# Density Evolution and Power Profile Optimization for Iterative Multiuser Decoders Based on Individually Optimum Multiuser Detectors*

Ralf R. Müller

Forschungszentrum Telekommunikation Wien
Tech Gate Vienna, Donau–City Str. 1/3
1220 Wien, Austria
mueller@ftw.at

Giuseppe Caire

Institute Eurecom
2229 Route des Crétes, B.P. 193
06904 Sophia–Antipolis CEDEX, France
caire@eurecom.fr

Toshiyuki Tanaka

Tokyo Metropolitan University
1-1 Minami-Osawa, Hachioji
Tokyo 192-0397, Japan
tanaka@eei.metro-u.ac.jp

**Abstract**

Iterative multiuser joint decoding based on exact Belief Propagation (BP) is analyzed in the large system limit by means of the replica method. It is shown that performance can be improved by appropriate power assignment to the users. The optimum power assignment can be found by linear programming in most technically relevant cases.

The performance of BP iterative multiuser joint decoding is compared to suboptimum approximations based on Interference Cancellation (IC). While IC receivers show a significant loss for equal-power users, they yield performance close to BP under optimum power assignment.

## 1 Introduction

The complexity of optimum multiuser joint decoding is exponential in both the code block-length (constraint length for convolutional codes) and the number of users [6]. Since this complexity is often considered infeasible, iterative methods have been proposed. In particular, the general framework of Belief-Propagation (BP) [11] can be applied to the coded multiuser channel [19, 2]. BP is known to yield optimum performance provided that the *Bayesian network* associated with the multiuser decoding problem is free of loops [9]. Even though the Bayesian network representing to multiuser decoding problem has cycles any case of interest (for more than one user and when user codes are non-trivial), driven by the fact that BP proved to yield excellent results also in the presence of cycles [1], the application of BP to multiuser decoding is a heuristically sound approach.

*The research of T. Tanaka was performed during his visit at Aston University, Birmingham, U.K. and supported by EPSRC research grant GR/N00562.

A straightforward implementation of BP is no longer exponential in the product of the blocklength (or constraint length) and the number of users, but only exponential in their maximum. However, exponential complexity in the number of users may still be too costly for many applications. Therefore, several approximations of exact BP have been proposed, many of which being based on interference cancellation (IC) possibly followed by linear MMSE filtering (see [2] and references therein).

Recently, other approaches for approximating the exact BP decoder have gained increasing attention. In particular, by exploiting the analogy between multiuser detection and minimum-distance decoding of lattices, soft-output versions of the Sphere-Decoder [18] have been proposed [17, 8]. When the number of users is not larger than the spreading gain, the Sphere-Decoder has polynomial complexity in the number of users, comparable to that of post-IC linear MMSE filtering. However, the quantification of the potential gain achievable by using more complicated BP approximation methods than post-IC linear MMSE filtering in iterative multiuser joint decoding remains unclear.

This work provides an answer to this question by analyzing the performance of the exact BP decoder for random spreading in the large system limit. The standard tool to analyze the limit performance of iterative decoding algorithms is Density Evolution (DE) [12]. In [2, 4], the large-system limit of IC-based iterative multiuser joint decoding was obtained by using DE in conjunction with *by-now classical* results on the large-system limit of linear multiuser detection [14]. In this work, we shall use DE in conjunction with results of statistical mechanics and the replica method [13, 7].

In addition to just providing an analysis of the performance of the BP decoder in the large system limit, our analysis also allows to solve the problem of optimum power assignment to the users in many cases of technical interest.

# 2 Synchronous CDMA system model

We consider the real-valued discrete-time channel model

$$\mathbf{Y} = \mathbf{SWX} + \mathbf{N} \tag{1}$$

originated by sampling at the chip-rate a synchronous CDMA system [16], where:

1. $\mathbf{Y}, \mathbf{N} \in \mathbb{R}^{L \times N}$, are the matrices of received chip-rate samples and the corresponding AWGN samples $\sim \mathcal{N}(0, \sigma_0^2)$;

2. $\mathbf{S} \in \mathbb{R}^{L \times K}$ contains the user spreading sequences by columns;

3. $\mathbf{W} = \mathrm{diag}(w_1, \ldots, w_K)$ contains the user amplitudes which are assumed to obey the normalization $\mathrm{trace}(\mathbf{W}^2) = K$ without loss of generality;

4. $\mathbf{X} \in \{+1, -1\}^{K \times N}$ is the matrix of the users' binary modulation symbols. The row $\mathbf{x}^k$ of $\mathbf{X}$ is the code word transmitted by user $k$. The column $\mathbf{x}_n$ of $\mathbf{X}$ is the vector of symbols transmitted by all users at the same time ($n$-th symbol interval);

5. $L, K$ and $N$ denote the spreading factor, the number of users and the code block length, respectively.

Spreading sequences are random with i.i.d. Gaussian elements with mean 0, variance $1/L$, and the $k$-th user *received* SNR is $\gamma_k = w_k^2/\sigma_0^2$. Users send information messages in the form of

binary uniformly distributed vectors $\mathbf{b}_k$ of length $B$. For the sake of simplicity, we assume that all users employ the same binary convolutional code $\mathcal{C} \subseteq \{+1,-1\}^N$ but each user employs a different random interleaver. The user coding rate is given by $R = B/N$ bit/symbol and the system spectral efficiency is given by $\rho = \alpha R$, where $\alpha = K/L$ (users per chip) is the *system load*.

We avoid addressing the complex valued AWGN channel in order to keep notation at a reasonable level of complication. Extending the results obtained for the real-valued channel to the complex-valued one is a tedious but straightforward exercise.

# 3 Iterative decoding algorithms

For a binary variable $c$ with pmf $(\Pr(c=+1), \Pr(c=-1))$ we define its log-ratio by

$$\mathcal{L} \triangleq \log \frac{\Pr(c=+1)}{\Pr(c=-1)}$$

The BP algorithm approximates iteratively the log-ratios $\mathcal{L}_{k,j}^{\text{bit}}$ corresponding to the marginals of the a posteriori joint pmf $\Pr(\mathbf{b}_1, \ldots, \mathbf{b}_K | \mathbf{Y})$. After a given number of iterations, a symbol-by-symbol decision is made according to the threshold rule $\widehat{b}_{k,j} = \text{sign}(\mathcal{L}_{k,j}^{\text{bit}})$.

The main building blocks of the BP iterative multiuser joint decoder are the SISO decoders and the individually optimum MAP multiuser detector (IO-MUD). SISO decoding is formally expressed by

$$\mathcal{L}_{k,n}^{\text{dec}} = \log \frac{\displaystyle\sum_{\mathbf{c} \in \mathcal{C}_k : c_n = +1} \exp\left( \frac{1}{2} \sum_{j \neq n} c_j \mathcal{L}_{k,j}^{\text{mud}} \right)}{\displaystyle\sum_{\mathbf{c} \in \mathcal{C}_k : c_n = -1} \exp\left( \frac{1}{2} \sum_{j \neq n} c_j \mathcal{L}_{k,j}^{\text{mud}} \right)} \tag{2}$$

where $\mathcal{L}_{k,j}^{\text{mud}}$ is the message (log-ratio) sent by the IO-MUD for user $k$ relative to coded symbol $c_{k,j}$ and $\mathcal{L}_{k,n}^{\text{dec}}$ is the so called decoder "extrinsic information". For convolutional codes, (2) is efficiently implemented by the forward-backward algorithm. The same forward-backward algorithm can compute the log-ratios $\{\mathcal{L}_{k,j}^{\text{bit}} : j = 1, \ldots, B\}$ for the user information bit while computing (2).

IO-MUD consists of calculating the log-ratios

$$\mathcal{L}_{k,n}^{\text{mud}} = \log \frac{\Pr(x_{k,n} = +1 | \mathbf{y}_n, \mathcal{L}_{1,n}^{\text{dec}}, \ldots, \mathcal{L}_{k-1,n}^{\text{dec}}, \mathcal{L}_{k+1,n}^{\text{dec}}, \ldots, \mathcal{L}_{K,n}^{\text{dec}})}{\Pr(x_{k,n} = -1 | \mathbf{y}_n, \mathcal{L}_{1,n}^{\text{dec}}, \ldots, \mathcal{L}_{k-1,n}^{\text{dec}}, \mathcal{L}_{k+1,n}^{\text{dec}}, \ldots, \mathcal{L}_{K,n}^{\text{dec}})} \tag{3}$$

$$= \log \frac{\displaystyle\sum_{\mathbf{x} \in \{\pm 1\}^K : x_k = +1} \exp\left( -\frac{1}{2\sigma_0^2} \left| \mathbf{y}_n - \sum_{j=1}^{K} w_j \mathbf{s}_j x_j \right|^2 + \frac{1}{2} \sum_{j \neq k} x_j \mathcal{L}_{j,n}^{\text{dec}} \right)}{\displaystyle\sum_{\mathbf{x} \in \{\pm 1\}^K : x_k = -1} \exp\left( -\frac{1}{2\sigma_0^2} \left| \mathbf{y}_n - \sum_{j=1}^{K} w_j \mathbf{s}_j x_j \right|^2 + \frac{1}{2} \sum_{j \neq k} x_j \mathcal{L}_{j,n}^{\text{dec}} \right)} \tag{4}$$

Unfortunately, there is no efficient way to perform this calculation, in general. However, the quality of the log-ratios can be analyzed in the large system limit using the replica method.

# 4 Performance of IO-MUD with non-uniform prior

The derivation of the performance is based on the replica method which is a common tool in statistical mechanics [10]. It was introduced into the analysis of multi-user systems in [13]. The analysis presented here makes use of the generalization to arbitrary powers in [7] and further generalizes the results to arbitrary non-uniform binary priors.

The distribution at the channel output at time instant $n$ conditioned on the signature sequences is proportional to

$$Z(\mathbf{y}_n, \mathbf{S}) = \sum_{\mathbf{x}_n} \Pr(\mathbf{x}_n) \exp\left(-\tfrac{1}{2\sigma^2}|\mathbf{y}_n - \mathbf{S}\mathbf{W}\mathbf{x}_n|^2\right) \tag{5}$$

if the fictious noise variance $\sigma^2$ is set to the true noise variance $\sigma_0^2$. Moreover (5) is independent of $n$ since the input is assumed to be stationary. Thus, the time index $n$ is dropped and $n$ is used for different purpose. In statistical mechanics the quantity

$$\mathcal{F}_K(\mathbf{y}, \mathbf{S}) = \tfrac{1}{K} \log Z(\mathbf{y}, \mathbf{S}) \tag{6}$$

is called the *free energy*. One of the fundamental principles of statistical mechanics is that the free energy is self-averaging in the large system limit. That is, it is identical to its average over the spreading sequences and the noise for almost all realizations

$$\lim_{K\to\infty} \mathcal{F}_K(\mathbf{y}, \mathbf{S}) = \lim_{K\to\infty} \mathrm{E}\ \tfrac{1}{K}\ \log Z(\mathbf{y}, \mathbf{S}) \triangleq \mathcal{F}. \tag{7}$$

A standard trick in statistical mechanics is to re-write the free energy in the following way

$$\mathcal{F} = \lim_{K\to\infty} \tfrac{1}{K} \lim_{n\to 0} \tfrac{\partial}{\partial n} \log\left(\mathrm{E}\, Z^n(\mathbf{y}, \mathbf{S})\right) \tag{8}$$

with the advantage that the expectation operator has moved into the argument of the logarithm. Now, the free energy is evaluated for integer $n$ and the results is assumed to generalize to positive real $n$. This procedure is called the replica method. Though it is still lacking rigorous mathematical justification, it is a standard technique in statistical physics, see [10, 13] for further discussion.

Let $x_{ka}$ denote the $a^{\text{th}}$ replica of user $k$'s binary symbol. Define $\lambda_k = \mathcal{L}_{k,n}^{\text{dec}}$ and $t_k = \tanh(\lambda_k/2)$. Then, it is shown in [13, 7] that the free energy can be expressed by the following supremum

$$\sup_{m,q} \left\{\alpha^{-1} G(m,q) - I(m,q)\right\} \tag{9}$$

where

$$G(m,q) \triangleq \frac{1}{2} \log \frac{(1 + \frac{\alpha}{\sigma^2}(1-q))^{1-n}}{1 + \frac{\alpha}{\sigma^2}(1-q) + \frac{n}{\sigma^2}(\sigma_0^2 + \alpha(1-2m+q))} \tag{10}$$

and

$$I(m,q) \triangleq \sup_{E,F} \left\{ nEm + \tfrac{1}{2}\, n(n-1)Fq + \tfrac{n}{2}\, F - \lim_{K\to\infty} \tfrac{1}{K} \sum_{k=1}^{K} \log M_k(w_k^2 E, w_k^2 F) \right\} \tag{11}$$

with

$$M_k(E,F) \triangleq \sum_{\{x_{ka}, a=1,\ldots,n\}} \prod_{a=1}^{n} \Pr(x_{ka}) \left\{ \tfrac{1+t_k}{2} \exp\left[ E \sum_{a=1}^{n} x_{ka} + \tfrac{F}{2} \left( \sum_{a=1}^{n} x_{ka} \right)^2 \right] \right. \tag{12}$$
$$\left. \tfrac{1-t_k}{2} \exp\left[ -E \sum_{a=1}^{n} x_{ka} + \tfrac{F}{2} \left( \sum_{a=1}^{n} x_{ka} \right)^2 \right] \right\}.$$

With the following property of the Gaussian measure $Dz \triangleq \exp(-z^2/2)/\sqrt{2\pi}dz$

$$\exp\left(F\,\frac{S^2}{2}\right) = \int \exp\left(\pm\sqrt{F}zS\right) Dz \tag{13}$$

$\forall\, S \in \mathbb{R}$, we get

$$\begin{aligned} M_k(E,F) &= \int \sum_{\{x_{ka},a=1,\ldots,n\}} \prod_{a=1}^{n} \Pr(x_{ka}) \left\{ \tfrac{1+t_k}{2} \exp\left[\left(z\sqrt{F}+E\right)\sum_{a=1}^{n} x_{ka}\right] \right. \\ &\qquad \left. \tfrac{1-t_k}{2} \exp\left[-\left(z\sqrt{F}+E\right)\sum_{a=1}^{n} x_{ka}\right]\right\} Dz. \end{aligned} \tag{14}$$

Since

$$\begin{aligned} f_n &\triangleq \sum_{\{x_{ka},a=1,\ldots,n\}} \prod_{a=1}^{n} \Pr(x_{ka}) \exp\left[\left(z\sqrt{F}+E\right)\sum_{a=1}^{n} x_{ka}\right] && (15)\\ &= \sum_{x_{kn}} \Pr(x_{kn}) f_{n-1} \exp\left[\left(z\sqrt{F}+E\right)x_{kn}\right] && (16)\\ &= f_{n-1}\,\frac{\cosh\left(z\sqrt{F}+E+\lambda_k/2\right)}{\cosh(\lambda_k/2)} = \frac{\cosh^n\left(z\sqrt{F}+E+\lambda_k/2\right)}{\cosh^n(\lambda_k/2)}, && (17) \end{aligned}$$

we find

$$M_k(E,F) = \frac{\int (1+t_k)\cosh^n\left(z\sqrt{F}+E+\frac{\lambda_k}{2}\right)+(1-t_k)\cosh^n\left(z\sqrt{F}+E-\frac{\lambda_k}{2}\right)Dz}{2\cosh^n\left(\frac{\lambda_k}{2}\right)} \tag{18}$$

Following the further development in [13], that is taking derivatives with respect to $m, q, E, F$ to find the supremum points, taking derivatives with respect to $n$ and letting $n \to 0$, we find

$$\begin{aligned} m &= \mathop{\mathrm{E}}_{\lambda,w} w^2 \int \tfrac{1+t}{2}\tanh\left(zw\sqrt{F}+w^2E+\tfrac{\lambda}{2}\right) + \tfrac{1-t}{2}\tanh\left(zw\sqrt{F}+w^2E-\tfrac{\lambda}{2}\right) Dz \\ q &= \mathop{\mathrm{E}}_{\lambda,w} w^2 \int \tfrac{1+t}{2}\tanh^2\left(zw\sqrt{F}+w^2E+\tfrac{\lambda}{2}\right) + \tfrac{1-t}{2}\tanh^2\left(zw\sqrt{F}+w^2E-\tfrac{\lambda}{2}\right) Dz \\ E &= \tfrac{1}{\sigma^2+\alpha(1-m)}, \qquad F = \tfrac{\sigma_0^2+\alpha(1-2m+q)}{[\sigma^2+\alpha(1-m)]^2} \end{aligned} \tag{19}$$

and the free energy becomes

$$\begin{aligned} \mathcal{F} &= \mathop{\mathrm{E}}_{\lambda,w} \gamma' \int \tfrac{1+t}{2}\log\cosh\left(zw\sqrt{F}+w^2E+\tfrac{\lambda}{2}\right) + \tfrac{1-t}{2}\log\cosh\left(zw\sqrt{F}+w^2E-\tfrac{\lambda}{2}\right) Dz \\ &\quad + \tfrac{1}{2}\log(1-t^2) - Em - \tfrac{F(1-q)}{2} - \tfrac{1}{2\alpha}\log\left(1+\tfrac{\alpha(1-m)}{\sigma^2}\right) - \tfrac{1}{2\alpha}\tfrac{\sigma_0^2+\alpha(1-2m+q)}{\sigma^2+\alpha(1-q)}. \end{aligned} \tag{20}$$

Since we are interested in the individually optimum detector, we let the fictious noise variance approach the true noise variance $\sigma \to \sigma_0$ and find $E = F$ and $m = q$.

Note that in context of iterative decoding the priors are not true priors in the sense that the transmission of zeros and ones were not equally likely. In the iterative decoding process the priors are conditional priors. They are conditioned on the output of the last iteration step and refer to a particular time instant and particular observations at other time instances only. Since only extrinsic information is propagated, the priors affect the interfering users only. Thus, the performance of all users with identical powers is identical due to the symmetry of the interference in the large system limit. Moreover, it can be shown in the same way as in [7] that the asymmetry in performance vanishes, if performance is measured in terms of multiuser efficiency. In the same way as in [13], it can be shown that there is an equivalent AWGN

superchannel and the impact of the other users is fully characterized by the multiuser efficiency which is given by

$$\eta = \sigma_0^2 E. \tag{21}$$

Thus, the system (19) gives

$$\begin{aligned}\frac{1}{\eta} &= 1 + \alpha \left[ \tfrac{1}{\sigma_0^2} - \mathop{\mathrm{E}}_{\lambda,\gamma} \gamma \int \tfrac{1+t}{2} \tanh\left(z\sqrt{\gamma\eta} + \gamma\eta + \tfrac{\lambda}{2}\right) + \tfrac{1-t}{2} \tanh\left(z\sqrt{\gamma\eta} + \gamma\eta - \tfrac{\lambda}{2}\right) Dz \right] \\ &= 1 + \alpha \mathop{\mathrm{E}}_{t,\gamma} \left[ \gamma\left(1-t^2\right) \int \frac{1 - \tanh\left(z\sqrt{\gamma\eta} + \gamma\eta\right)}{1 - t^2 \tanh^2\left(z\sqrt{\gamma\eta} + \gamma\eta\right)} Dz \right]. \end{aligned} \tag{22}$$

In terms of multiuser efficiency, the free energy can be further simplified to read

$$\begin{aligned}\mathcal{F} &= \mathop{\mathrm{E}}_{t,\gamma} \int \frac{1+t}{2} \log\left[\left(1-t^2\right)\cosh^2\left(z\sqrt{\gamma\eta} + \gamma\eta\right) + t^2\right] \\ &\quad + t \log\left[\cosh\left(z\sqrt{\gamma\eta} + \gamma\eta\right) + t \sinh\left(z\sqrt{\gamma\eta} + \gamma\eta\right)\right] Dz - \gamma\eta + \frac{\log(\eta) - \eta}{2\alpha}. \end{aligned} \tag{23}$$

The fixed-point equation (22) may have either one or three positive solutions. When there are three solutions, all three solutions are extremum points of the free energy. One of them is the global maximum of the free energy. One is a local maximum of the free energy, and one is a local minimum of the free energy. Two of these three solutions can be found by iterating the fixed-point equation (22) setting the initial value for the multiuser efficiency to 0 and 1, respectively.

One of these two solutions globally maximizes the free energy (23). It is the multiuser efficiency of the individually optimum multiuser detector (4) which requires the summation of at least $2^{K-1}$ terms. It is proven that this task cannot be performed in polynomial time [16].

The solution corresponding to initial value $\eta_0 = 0$ is the multiuser efficiency of a non-linear gradient based iterative search approximating the np-complete detector (4). It may be different from the solution maximizing the free energy globally. In such cases it gives smaller multiuser efficiency than the solution obtained for initial value $\eta_0 = 1$. In exchange for this, it can be achieved with polynomial time complexity using Stochastic Hopfield Neural Networks (SHNN).

All three solutions coincide for a wide range of parameters (load, noise variance, prior distribution). In this case, the performances of the gradient search-based detector with polynomial complexity and the np-complete individually optimum detector (4) coincide, too. At first glance, this seems to contradict the np-completeness of the individually optimum detector. Note, however, that previous considerations hold for an infinite user population only.

# 5 Density evolution and Gaussian approximation

For finite number of users, the messages $\mathcal{L}_{k,n}^{\mathrm{mud}}$ and $\mathcal{L}_{k,\ell}^{\mathrm{dec}}$ are random variables whose joint pdf is induced by the joint probability measure of the users information bits, of the channel noise, of the users' spreading sequences, and of the random interleavers of the user codes.

The DE approach to the analysis of message-passing iterative decoding algorithms consists of propagating from iteration to iteration the pdf of the messages [12]. The bit-error probability performance of the decoder can be derived by the limiting pdf of the messages after a large number of iterations. Under mild conditions, as $N \to \infty$ a general concentration result [12] ensures that the messages arriving at each node are mutually statistically independent, and their

marginal pdfs converge in probability to the marginal pdfs computed on a cycle-free average graph, where in our case "averaging" is with respect to the graph structure defined by the bit-interleavers, the random information bits and the channel noise.

In order to remove the randomness due to the random selection of the users spreading sequences, and information messages, we study the IO-MUD in the large-system limit and make use of the self-averaging property of the free energy. (Notice the order of the limits: first we let $N \to \infty$ and then $L, K \to \infty$). We assume that the users are grouped into a finite number $J$ of classes of cardinality $K_1, \ldots, K_J$, where $K = \sum_{j=1}^{J} K_j$, with received SNR levels $g_1, \ldots, g_J$, i.e., $\gamma_k = g_j$ if user $k$ belongs to class $j$, and we assume that the ratio $\alpha_j = K_j/L$ remains fixed for all $j$, as $L \to \infty$.

The prior distribution of the information bit of user $k$ at time instant $n$ is uniquely characterized by the parameter $t_{k,n} = \tanh(\mathcal{L}_{k,n}^{\mathrm{dec}}/2)$. By following in the footsteps of [2] we can show that, at any iteration $m$, the empirical distribution of the $t_{k,n}$'s over all users $k$ in class $j$ converges to a given deterministic distribution $F_j^{(m)}(t)$, as $K \to \infty$. Thus, the fixed point equation for the multiuser efficiency (22) at decoder iteration $m$ can be expressed as

$$\frac{1}{\eta^{(m)}} = 1 + \sum_{j=1}^{J} \alpha_j g_j \int\int (1-t^2) \, \frac{1 - \tanh\left(z\sqrt{g_j\eta^{(m)}} + g_j\eta^{(m)}\right)}{1 - t^2 \tanh^2\left(z\sqrt{g_j\eta^{(m)}} + g_j\eta^{(m)}\right)} \, Dz \, dF_j^{(m)}(t). \quad (24)$$

Thus, the conditional distribution of $\mathcal{L}_{k,n}^{\mathrm{mud}}$ given $c_{k,n} = 0$ is $\mathcal{N}(2g_j\eta^{(m)}, 4g_j\eta^{(m)})$, for user $k$ in class $j$, in the large system limit. Hence, the DE is completely expressed by the evolution of the single parameter $\eta^{(m)}$, for $m = 0, 1, 2, \ldots$.

For general linear convolutional codes, the SISO decoder is too complicated to compute the pdf of $\mathcal{L}_{k,n}^{\mathrm{dec}}$ from the pdf of $\mathcal{L}_{k,n}^{\mathrm{mud}}$ in closed form. A semianalytic technique to the DE consists of approximating the pdf of $\mathcal{L}_{k,n}^{\mathrm{dec}}$ by a Monte Carlo generated histogram, obtained directly by the forward-backward algorithm applied to randomly generated i.i.d. input log-ratios $\mathcal{L}_{k,n}^{\mathrm{mud}} \sim \mathcal{N}(2g_j\eta^{(m)}, 4g_j\eta^{(m)})$. A simpler approach consists of a Gaussian Approximation (GA) of the SISO output messages [5, 3, 2]. Here, we make use of the "Gaussian tail matching" approximation of [2]. Let $\epsilon$ denote the symbol-error rate (SER) at the SISO decoder output, given by [1] $\epsilon = \Pr\left(\mathrm{sign}(\mathcal{L}_{k,n}^{\mathrm{dec}}) \neq c_{k,n}\right)$. Assuming $\mathcal{L}_{k,n}^{\mathrm{dec}} \sim \mathcal{N}(2\mu, 4\mu)$ then $\epsilon = Q(\sqrt{\mu/2})$. For a given convolutional code over AWGN, the SER $\epsilon$ is a known function of the decoder input SNR, which in our case is given by $g_j\eta^{(m)}$ for a user in class $j$ at iteration $m$. Hence, the pdf of the log-ratios at the output of its SISO decoder (under the GA assumption) is uniquely identified by the single parameter

$$\mu_j\left(\eta^{(m)}\right) \triangleq \left[Q^{-1}(\epsilon(g_j\eta^{(m)}))\right]^2 \quad (25)$$

By putting (24) and (25) together, we can express the full DE-GA by a one-dimensional dynamic system in the form $\eta^{(m+1)} = \Psi(\mathbf{g}, \boldsymbol{\alpha}, \eta^{(m)})$, with initial condition $\eta^{(0)} = 0$, where the mapping function $\Psi(\mathbf{g}, \boldsymbol{\alpha}, \eta)$, is obtained implicitly by solving the equation

$$\frac{1}{\Psi} = 1 + \sum_{j=1}^{J} \alpha_j g_j \int_{\mathbb{R}^2} \frac{\left(1 - \tanh^2\left(y\sqrt{\mu_j(\eta)} + \mu_j(\eta)\right)\right)\left(1 - \tanh\left(z\sqrt{g_j\Psi} + g_j\Psi\right)\right)}{1 - \tanh^2\left(y\sqrt{\mu_j(\eta)} + \mu_j(\eta)\right)\tanh^2\left(z\sqrt{g_j\Psi} + g_j\Psi\right)} \, Dz \, Dy. \quad (26)$$

for all $\eta \in [0, 1]$.

[1] Notice that SER refers to decisions based on extrinsic information, not on a posteriori probabilities of the SISO decoders.

# 6 Optimal received power distribution

Within the limits of the assumptions made in order to obtain the DE-GA, the decoder performance is completely characterized by the fixed points of the system defined by the mapping function $\Psi(\mathbf{g}, \boldsymbol{\alpha}, \eta)$. This is continuous and non-decreasing in $\eta$, with $\Psi(\mathbf{g}, \boldsymbol{\alpha}, 0) > 0$ and $\Psi(\mathbf{g}, \boldsymbol{\alpha}, 1) \leq 1$. Then, the trajectory with initial condition $\eta^{(0)} = 0$ converges to the fixed point given by the smallest solution of the equation

$$\Psi(\mathbf{g}, \boldsymbol{\alpha}, \eta) = \eta, \quad \eta \in [0, 1] \tag{27}$$

Next, we optimize the system spectral efficiency with respect to the received power distribution, defined by $(\mathbf{g}, \boldsymbol{\alpha})$. We fix a target maximum BER, to be achieved by all users in the system. This implies that for all users, after the iterative decoder has converged to a stationary point, the SINR at the SISO decoder inputs must be not smaller than a given threshold value $\mathrm{SINR}_{\mathrm{th}}$, which depends on the code and on the target BER. We discretize the SNR values such that $g_1 < g_2 < \cdots < g_J$, for some integer $J$, we select a desired channel load $\alpha$, a *constraint interval*[2] $[\delta_1, \delta_2] \subseteq [0, 1)$ and a margin[3] $\varepsilon > 0$. Then, we solve the following constrained optimization problem with respect to $\boldsymbol{\alpha}$

$$\text{minimize} \quad \sum_{j=1}^{J} \alpha_j g_j \qquad \text{subject to} \begin{cases} \Psi(\mathbf{g}, \boldsymbol{\alpha}, \eta) \geq \eta + \varepsilon, & \forall \ \ \eta \in [\delta_1, \delta_2] \\ \sum_{j=1}^{J} \alpha_j = \alpha, & \\ \alpha_j \geq 0, & \forall \ \ j \end{cases} \tag{28}$$

The solution $\boldsymbol{\alpha}^\star$ can be accepted if the fixed point $\eta^\star$, i.e., the smallest solution of the equation $\Psi(\mathbf{g}, \boldsymbol{\alpha}^\star, \eta) = \eta$ for $\eta \in [0, 1]$, is such that $g_1 \eta^\star \geq \mathrm{SINR}_{\mathrm{th}}$. Otherwise, the program is run again by changing the SNR values $\mathbf{g}$ and the design parameters $\delta_1, \delta_2$.

The program (28) is linear for the SHNN solution of the fixed point equation for the multiuser efficiency (22), as we show in the following. The objective function, the non-negativity constraint and the equality constraint are obviously linear. The inequality constraint $\Psi(\mathbf{g}, \boldsymbol{\alpha}, \eta) \geq \eta + \varepsilon, \quad \forall \ \ \eta \in [\delta_1, \delta_2]$ is also linear in $\boldsymbol{\alpha}$.

Though there is no explicit expression for $\Psi(\mathbf{g}, \boldsymbol{\alpha}, \eta)$, we can use the same argument as for post-IC linear MMSE filtering in [4]. Let $\Psi = G_\eta(\Psi)$ be the fixed point equation (26) yielding $\Psi$ as a function of $\eta$, and let $\Psi^\star$ be the smallest non-negative solution—this is the solution corresponding to the SHNN solution—for $\eta$ fixed. Then, the following implication holds:

$$\Psi \leq \Psi^\star \iff \Psi \leq G_\eta(\Psi). \tag{29}$$

Hence, we conclude that the inequality $\Psi^\star \geq \eta$ is equivalent to $\eta \leq G_\eta(\eta)$, which yields the linear constraint

$$\sum_{j=1}^{J} \alpha_j g_j \int_{\mathbb{R}^2} \frac{\Big(1 - \tanh^2\big(y\sqrt{\mu_j(\eta)} + \mu_j(\eta)\big)\Big)\Big(1 - \tanh\big(z\sqrt{g_j\eta} + g_j\eta\big)\Big)}{1 - \tanh^2\big(y\sqrt{\mu_j(\eta)} + \mu_j(\eta)\big)\tanh^2\big(z\sqrt{g_j\eta} + g_j\eta\big)} \, Dz\, Dy \leq \frac{1}{\eta + \varepsilon} - 1. \tag{30}$$

for $\eta \in [\delta_1, \delta_2]$ to be used in (28). However, if the free energy does not favor the smallest solution of the fixed point equation (22), we find that (29) does not hold. Fortunately, the free energy favors the SHNN solution for convolutional codes unless ones demands for extremely low bit error rates (typically $10^{-10}$ or below) with weak codes.

[2]The contraint interval should be a subset of the unit interval since it is sufficient for the fixed-point multiuser efficiency to be close to 1. Forcing it to approach unity arbitrarily close would result in a huge degradation in power efficiency. It is also reasonable to exclude 0 from the contraint interval for numerical reasons.

[3]The margin $\varepsilon$ is trading number of iterations against power efficiency.

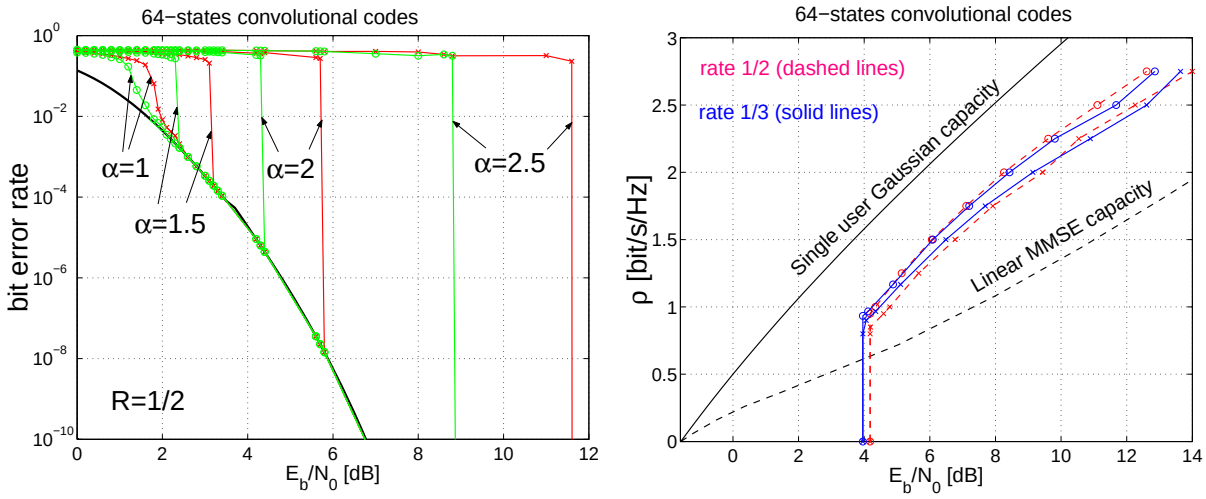


Figure 1: Bit error rate for constant power profile (left hand side) and spectral efficiency for optimized power profile (right hand side) of iterative multiuser decoding. IO-MUD and MMSE-IC are marked by circles and crosses, respectively. The power optimization uses the parameter $\varepsilon = 10^{-2}$. The target bit error rate for plotting the spectral efficiency is $10^{-5}$.

# 7 Results

First, we compare the performance of iterative multiuser decoding employing IO-MUD to MMSE-IC without power profile optimization. The respective bit error rates for rate $R = \frac{1}{2}$ convolutional codes with 64 states are depicted on the left hand side of Fig. 1 for several system loads $\alpha$. It can be observed that for increasing load the gap widens. That is MMSE-IC becomes more and more suboptimum. At a moderate load of $\alpha = 2.5$ corresponding to a spectral efficiency of $\rho = 1.25$ bit/s/Hz, the penalty due to MMSE-IC is 2.75 dB compared to the exact belief propagation algorithm.

It is worth to remark that the figure remains unchanged, if the IO-MUD is replaced by the SHNN solution, since phase transitions occur only for those signal-to-noise ratios which lead to bit error rates below $10^{-10}$.

The significant gap between MMSE-IC and IO-MUD becomes considerably small, when the power profile of the users is optimized. This is illustrated on the right hand side of Figure 1. For the same half rate convolutional code with 64 states leading to 2.75 dB loss without power profile optimization, a loss of only 0.5 dB occurs after power optimization. For large spectral efficiency the gap slightly increases, but still remains small. No phase transitions could be found for the optimized power profiles at the target bit error rate of $10^{-5}$.

The narrowing gap for optimized power profile is quite intutive having in mind the result of reference [15] which shows that for infinite length random codes and optimized power profile among the users MMSE-IC incurs no suboptimality. Thus, we can expect that for stronger codes the gap continues to shrink.

# 8 Summary and conclusions

A fixed point equation for the large system multiuser efficiency of the IO-MUD with non-uniform priors has been derived. In general, the fixed point equation has multiple solutions. For the smallest solution, power profile optimization has been shown to be a linear program. For a wide range of practical parameter settings in iterative multiuser decoding, the fixed point

equation has a unique solution and phase transitions do not occur.

For optimized power profile, the gap between exact belief propagation and the approximation of the IO-MUD by MMSE-IC is considerably small. Thus, there is only little left to be gained in performance due to multiuser detection algorithms with higher complexity than MMSE-IC, e.g. sphere decoding based algorithms as proposed in [17, 8].

# References


[1] Special issue on iterative decoding. *IEEE Trans. Inform. Theory*, 47(2), Feb. 2001.

[2] J. Boutros and G. Caire. Iterative multiuser joint decoding: Unified framework and asymptotic analysis. *IEEE Trans. Inform. Theory*, 48(7):1772–1793, Jul. 2002.

[3] S. ten Brink. Convergence of iterative decoding. *Elect. Letters*, 35(13):1117–1119, Jun. 1999.

[4] G. Caire and R. R. Müller. The optimal received power distribution for IC–based iterative multiuser joint decoders. In *Proc. 39th Allerton Conf. Commun., Contr. & Comp.*, Oct. 2001.

[5] S. Y. Chung, T. Richardson, and R. Urbanke. Analysis of sum-product decoding of low-density parity-check codes using a Gaussian approximation. *IEEE Trans. Inform. Theory*, 47(2):657–670, Feb. 2001.

[6] T. R. Giallorenzi and S. G. Wilson. Multiuser ML sequence estimator for convolutionally coded asynchronous DS–CDMA systems. *IEEE Trans. Commun.*, 44(8):997–1008, Aug. 1996.

[7] D. Guo and S. Verdú. Multiuser detection and statistical mechanics. In *Ian Blake Festschrift*, 2002.

[8] B. Hochwald and S. ten Brink. List sphere decoding to attain capacity on a multi-antenna link. In *Proc. 39th Allerton Conf. Commun., Contr. & Comp.*, Oct. 2001.

[9] F. Kschischang, B. Frey, and H.-A. Loeliger. Factor graphs and the sum-product algorithm. *IEEE Trans. Inform. Theory*, 47(2):498–519, Feb. 2001.

[10] H. Nishimori. *Statistical Physics of Spin Glasses and Information Processing*. Oxford University Press, Oxford, U.K., 2001.

[11] J. Pearl. *Probabilistic reasoning in intelligent systems: networks of plausible inference*. Morgan Kaufmann, San Francisco, 1988.

[12] T. Richardson and R. Urbanke. The capacity of low-density parity check codes under message passing decoding. *IEEE Trans. Inform. Theory*, 47(2):599–618, Feb. 2001.

[13] T. Tanaka. A statistical mechanics approach to large-system analysis of CDMA multiuser detectors. *IEEE Trans. Inform. Theory*, to appear.

[14] D. Tse and S. Hanly. Linear multiuser receivers: Effective interference, effective bandwidth and user capacity. *IEEE Trans. Inform. Theory*, 45(2):641–657, Mar. 1999.

[15] M. K. Varanasi and T. Guess. Optimum decision feedback multiuser equalization with successive decoding achieves the total capacity of the Gaussian multiple-access channel. In *Proc. Asilomar Conf. Signals, Syst. & Comp.*, Nov. 1997.

[16] S. Verdú. *Multiuser Detection*. Cambridge University Press, New York, 1998.

[17] H. Vikalo and B. Hassibi. Low-complexity iterative decoding over multiple antenna channels via a modified sphere decoder. In *Proc. 39th Allerton Conf. Commun., Contr. & Comp.*, Oct. 2001.

[18] E. Viterbo and J. Boutros. A universal lattice code decoder for fading channels. *IEEE Trans. Inform. Theory*, 45(6):1639–1642, Jul. 1999.

[19] X. Wang and H. V. Poor. Iterative (turbo) soft interference cancellation and decoding for coded CDMA. *IEEE Trans. Commun.*, 47(7):1046–1061, Jul. 1999.