y+\sqrt{\gamma})^2}\right]\right) + (1-\alpha)e^{\frac{\eta}{2}y^2}}\mathrm{d}y \\
&= \alpha - \int \frac{\left(\sqrt{\frac{\eta}{2\pi}}\frac{\alpha}{2}\left[e^{-\frac{\eta}{2}(y-\sqrt{\gamma})^2} - e^{-\frac{\eta}{2}(y+\sqrt{\gamma})^2}\right]\right)\alpha\sinh(\eta y\sqrt{\gamma})}{\alpha\cosh(\eta y\sqrt{\gamma}) + (1-\alpha)e^{\frac{\eta\gamma}{2}}}\mathrm{d}y \\
&= \alpha - \frac{1}{2}\sqrt{\frac{\eta}{2\pi}}\left[\int \frac{e^{-\frac{\eta}{2}(y-\sqrt{\gamma})^2}\alpha^2\sinh(\eta y\sqrt{\gamma})}{\alpha\cosh(\eta y\sqrt{\gamma}) + (1-\alpha)e^{\frac{\eta\gamma}{2}}}\mathrm{d}y - \int \frac{e^{-\frac{\eta}{2}(y'+\sqrt{\gamma})^2}\alpha^2\sinh(\eta y'\sqrt{\gamma})}{\alpha\cosh(\eta y'\sqrt{\gamma}) + (1-\alpha)e^{\frac{\eta\gamma}{2}}}\mathrm{d}y'\right]
\end{aligned}$$

At this stage, we use the change of variables $\left\{z = \sqrt{\eta}(\sqrt{\gamma} - y),\ z' = -y'\right\}$, which yields

$$\alpha - \frac{1}{2}\sqrt{\frac{\eta}{2\pi}}\left[\int \frac{e^{-\frac{z^2}{2}}\alpha^2\sinh(\eta\gamma - z\sqrt{\eta\gamma})}{\alpha\cosh(\eta\gamma - z\sqrt{\eta\gamma}) + (1-\alpha)e^{\frac{\eta\gamma}{2}}}\mathrm{d}z + \int \frac{e^{-\frac{\eta}{2}(-z'+\sqrt{\gamma})^2}\alpha^2\sinh(\eta z'\sqrt{\gamma})}{\alpha\cosh(\eta z'\sqrt{\gamma}) + (1-\alpha)e^{\frac{\eta\gamma}{2}}}\mathrm{d}z'\right]$$

Finally, the change of variable $\left\{\hat{z} = \sqrt{\eta}(\sqrt{\gamma} - z')\right\}$ leads to the following MMSE:

$$\begin{aligned}
\text{MMSE}(\gamma,\eta,\alpha) &= \alpha - \frac{1}{2}\frac{1}{\sqrt{2\pi}}\left[\int \frac{e^{-\frac{z^2}{2}}\alpha^2\sinh(\eta\gamma - z\sqrt{\eta\gamma})}{\alpha\cosh(\eta\gamma - z\sqrt{\eta\gamma}) + (1-\alpha)e^{\eta\gamma/2}}\mathrm{d}z\right. \\
&\qquad \left. + \int \frac{e^{\frac{-\hat{z}^2}{2}}\alpha^2\sinh(\eta\gamma - \hat{z}\sqrt{\eta\gamma})}{\alpha\cosh(\eta\gamma - \hat{z}\sqrt{\eta\gamma}) + (1-\alpha)e^{\eta\gamma/2}}\mathrm{d}z'\right] \\
&= \alpha - \int \frac{1}{\sqrt{2\pi}}e^{\frac{-y^2}{2}}\frac{\alpha^2\sinh[\eta\gamma - y\sqrt{\eta\gamma}]}{\alpha\cosh[\eta\gamma - y\sqrt{\eta\gamma}] + (1-\alpha)e^{\eta\gamma/2}}\mathrm{d}y
\end{aligned}$$

Since the SNR is constant among users, it follows naturally that the large-system fixed-point equation is given by (18).

APPENDIX II
PROOF OF THEOREM 3.3

From now on, for simpliciy's sake we omit the explicit indication of the arguments of the MMSE function. The lower bound is obtained by noting that, for large SNR, the general MMSE, denoted MMSE_{α}, is lower-bounded by $\mathrm{MMSE}_{\{0,1\}}$, which describes the detection performance when the transmitted symbols are $\{0,1\}$, with probabilities $\{1-\alpha,\alpha\}$. In this case the MMSE has the following form:

$$\begin{aligned}\mathrm{MMSE}_{\{0,1\}} &= \alpha - \int \frac{\left(\sqrt{\frac{\eta}{2\pi}}\alpha e^{-\frac{\eta}{2}(y-\sqrt{\gamma})^2}\right)^2}{\sqrt{\frac{\eta}{2\pi}}\left(\alpha e^{-\frac{\eta}{2}(y-\sqrt{\gamma})^2}+(1-\alpha)e^{-\frac{\eta}{2}y^2}\right)}\mathrm{d}y \\ &= \alpha - \alpha\sqrt{\frac{\eta}{2\pi}}\int_{-\infty}^{\infty}\frac{e^{-\frac{\eta}{2}(y-\sqrt{\gamma})^2}}{1+e^{-y\eta\sqrt{\gamma}+\eta\gamma/2+2\lambda_\alpha}}\mathrm{d}y\end{aligned}$$

where $\lambda_\alpha \triangleq \frac{1}{2}\ln\left(\frac{1-\alpha}{\alpha}\right)$. We make the change of variable $\hat{y}=\sqrt{\eta}y$ to obtain

$$\begin{aligned}\mathrm{MMSE}_{\{0,1\}} &= \alpha - \alpha\sqrt{\frac{1}{2\pi}}\int_{-\infty}^{\infty}\frac{e^{-\frac{1}{2}(\hat{y}-\sqrt{\eta\gamma})^2}}{1+e^{-\hat{y}\sqrt{\eta\gamma}+\eta\gamma/2+2\lambda_\alpha}}\mathrm{d}\hat{y} \\ &= \alpha - \alpha\sqrt{\frac{1}{2\pi}}\int_{-\infty}^{\infty}\frac{e^{-\frac{1}{2}(\hat{y}-\sqrt{\eta\gamma})^2+\frac{\hat{y}\sqrt{\eta\gamma}}{2}-\frac{\eta\gamma}{4}-\lambda_\alpha}}{2\cosh\left(-\frac{\hat{y}\sqrt{\eta\gamma}}{2}+\frac{\eta\gamma}{4}+\lambda_\alpha\right)}\mathrm{d}\hat{y}\end{aligned}$$

Manipulation of the denominator yields

$$\mathrm{MMSE}_{\{0,1\}} = \alpha - \alpha e^{-\frac{3\eta\gamma}{8}-\lambda_\alpha}\sqrt{\frac{1}{8\pi}}\int_{-\infty}^{\infty}e^{-\frac{1}{2}(\hat{y}-\frac{3\sqrt{\eta\gamma}}{2})^2}\,\mathrm{sech}\left(-\frac{\hat{y}\sqrt{\eta\gamma}}{2}+\frac{\eta\gamma}{4}+\lambda_\alpha\right)\mathrm{d}\hat{y}$$

After the change of variable $\hat{y}=2z+\frac{\sqrt{\eta\gamma}}{2}+\frac{2\lambda_\alpha}{\sqrt{\eta\gamma}}$, we obtain

$$\mathrm{MMSE}_{\{0,1\}} = \alpha - \alpha e^{-\frac{3\eta\gamma}{8}-\lambda_\alpha}\sqrt{\frac{1}{2\pi}}\int_{-\infty}^{\infty}e^{-\frac{1}{2}(2z-\sqrt{\eta\gamma}+\frac{2\lambda_\alpha}{\sqrt{\eta\gamma}})^2}\,\mathrm{sech}\left(z\sqrt{\eta\gamma}\right)\mathrm{d}z$$

Use now the following asymptotic expansion for $\mathrm{sech}(z)$: $|z|\to\infty$:

$$\mathrm{sech}(z) = 2e^{-|z|}\left(1+\sum_{\ell=1}^{\infty}(-1)^\ell e^{-2\ell|z|}\right)$$

and obtain

$$\begin{aligned}\mathrm{MMSE}_{\{0,1\}} = \alpha - \alpha e^{-\frac{3\eta\gamma}{8}-\lambda_\alpha}\sqrt{\frac{1}{2\pi}}\Bigg[&\int_{-\infty}^{0}e^{-\frac{1}{2}(2z-\sqrt{\eta\gamma}+\frac{2\lambda_\alpha}{\sqrt{\eta\gamma}})^2}2e^{z\sqrt{\eta\gamma}}\left(1+\sum_{\ell=1}^{\infty}(-1)^\ell e^{2\ell z\sqrt{\eta\gamma}}\right)\mathrm{d}z \\ &+\int_{0}^{\infty}e^{-\frac{1}{2}(2z-\sqrt{\eta\gamma}+\frac{2\lambda_\alpha}{\sqrt{\eta\gamma}})^2}2e^{-z\sqrt{\eta\gamma}}\left(1+\sum_{\ell=1}^{\infty}(-1)^\ell e^{-2\ell z\sqrt{\eta\gamma}}\right)\mathrm{d}z\Bigg]\end{aligned}$$

We now use the expansion

$$\text{MMSE}_{\{0,1\}} = \alpha - \alpha e^{-\eta\gamma/8+\lambda_\alpha-\frac{2F_\alpha^2}{\eta\gamma}}\left[\sum_{\ell=0}^{\infty}(-1)^\ell\sqrt{\frac{2}{\pi}}\int_{-\infty}^{0} e^{-2z^2+\left((3+2\ell)\sqrt{\eta\gamma}-\frac{4\lambda_\alpha}{\sqrt{\eta\gamma}}\right)z}\mathrm{d}z + \sum_{\ell=0}^{\infty}(-1)^\ell\sqrt{\frac{2}{\pi}}\int_{0}^{\infty} e^{-2z^2+\left((1-2\ell)\sqrt{\eta\gamma}-\frac{4\lambda_\alpha}{\sqrt{\eta\gamma}}\right)z}\mathrm{d}z\right]$$

Express now the integrals in terms of the Q-function, $Q(x) \triangleq \frac{1}{\sqrt{2\pi}}\int_x^\infty e^{-\frac{t^2}{2}}dt$:

$$\begin{aligned}\text{MMSE}_{\{0,1\}} = \alpha e^{-\eta\gamma/8+\lambda_\alpha-\frac{2F_\alpha^2}{\eta\gamma}}\Bigg[&e^{\frac{\left(\frac{\sqrt{\eta\gamma}}{2}-\frac{2\lambda_\alpha}{\sqrt{\eta\gamma}}\right)^2}{2}}Q\left(\frac{\sqrt{\eta\gamma}}{2}-\frac{2\lambda_\alpha}{\sqrt{\eta\gamma}}\right) - e^{\frac{\left(\frac{3\sqrt{\eta\gamma}}{2}-\frac{2\lambda_\alpha}{\sqrt{\eta\gamma}}\right)^2}{2}}Q\left(\frac{3\sqrt{\eta\gamma}}{2}-\frac{2\lambda_\alpha}{\sqrt{\eta\gamma}}\right)\\ &-\sum_{\ell=1}^{\infty}(-1)^\ell\Bigg[e^{\frac{\left(\frac{(2\ell-1)\sqrt{\eta\gamma}}{2}+\frac{2\lambda_\alpha}{\sqrt{\eta\gamma}}\right)^2}{2}}Q\left(\frac{(2\ell-1)\sqrt{\eta\gamma}}{2}+\frac{2\lambda_\alpha}{\sqrt{\eta\gamma}}\right)\\ &+e^{\frac{\left(\frac{(3+2\ell)\sqrt{\eta\gamma}}{2}-\frac{2\lambda_\alpha}{\sqrt{\eta\gamma}}\right)^2}{2}}\left(Q\left(\frac{(3+2\ell)\sqrt{\eta\gamma}}{2}-\frac{2\lambda_\alpha}{\sqrt{\eta\gamma}}\right)\right)\Bigg]\Bigg]\end{aligned}$$

Next, use the expansion of the Q function, [1]:

$$Q(x) = \frac{e^{-x^2/2}}{\sqrt{2\pi}x}\left(1+\sum_{\ell=1}^{\infty}(-1)^\ell\frac{\prod_{q=1}^{\ell}(2q-1)}{x^{2\ell}}\right) \tag{42}$$

to obtain

$$\text{MMSE}_{\{0,1\}} = 2\sqrt{\frac{\alpha(1-\alpha)}{2\pi\eta\gamma}}e^{-\eta\gamma/8-\frac{2F_\alpha^2}{\eta\gamma}}\left[\frac{1}{\left(1-\frac{4\lambda_\alpha}{\eta\gamma}\right)}+\frac{1}{\left(1+\frac{4\lambda_\alpha}{\eta\gamma}\right)}-\frac{1}{\left(3-\frac{4\lambda_\alpha}{\eta\gamma}\right)}-\frac{1}{\left(3+\frac{4\lambda_\alpha}{\eta\gamma}\right)} + \cdots + \mathcal{O}\left(\frac{1}{\sqrt{\eta\gamma}}\right)\right]$$

where the linear term in λ_α is substituted in the common factor. Assuming large enough $\eta\gamma$, and using the result

$$2\sum_{n=0}^{\infty}\frac{(-1)^{n+1}}{2n+1} = \frac{\pi}{2}$$

we obtain the following lower bound:

$$\text{MMSE}_{\{0,1\}} > 2\sqrt{\frac{\alpha(1-\alpha)}{2\pi\eta\gamma}}e^{-\eta\gamma/8}\sqrt{2} = 2\sqrt{\frac{\alpha(1-\alpha)}{\pi\eta\gamma}}e^{-\eta\gamma/8}$$

As far as the upper-bound is concerned, we derive the general MMSE, denoted MMSE_α, and its particularization to the case when all users are assumed to be active, denoted MMSE_1. Hence, we express the corresponding integrals in an analogous manner:

$$\begin{aligned}\zeta_\alpha &= \int\frac{1}{\sqrt{2\pi}}e^{\frac{-(y-\sqrt{\eta\gamma})^2}{2}}\frac{\alpha\sinh(y\sqrt{\eta\gamma})}{\alpha\cosh(y\sqrt{\eta\gamma})+(1-\alpha)e^{\eta\frac{\gamma}{2}}}\mathrm{d}y.\\ \zeta_1 &= \int\frac{1}{\sqrt{2\pi}}e^{\frac{-(y-\sqrt{\eta\gamma})^2}{2}}\tanh(y\sqrt{\eta\gamma})\mathrm{d}y\end{aligned}$$

We obtain:

$$\begin{aligned}\text{MMSE}_\alpha &= \alpha(1-\zeta_\alpha) = \alpha\left(1+\zeta_1-\zeta_1-\zeta_\alpha\right) = \alpha\left((1-\zeta_1)+(\zeta_1-\zeta_\alpha)\right)\\ &= \alpha\left(\text{MMSE}_1+(\zeta_1-\zeta_\alpha)\right)\end{aligned}$$

Next, we expand $\alpha(\zeta_1-\zeta_\alpha)$, which yields

$$\begin{aligned}\alpha(\zeta_1-\zeta_\alpha) &= (1-\alpha)e^{\eta\gamma/2}\int\frac{1}{\sqrt{2\pi}}e^{\frac{-(y-\sqrt{\eta\gamma})^2}{2}}\frac{\alpha\sinh[y\sqrt{\eta\gamma}]}{\alpha\cosh^2[y\sqrt{\eta\gamma}]+(1-\alpha)e^{\eta\gamma/2}\cosh(y\sqrt{\eta\gamma})}\mathrm{d}y\\ &= (1-\alpha)e^{\eta\gamma/2}\int\frac{1}{\sqrt{2\pi}}e^{\frac{-(y-\sqrt{\eta\gamma})^2}{2}}\frac{\tanh[y\sqrt{\eta\gamma}]}{\cosh[y\sqrt{\eta\gamma}]+e^{\eta\gamma/2+\ln\left(\frac{1-\alpha}{\alpha}\right)}}\mathrm{d}y\end{aligned} \tag{43}$$

Consider now the following inequalities:

$$\begin{aligned}\tanh(z) &\le 1\\ \cosh(z) &\ge \frac{e^z}{2}\end{aligned}$$

After substitution and manipulation of the denominator of (43), we obtain

$$\begin{aligned}\alpha\left(\zeta_1-\zeta_\alpha\right) \le& (1-\alpha)e^{\eta\gamma/2}\int_{-\infty}^{\infty}\frac{e^{\frac{-(y-\sqrt{\eta\gamma})^2}{2}}}{\sqrt{2\pi}}\frac{1}{\frac{e^{y\sqrt{\eta\gamma}}}{2}+e^{\eta\gamma/2+\ln\left(\frac{(1-\alpha)}{\alpha}\right)}}\mathrm{d}y\\ =& (1-\alpha)e^{\eta\gamma/2}\int_{-\infty}^{\infty}\frac{e^{\frac{-(y-\sqrt{\eta\gamma})^2}{2}}}{\sqrt{2\pi}}\frac{e^{-\frac{y\sqrt{\eta\gamma}}{2}-\frac{\eta\gamma}{4}-\phi_\alpha}}{\cosh\left(\frac{y\sqrt{\eta\gamma}}{2}-\frac{\eta\gamma}{4}-\phi_\alpha\right)}\mathrm{d}y\end{aligned}$$

where $\phi_\alpha \triangleq \frac{1}{2}\ln\left(\frac{2(1-\alpha)}{\alpha}\right)$. We readjust terms to express the integral in a convenient form:

$$\alpha\left(\zeta_1-\zeta_\alpha\right) \le (1-\alpha)e^{-\eta\gamma/8-\phi_\alpha}\int_{-\infty}^{\infty}\frac{e^{\frac{-\left(y-\frac{\sqrt{\eta\gamma}}{2}\right)^2}{2}}}{\sqrt{2\pi}}\operatorname{sech}\left(\frac{y\sqrt{\eta\gamma}}{2}-\frac{\eta\gamma}{4}-\phi_\alpha\right)\mathrm{d}y$$

We now use the following transformation in the variable of integration: $y = 2z+\frac{\sqrt{\eta\gamma}}{2}+\frac{2\phi_\alpha}{\sqrt{\eta\gamma}}$.

$$\begin{aligned}\alpha\left(\zeta_1-\zeta_\alpha\right) \le& 2(1-\alpha)e^{\frac{7\eta\gamma}{8}-\phi_\alpha}\int_{-\infty}^{0}\frac{e^{\frac{-\left(2\hat{z}-2\sqrt{\eta\gamma}-\frac{2\phi_\alpha}{\sqrt{\eta\gamma}}\right)^2}{2}}}{\sqrt{2\pi}}\operatorname{sech}\left(z\sqrt{\eta\gamma}\right)\mathrm{d}\hat{z}\\ &+2(1-\alpha)e^{-\eta\gamma/8-\phi_\alpha}\int_{0}^{\infty}\frac{e^{\frac{-\left(2z+\frac{2\phi_\alpha}{\sqrt{\eta\gamma}}\right)^2}{2}}}{\sqrt{2\pi}}\operatorname{sech}\left(z\sqrt{\eta\gamma}\right)\mathrm{d}z\end{aligned}$$

and the asymptotic expansion for $\operatorname{sech}(z)$ in the above derivation:

$$\operatorname{sech}(z) = 2e^{-|z|}\left(1+\sum_{\ell=1}^{\infty}(-1)^\ell e^{-2\ell|z|}\right)$$

This yields:

$$\begin{aligned}\alpha\left(\zeta_1-\zeta_\alpha\right) \le& 4(1-\alpha)e^{-\eta\gamma/8-\phi_\alpha-\frac{2\phi_\alpha^2}{\eta\gamma}}\sum_{\ell=0}^{\infty}(-1)^\ell\int_{0}^{\infty}\frac{e^{-2\hat{z}^2+\left((1+2\ell)\sqrt{\eta\gamma}-\frac{4\phi_\alpha}{\sqrt{\eta\gamma}}\right)\hat{z}}}{\sqrt{2\pi}}\mathrm{d}z\\ &+4(1-\alpha)e^{-\eta\gamma/8-\phi_\alpha-\frac{2\phi_\alpha^2}{\eta\gamma}}\sum_{\ell=0}^{\infty}(-1)^\ell\int_{0}^{\infty}\frac{e^{-2\hat{z}^2-\left((1+2\ell)\sqrt{\eta\gamma}+\frac{4\phi_\alpha}{\sqrt{\eta\gamma}}\right)\hat{z}}}{\sqrt{2\pi}}\mathrm{d}z\end{aligned}$$

Finally, expressing the integrals in terms of the Q function:

$$\alpha\left(\zeta_1-\zeta_\alpha\right) \leq 2(1-\alpha)e^{-\eta\gamma/8-\phi_\alpha-\frac{2\phi_\alpha^2}{\eta\gamma}}\left[\sum_{\ell=0}^{\infty}(-1)^\ell e^{\frac{\left(\frac{(1+2\ell)\sqrt{\eta\gamma}}{2}-\frac{2\phi_\alpha}{\sqrt{\eta\gamma}}\right)^2}{2}}Q\left(\frac{(1+2\ell)\sqrt{\eta\gamma}}{2}-\frac{2\phi_\alpha}{\sqrt{\eta\gamma}}\right)\right.$$
$$\left.+\sum_{\ell=0}^{\infty}(-1)^\ell e^{\frac{\left(\frac{(1+2\ell)\sqrt{\eta\gamma}}{2}+\frac{2\phi_\alpha}{\sqrt{\eta\gamma}}\right)^2}{2}}Q\left(\frac{(1+2\ell)\sqrt{\eta\gamma}}{2}+\frac{2\phi_\alpha}{\sqrt{\eta\gamma}}\right)\right]$$

and manipulating the expansion:

$$\begin{aligned}\alpha\left(\zeta_1-\zeta_\alpha\right) \leq & 2(1-\alpha)e^{-\eta\gamma/8-\phi_\alpha-\frac{2\phi_\alpha^2}{\eta\gamma}}\left[e^{\frac{\left(\frac{\sqrt{\eta\gamma}}{2}-\frac{2\phi_\alpha}{\sqrt{\eta\gamma}}\right)^2}{2}}Q\left(\frac{\sqrt{\eta\gamma}}{2}-\frac{2\phi_\alpha}{\sqrt{\eta\gamma}}\right)\right.\\
&+e^{\frac{\left(\frac{\sqrt{\eta\gamma}}{2}+\frac{2\phi_\alpha}{\sqrt{\eta\gamma}}\right)^2}{2}}Q\left(\frac{\sqrt{\eta\gamma}}{2}+\frac{2\phi_\alpha}{\sqrt{\eta\gamma}}\right)\\
&+\sum_{\ell=1}^{\infty}(-1)^\ell e^{\frac{\left(\frac{(1+2\ell)\sqrt{\eta\gamma}}{2}-\frac{2\phi_\alpha}{\sqrt{\eta\gamma}}\right)^2}{2}}Q\left(\frac{(1+2\ell)\sqrt{\eta\gamma}}{2}-\frac{2\phi_\alpha}{\sqrt{\eta\gamma}}\right)\\
&\left.+\sum_{\ell=1}^{\infty}(-1)^\ell e^{\frac{\left(\frac{(1+2\ell)\sqrt{\eta\gamma}}{2}+\frac{2\phi_\alpha}{\sqrt{\eta\gamma}}\right)^2}{2}}Q\left(\frac{(1+2\ell)\sqrt{\eta\gamma}}{2}+\frac{2\phi_\alpha}{\sqrt{\eta\gamma}}\right)\right]\end{aligned}$$

we obtain, after using the series expansion (42),

$$\begin{aligned}\alpha\left(\zeta_1-\zeta_\alpha\right) \leq & 2\sqrt{\frac{(1-\alpha)\alpha}{\pi\eta\gamma}}e^{-\eta\gamma/8}\left[\frac{1}{\left(1-\frac{4\phi_\alpha}{\eta\gamma}\right)}+\frac{1}{\left(1+\frac{4\phi_\alpha}{\eta\gamma}\right)}-\frac{1}{\left(3-\frac{4\phi_\alpha}{\eta\gamma}\right)}-\frac{1}{\left(3+\frac{4\phi_\alpha}{\eta\gamma}\right)}\right.\\
&\left.+\cdots+\mathcal{O}\left(\frac{1}{\sqrt{\eta\gamma}}\right)\right]\end{aligned}$$

where the linear term in $\phi_\alpha=\frac{1}{2}\ln\left(\frac{2(1-\alpha)}{\alpha}\right)$ is substituted in the common factor, and quadratic terms are neglected.

Using the same result as before on the series $2\sum_{n=0}^{\infty}\frac{(-1)^{n+1}}{2n+1}$, we obtain the following upper bound:

$$\alpha\left(\zeta_1-\zeta_\alpha\right)<\sqrt{\frac{\pi\alpha(1-\alpha)}{\eta\gamma}}e^{-\eta\gamma/8}$$

Finally, using the upper bound given in Lemma 3.2 for BPSK, the overall MMSE can be upper-bounded by

$$\text{MMSE}_\alpha \leq 2\alpha e^{-\eta\gamma/2}+\sqrt{\frac{\pi\alpha(1-\alpha)}{\eta\gamma}}e^{-\eta\gamma/8}$$

where ρ is a constant.

APPENDIX III
PROOF OF THEOREM 4.4

We analyze the function:

$$G(\eta)=(1-\eta)e^{u\eta}/\sqrt{\eta} \tag{44}$$

where u is a constant, which entirely describes the dependence of Υ_β on η for large enough $\eta\gamma$. By simple derivation of (44), it is easy to observe that the solution has critical points in the domain $(0,1)$ if and only if $u \geq (3+\sqrt{8})/2$. These points are:

$$\begin{aligned} \eta_m &= \frac{u-1/2-\sqrt{u^2-3u+1/4}}{2u} \\ \eta_M &= \frac{u-1/2+\sqrt{u^2-3u+1/4}}{2u} \end{aligned}$$

and lie in the domain $(0,1)$. In fact, note that $u^2-3u+1/4 < (u-3/2)^2$, and thus it can be verified that $0 < 1/2u < \eta_m < \eta_M < 1-1/u < 1$. By using the second derivative of (44) we observe that these solutions correspond to a local minimum and a local maximum, respectively. Let us now study the function (44) to justify the range of values of the critical system load. $G(\eta)$ is a continuous function in $(0,1]$ that takes positive values. Since $G(\eta)$ tends to 0 as η approaches 1, and tends to infinity as η approaches 0, it can be concluded that the range for which $G(\eta)$ has only one pre-image is $(0, G(\eta_m)) \cup (G(\eta_M), \infty)$. Hence, there are single pre-images in the ranges $(0, \min\{G^{-1}(G(\eta_M))\})$ and $(\max\{G^{-1}(G(\eta_m))\}, 1)$. For $G(\eta_m) \leq G(.) \leq G(\eta_M)$, there are multiple pre-images (exactly, 3) due to the local minimum and maximum. Then, the smallest pre-image lies on $[\min\{G^{-1}(G(\eta_m))\}, \eta_M]$ whereas the largest lie on $[\eta_M, \max\{G^{-1}(G(\eta_m))\}]$. In conclusion, the overall *smallest* pre-images belong to $(0, \eta_m]$, whereas the *largest* pre-images belong to $[\eta_M, 1)$. In particular, single small and large solutions satisfy $(0, \min\{G^{-1}(G(\eta_M))\} \subset (0, \eta_m]$ and $(\max\{G^{-1}(G(\eta_m))\}, 1) \subset [\eta_M, 1)$ respectively. By bounding MMSE by (22) and (23), and replacing $u = \gamma/8$, we obtain the expected results.

REFERENCES

[1] S. Verdú, *Multiuser Detection*. Cambridge University Press, UK, 1998.

[2] K. Halford and M. Brandt-Pearce, "New-user identification in a CDMA system," *IEEE Trans. Commun.*, vol. 46, pp. 144–155, 1998.

[3] M. L. Honig and H. V. Poor, "Adaptive interference suppression in wireless communication systems," in *Wireless Communications: Signal Processing Perspectives*. 1998, H.V. Poor and G.W. Wornell, Eds Englewood Cliffs, NJ:Prentice Hall.

[4] E. Biglieri and M. Lops, "Multiuser detection in a dynamic environment- Part I: User identification and data detection," *IEEE Trans. Inform. Theory*, vol. 53, no. 9, pp. 3158–3170, September 2007.

[5] T. Tanaka, "A statistical-mechanics approach to large-system analysis of CDMA multiuser detectors," *IEEE Trans. Inform. Theory*, vol. 48, no. 11, pp. 2888–2910, November 2002.

[6] D. Guo and S. Verdú, "Randomly spread CDMA: Asymptotics via statistical physics," *IEEE Trans. Inform. Theory*, vol. 51, no. 6, pp. 1983–2007, June 2005.

[7] H. Nishimori, *Statistical Physics of Spin Glasses and Information Processing: An Introduction*, Oxford Univ. Press, Oxford, UK, 2001.

[8] D. Guo and T. Tanaka, "Generic multiuser detection and statistical physics," in *Advances in Multiuser Design*. M. L. Honig (Ed.). John Wiley & Sons, 2008.

[9] M. Mézard and A. Montanari, *Information, Physics and Computation*, Oxford Univ. Press, 2008.

[10] A. Tauste Campo and E. Biglieri, "Large-system analysis of static multiuser detection with an unknown number of users," in *Proc. of IEEE Int. Workshop on Comp. Advances in Multi-Sensor Adapt. Process. (CAMSAP'07)*, Saint Thomas, US, Dec. 2007.

[11] A. Tauste Campo and E. Biglieri, "Asymptotic capacity of a static multiuser channel with an unknown number of users," in *Proc. IEEE Int. Symp. on Wir. Per. Mult. Comms. (WPMC'08)*, Sept. 2008.

[12] E. Biglieri and M. Lops, "A new approach to multiuser detection in mobile transmission systems," in *Proc. IEEE Int. Symp. on Wir. Per. Mult. Comms. (WPMC'06)*, San Diego, CA, Sept. 2006.

[13] D. Tse and S. Hanly, "Linear multiuser receivers: effective interference, effective bandwidth, and user capacity," *IEEE Trans. Inf. Theory*, vol. 45, no. 2, pp. 641–657, March 1999.

[14] R. R. Müller, "Channel capacity and minimum probability of error in large dual antenna array systems with binary modulation," *IEEE Trans. Signal Processing*, vol. 51, no. 11, pp. 2821–2828, November 2003.

[15] R.S. Ellis, *Entropy, Large Deviations, and Statistical Mechanics*, vol. 271, A series of comprehensive studies in mathematics. Springer-Verlag, 1985.

[16] V. Poor, *An introduction to Signal Detection and Estimation*, Springer-Verlag, New York, 1988.

[17] D. Tse and S. Verdú, "Optimum asymptotic multiuser efficiency of randomly spread cdma," *IEEE Trans. Inform. Theory*, vol. 46, no. 11, pp. 2718–2722, Nov. 2000.

[18] S. Verdú and S. (Shitz) Shamai, "Spectral efficiency of CDMA with random spreading," *IEEE Trans. Inform. Theory*, vol. 45, no. 2, pp. 622–640, March 1999.

[19] M. Talagrand, *Spin Glasses: A Challenge for Mathematicians*, Springer, 2003.

[20] A. Montanari and D. Tse, "Analysis of belief propagation for non-linear problems: The example of CDMA (or: How to prove tanaka's formula)," in *Proc. IEEE Inform. Theory Workshop*, Punta del Este, Uruguay, March 2006.

[21] D. Guo and C.-C. Wang, "Random sparse linear systems observed via arbitrary channels: A decoupling principle," in *Proc. IEEE Int. Symp. Inform. Theory*, Nice, France, June 2007.

[22] A. Lozano, A. M. Tulino, and S. Verdú, "Optimum power allocation for parallel gaussian channels with arbitrary input distributions," *IEEE Trans. Inform. Theory*, vol. 52, no. 7, pp. 3033–3051, July 2006.