2000 Conference on Information Sciences and Systems, Princeton University, March 15-17, 2000

Combining Multiuser Detection with Coding: Promises and Problems

Ralf R. Müller[1]
Department of Electrical Engineering
Princeton University
Engineering Quadrangle
Princeton, NJ 08544
e-mail: rmueller@ee.princeton.edu

Abstract — **Though proven optimum for infinite code word length in [1], linear MMSE detection for CDMA followed by a successive decoding chain is shown to degrade severely in comparison to a joint coding approach for short codes. Independent decoders are demonstrated to be incapable to cope with multiuser interference in a near–optimum way under reasonable system assumptions.**

I. Introduction

In the early days of information theory, Shannon raised the question whether source and channel coding are separable without loss in performance and proved they were. However, his proof assumed codewords of infinite length. In practice, joint source and channel coding schemes turned out to exhibit superior performances than separated approaches [2, p. 534].

Varanasi and Guess [1] proved that multiuser detection and coding are also separable with respect to the achievability of the vertices of the capacity region of the Gaussian multiple–access channel with correlated waveforms (GMACCW). As their result for the vertices can be generalized via time sharing or rate–splitting [3, 4] to hold for the whole capacity region, it seems a natural question how far this result extends to practical coding schemes that do not use codewords of infinite lengths.

Preliminaries on separation between multiuser detection and coding are given in Section II of this paper. In Section III, the separation between multiuser detection and coding proposed by Varanasi and Guess is found to show severe drawbacks in performance in comparison to joint multiuser decoding when the codeword length is limited, in general, and when the number of users is larger than the spreading gain, in particular.

Section IV considers a more pragmatic split between multiuser detection and coding, i.e. a maximum likelihood (ML) detector without knowledge about the code laws followed by independent single–user decoders. Conclusions from these considerations are pointed out in Section V.

II. Preliminaries

Consider a synchronous Gaussian multiple–access channel with correlated waveforms and K users as depicted in Fig. 1. It is well–known [5] that a filter bank matched to the K waveforms $s_k(t), 1 \leq k \leq K$, provides a sufficient discrete–time statistics $\boldsymbol{r}[\nu]$ for the signal $y(t)$. As long as the users' individual information rates lie within the capacity region of the GMACCW there is a coding schemes that enables perfect reconstruction of the transmitted information $x_k[\mu], 1 \leq k \leq K$, from the knowledge of the vector series $\boldsymbol{r}[\nu]$.

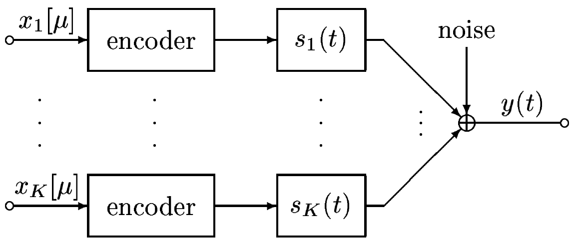

Fig. 1: GMACCW with K users.

With respect to the data processing lemma, it is not obvious whether this reconstruction can be handled by the structure in Fig. 2. Hereby, the "multiuser detector" is allowed to be any algorithm that has no knowledge about the code laws while the independent decoders are assumed to have no knowledge about the users' waveforms. However, as far as channel capacity is addressed the codewords are finite in length. Thus, the decoders could perfectly adapt to all the users' waveforms if they were provided the signal $\boldsymbol{r}[\nu]$. On the other hand, the "multiuser detector" is conjectured to not be able to learn the code laws as almost all of the allowed codewords are never transmitted. Nevertheless, it follows that the structure in Fig. 2 does indeed incur no loss of generality if the "multiuser detector" is able to provide all the information contained in $\boldsymbol{r}[\nu]$ to each of the decoders. This, however, is a trivial task: It just needs to quantize the vector signal $\boldsymbol{r}[\nu]$ with an arbitrary average error measure smaller than ϵ and multiplex the quantized signals in amplitude. As $\epsilon \to 0$, each of the decoders can reconstruct the vector signal $\boldsymbol{r}[\nu]$ with sufficient precision and then perform joint multiuser decoding.

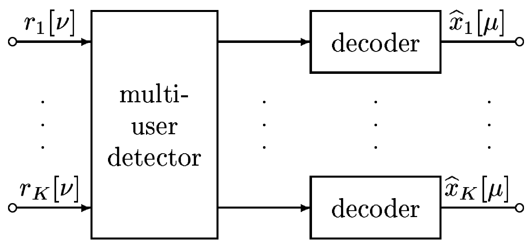

Fig. 2: Separation of multiuser detection and decoding.

It should be clear from these considerations that the simple definition of a multiuser detector and a decoder as general machines that simply do not know about the code laws and

[1]This work was supported in part by the German Academic Exchange service (DAAD) under grant 332 4 00 510.

the users' waveforms, respectively, is not very helpful. Instead, more pragmatic and useful approaches will be considered within the next two sections.

III. MMSE–SIC

As shown in [1], the multiuser detector can be a linear MMSE detector without loss of performance if the decoders make use of successive cancellation, such as depicted for two users in Fig. 3. Besides the need for perfect cancellation, this scheme is

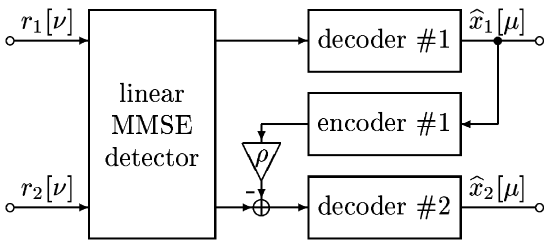

Fig. 3: Successive cancellation for two user with ρ denoting the correlation coefficient between the two users' waveforms.

affected by another drawback which also inherently arises from the cancellation procedure and here is called *power penalty propagation*. This effect is most easily explained for a simple successive cancellation scheme without spreading as used in [6] to find the capacity region of the Gaussian multiple–access channel.

Let a finite length channel code be characterized by the fact that it needs $V \geq 1$ times the signal power in order to achieve the same error probability as an infinite length random code. Using such a code in a chain of successive decoding involves the following trouble: While the last user in the chain is simply affected by its power penalty V in comparison to the infinite length random code, the last but one user has to cope with its own power penalty in addition to the unnecessary high interference resulting from the last user's power penalty. This deleterious effect is obviously the worse the more users are involved.

It is straightforward to show that the total required signal–to–noise ratio when each user gets an identical rate share Γ/K from the total rate Γ is given by

$$\frac{E_{\rm s}}{N_0} = \frac{1}{2}\left(\left(V\left(4^{\Gamma/K}-1\right)+1\right)^K - 1\right) \qquad (1)$$

with the particular cases:

$$K = 1: \qquad \frac{E_{\rm s}}{N_0} = V \cdot \frac{4^{\Gamma}-1}{2} \qquad (2)$$

$$K \to \infty: \qquad \frac{E_{\rm s}}{N_0} = \frac{4^{V\Gamma}-1}{2} \qquad (3)$$

Obviously, the power penalty is no longer a linear factor like in the single user case (2), but becomes even exponentially involved in (3) as the number of users grows large. The multiuser power penalty $V_{\rm MPP}$ is therefore larger than the single user power penalty V. Moreover, it is depended on the number of users. Some examples for this power penalty propagation are depicted in Fig. 4 where signal–to–noise ratio is normalized by $E_{\rm b} = KE_{\rm s}/\Gamma$. Note that this effect causes severe degradation when the number of users and the power penalty are not very small. Power penalty propagation is also an obstacle when

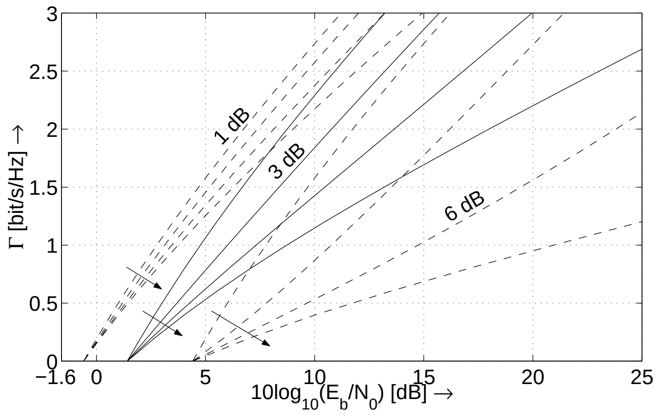

Fig. 4: Spectral efficiencies for several numbers of users $K = 1, 2, 5, \infty$ (following arrows) and power penalties $10\log_{10}(V) = 1, 3, 6$ dB.

rate–splitting [3] shall be used in order to extend achievability of the vertices of the capacity region to any point on its dominant face via successive decoding, as this method almost doubles the effective number of users.

For CDMA with correlated waveforms a similar effect occurs, but it is more difficult to analyze. The power penalty propagation depends on the crosscorrelations between the waveforms and the sorting of the users. Even the question which sorting is best is a difficult task and has not been solved according to the knowledge of the author. Simplifications are possible for random signature waveforms in the asymptotic regime if spreading factor N and number of users K grow to infinity with a fixed ratio ζ called *load*. In that case, results found in [7] for linear MMSE detection can be easily generalized to apply for MMSE detection with successive decoding. In that way, reference [8] showed that the demand for equal rates for all users does not cause significant degradation for large random spreading systems even if methods like rate–splitting or time–sharing that introduce virtual users are not taken into account.

Following [8], but including the power penalty into the considerations, spectral efficiency $\Gamma = K/N \cdot C$ can be found as a function of $E_{\rm b}/N_0$ solving an integral equation, numerically, (see [8] for details). The obtained results strongly depend on the load $\zeta = K/N$ of the system which also defines the spreading–coding tradeoff for a fixed number of users. While for $V = 1$, i.e. no power penalty, infinite load results in highest spectral efficiency, $V > 1$ yields a finite optimum for the load. For large signal–to–noise ratios, the optimum load converges towards $\zeta = 1$. Spectral efficiency for optimized load and several power penalties is depicted in Fig. 5 for the asymptotic limit of $K = \zeta N \to \infty$. In contrast to Fig. 4, the slope of the curves does not flatten out. In order to understand this, note that a lower bound for spectral efficiency of MMSE–SIC is given by the spectral efficiency of a decorrelating detector followed by SIC and the decorrelating detector is not sensitive to the interfering users' powers. The limit in spectral efficiency of MMSE–SIC as $V \to \infty$ is actually identical to those of decorrelating SIC. It is also remarkable that the MMSE detector as front–end shows significant improvement in spectral efficiency over the decorrelating front–end only if the power penalty is smaller than about 2 to 3 dB. This result is particu-

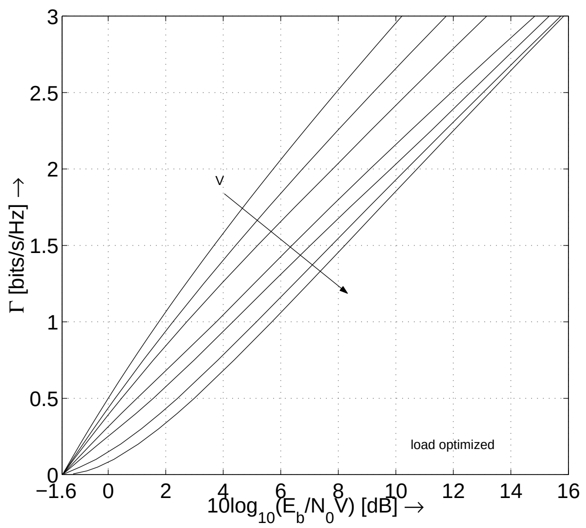

Fig. 5: Spectral efficiency of CDMA with random spreading for several power penalties $10\log_{10}(V) = 0, 0.5, 1, 2, 3, 6, \infty$ dB (following arrow) and optimized load as function of signal–to–noise ratio normalized to the respective power penalty.

larly interesting from the viewpoint of implementation, as the decorrelator does not need knowledge about the interfering users' powers.

The loss due to power penalty propagation can become arbitrary large for fixed load [9], but saturates if comparison is based on optimized load. This can be observed from Fig. 5 where decorrelating SIC serves as a lower bound on performance. It is shown more clearly in Fig. 6. Particularly for

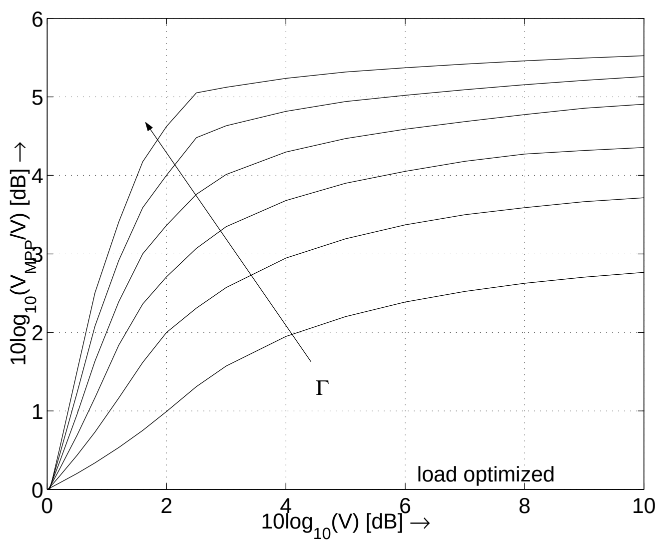

Fig. 6: Additional power penalty as function of single–user power penalty for random spreading, $K \to \infty$, and several spectral efficiencies $\Gamma = 0.5, 1.0, 1.5, 2.0, 2.5, 3.0$ bits/s/Hz (following arrow).

large spectral efficiencies, the single user power penalty V needs to fall below a certain threshhold of about 2 to 3 dB in order to combat power penalty propagation effectively.

Introducing the power penalty in order to model the influence of finite length coding onto the performance of MMSE–SIC, linear detection followed by SIC was shown to suffer from severe drawbacks when used as a means to separate multiuser coding into multiuser detection and single–user coding.

IV. ML–Detection

In Section III, the decoders were allowed to communicate through the re–encoded feedback signals originating from the previous decoders. This provided at least theoretically optimum performance for any choice of the signature waveforms. However, each decoder has to wait for its processor's output. This prohibits parallel implementations, in principle.

In contrast, let now be considered a structure like it is depicted in Fig. 2 where all decoders operate independently from each other. In order to overcome the trouble with this structure outlined in Section II, the decoders are assumed to be single–user decoders which do not even know about the presence of other interfering users and have no capabilities to learn anything from their input signals other than the data of the user they are supposed to decode.

With these preliminaries on the decoders, the multiuser detector has to recover the coded signal of each user as closely as possible without knowledge about the code laws. With the assumption that the decoder output symbols appear to be equiprobable[1] random choices from the code's output alphabet, the multiuser detector can apply the *maximum likelihood* decision rule. However, it should also provide soft information [10] to the decoder, i.e. an estimate on the probability that the taken decision was correct. In the following, both hard and soft output ML detectors are addressed.

Although those detectors involve exponential complexity and are unlikely to be actually implemented in a variety of applications, their performance serves as an upper bound on the potentials of suboptimum structures that approximate either soft or hard ML detection. This is particularly interesting for nonlinear iterative multiuser detectors, e.g. [11, 12, 13], as their performance is known in terms of uncoded bit error rate only which does not straightforwardly relate to performance when embedded in systems with outer error correction coding which is able to benefit from reliability information.

Spectral efficiency for random waveforms is the most reasonable and accepted measure of performance whenever performance of coded multiuser detection is addressed. However, analytical results like those found in [14, 15, 9] are hard to obtain in the present case. Analytical results on the performance of ML detectors have been an open problem in literature for a long time. Bounds on the uncoded error probability for given waveforms are due to Verdú [5]. These bounds are tight for low noise variance. For random spreading Tse and Verdú [16] recently showed the following result: if the number of users as well as the spreading gain grow to infinity with a fixed ratio ζ, the asymptotic error probability for vanishing noise variance approaches that of a single user communication channel regardless of the ratio ζ. Neither the latter result nor the bounds on the error probability are useful for the targeted problem, as in combination with coding, one wants to operate at low signal–to–ratios, i.e. high uncoded error probability. Therefore, computer simulations (which may also be called Monte–Carlo integration) are used to obtain spectral efficiency.

[1] Note that it is not clear whether equiprobable signaling is optimum in the considered situation.

For binary modulation with hard ML detection, one can obviously simulate the bit error probability $\overline{P}_{\rm b}$ averaged over various realizations of the spreading sequences and plug it into the formula for the capacity of a binary symmetric channel (BSC)

$$C_{\rm BSC} = 1 - {\rm e}_2\left(\overline{P}_{\rm b}\right) \qquad (4)$$

with ${\rm e}_2(x) \triangleq -x\log_2(x) - (1-x)\log_2(1-x)$ denoting the binary entropy function.

In order to evaluate the performance of soft ML detection a slightly more subtle approach is needed which, as a byproduct, also gives the result for hard detection: Let each user transmit a single bit with an instant realization of random waveforms and noise. Then, one can calculate the probabilities for the 2^K hypotheses the detector has to chose from conditioned on the received signal. By summation, these 2^K probabilities yield the K marginal probabilities for the K hypotheses on the K users' symbols. It remains to average the capacity of the BSC over these probabilities and other realizations of the waveforms and noise. This gives

$$C_{\rm soft} = 1 - \underset{\boldsymbol{r}}{\mathsf{E}}\left\{{\rm e}_2\left(P_{\rm b}|\boldsymbol{r}\right)\right\}. \qquad (5)$$

Comparing with (4), the only difference for hard detection is that the probabilities are averaged[2] first and then capacity of BSC is calculated at the mean probability. The loss due to hard detection is hereby identified as a result of the convexity of (4) and Jensen's inequality. This method can easily be generalized to non–binary alphabets.

As the method outlined above becomes exhaustive when many users are involved, in the following the number of users is limited to $K = 12$. For such a small number of users, the probability that two or more users are assigned identical signature waveforms is not negligible if binary spreading sequences are applied. If Gaussian random sequences were used, all signals would be subject to some implicit fading. Spherical random sequences [17] are used in order to avoid this deleterious effects.

Spectral efficiency $\Gamma = K/N \cdot C$ for transmission over the GMACCW is plotted in Fig. 7 versus the normalized signal–to–noise ratio $E_{\rm b}/N_0$ for $K = 12$ users and various spreading factors. In contrast to conventional detection [18] and joint multiuser decoding [15], the tradeoff between spreading and coding is not at all trivial, but there is an optimum partitioning of the overall redundancy into spreading and coding like it was observed for the linear MMSE detector in [19, 20] and later[3] by bounds and fully analytic means in [22] and [15], respectively. Results for $N > K = 12$ are not shown as they keep behind the spectral efficiency of $K = N = 12$ for any signal–to–noise ratio. While linear multiuser detectors favor $K < N$, ML detectors seem to require $K \geq N$ in order to fully exploit their potentials.

Spectral efficiency in the presence of flat Rayleigh fading[4], is shown in Fig. 8 for the same parameters as before. Somewhat surprisingly, fading is beneficial for some signal–to–noise ratios and small N. This effect has also been observed for linear MMSE detection in [23]. It is due to the fact that fading

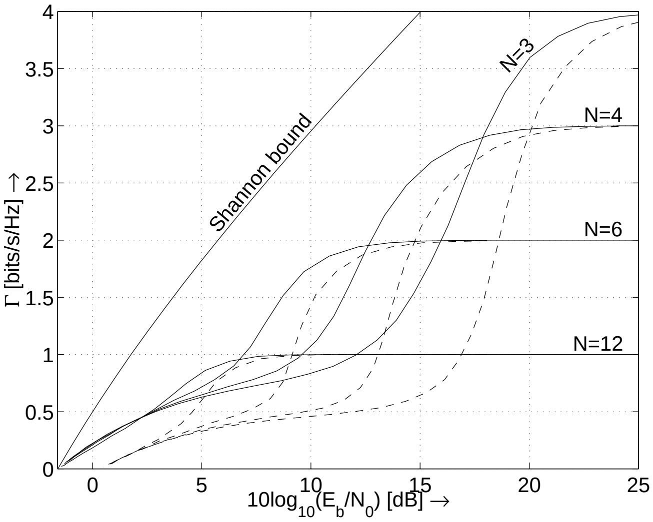

Fig. 7: Spectral efficiency of soft (solid lines) and hard (dashed lines) ML detection for $K = 12$ users.

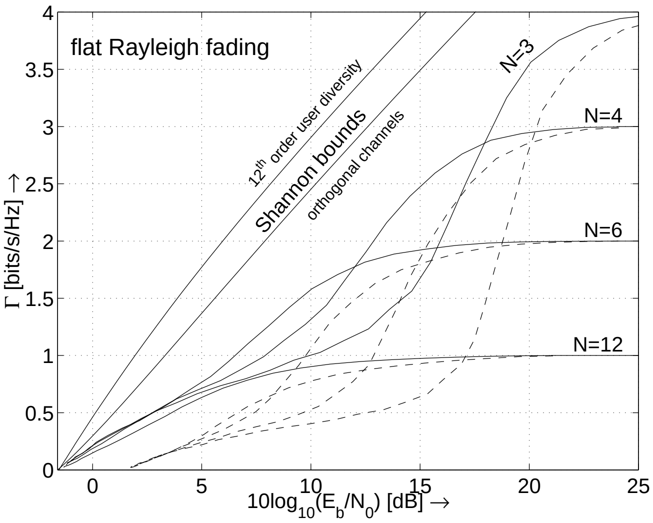

Fig. 8: Spectral efficiency of ML detection with flat Rayleigh fading, parameter as in Fig. 7.

creates a smaller number of significantly active users and may help if a smaller load is preferable at the given signal–to–noise ratio.

It is illuminating to compare the spectral efficiency of soft ML detection to those of linear MMSE detection and joint decoding which were previously reported in [20, 15] and [15], respectively, and extended to flat Rayleigh fading in [23]. For such a comparison, one has to take into account that the ML detector operates well for binary modulation, in particular, and for symbol alphabets that are finite sets, in general, while the MMSE detector as well as the joint decoder prefer continuously Gaussian distributed symbol alphabets. In the later case, there would be even no difference between linear MMSE and soft ML detection. Fig. 9 compares the detectors discussed above for unit load with $K = 12$ users with and without Rayleigh fading. As the ML detector was restricted to binary data modulation, its spectral efficiency is upper bounded by 1 bit/s/Hz. Moreover, there is little impact due to soft ML

[2] In this context, averaging means $\mathsf{E}\{\max\{P, 1-P\}\}$, as the decision device reverts its result if the probability of error exceeds one half. Note that this is irrelevant for soft decision because of the symmetry of the binary entropy function.

[3] For the correct date [20] was submitted, see [21].

[4] For ease of comparison to Fig. 7, the channel is assumed to fade in amplitude only, but not in phase.

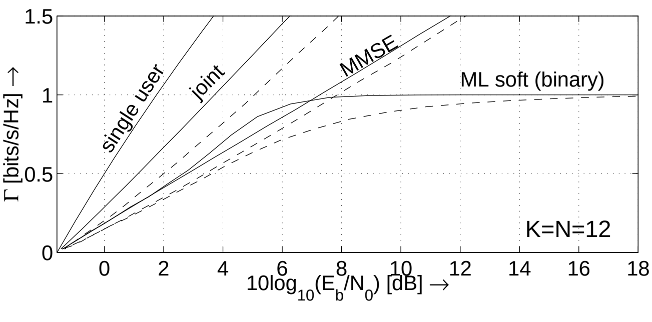

Fig. 9: Spectral efficiencies of several multiuser detectors with (dashed lines) and without (solid lines) flat Rayleigh fading.

detection when compared to linear MMSE detection, particularly for small signal–to–noise ratios. For systems with smaller load which were also studied, the impact of soft ML detection has been found to be even smaller. In the presence of flat Rayleigh fading soft ML detection with binary data modulation even degrades behind the linear MMSE detector because of the limitation to bear not more than a single bit per channel use. This is particularly malificious on fading channels, as there might be large potential with higher order data modulation when there are instantaneously good channel conditions. When compared with linear MMSE detection that has also been restricted to binary modulation, soft ML detection still realizes a small gain that would be visible when both curves were plotted in Fig. 9.

A large gain in spectral efficiency over linear MMSE detection is realized for loads bigger than unity and medium to large signal–to–noise ratios (not shown in figures). On the one hand, this is due to degradation caused by the interference limitation of the linear MMSE detector for $\zeta > 1+1/SIR$ where SIR denotes the targeted signal–to–interference ratio. On the other hand, the ML detector favors loads of 1 and more, cf. Figs. 7 and 8. A fair comparison, however, should address both detectors operating in their preferred interference environment, i.e. the load to be adjusted sensibly. Spectral efficiency for $K = 12$ users and optimized spreading factor N is depicted in Fig. 10 for soft ML and linear MMSE detection. While soft ML detection favors larger load the larger the signal–to–noise ratio, linear MMSE detection favors $N = 14$ to $N = 15$ for 10 dB $< 10\log_{10}(E_b/N_0) <$ 25 dB regardless whether the channel fades or not.

In general, the gain in spectral efficiency realized by soft ML detection in comparison to linear MMSE detection is surprisingly small. It grows bigger for large signal–to–noise ratios, particularly in the presence of fading. For soft ML detection the loss due to fading vanishes more and more as the spectral efficiency increases. This indicates that the soft ML detector is capable to benefit from *user diversity*. As shown in [24], user diversity provides the influence of fading onto the total capacity of multiple–access channels to vanish as the number of users grows towards infinity. For multiple–access channels with random spreading, user diversity occurs only if the number of users per spreading factor, i.e. the load, increases [23]. Thus, the influence of fading onto spectral efficiency of soft ML detection is only negligible for small spreading factors, as also observed in Fig. 10. Except for very low signal–to–noise ratios, linear MMSE is not capable to gain from user diversity due to the small load that is preferred by other more influential reasons.

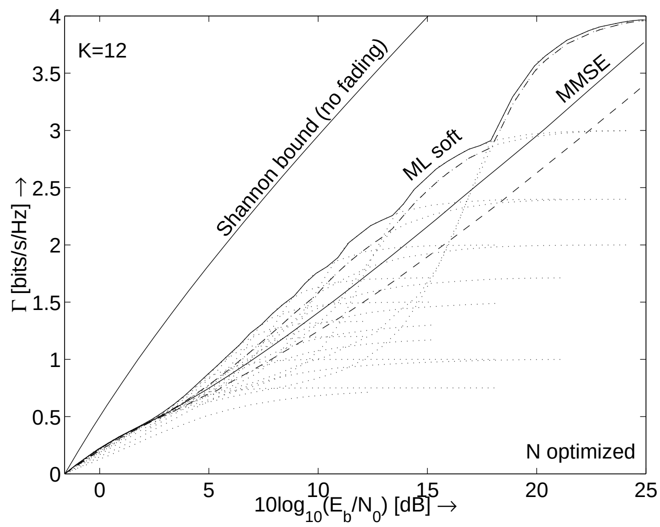

Fig. 10: Spectral efficiency of MMSE and soft ML detection with (dashed line) and without (solid line) flat Rayleigh fading for optimized spreading factor N.

V. Conclusions

The problem of multiuser coding cannot be broken into multiuser detection and *independent* single–user coding. Receiver structures like in Fig. 2 with independent single–user decoders suffer from severe drawback for any kind of multiuser detector.

The possible impact in spectral efficiency over linear MMSE detection by means of more sophisticated methods is surprisingly small and it is not clear whether one is willing to tolerate significant additional complexity in order to gain some fractions of it.

If the single–user decoders are not independent, e.g. they form a chain of successive decoding like in Fig. 3, the separation of multiuser detection and coding is enabled for codes of infinite lengths. For finite–length codes, however, linear MMSE detection followed by a successive decoding chain degrades significantly.

Acknowledgments

The author would like to thank Prof. Johannes Huber and the Telecommunications Laboratory of the University of Erlangen–Nuremberg, Germany, were parts of this work were performed.

References

[1] Mahesh K. Varanasi and Tommy Guess. Optimum decision feedback multiuser equalization with successive decoding achieves the total capacity of the Gaussian multiple–access channel. In *Proc. of Asilomar Conference on Signals, Systems and Computers*, pages 1405–1409, Monterey, CA, November 1997.

[2] Robert G. Gallager. *Information Theory and Reliable Communication.* John Wiley & Sons, New York, 1968.

[3] Bixio Rimoldi and Rüdiger Urbanke. A rate–splitting approach to the Gaussian multiple–access channel. *IEEE Transactions on Information Theory*, 42(2):364–375, March 1996.

[4] Alexander Lampe, Ralf R. Müller, and Johannes B. Huber. Transmit power allocation for Gaussian multiple access channels with diversity. In *Proc. of IEEE Information Theory Workshop (ITW)*, page 101, Kruger National Park, South Africa, June 1999.

[5] Sergio Verdú. Minimum probability of error for asynchronous Gaussian multiple–access channels. *IEEE Transactions on Information Theory*, IT–32(1):85–96, January 1986.

[6] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. John Wiley & Sons, New York, 1991.

[7] David Tse and Stephen Hanly. Linear multiuser receivers: Effective interference, effective bandwidth and user capacity. *IEEE Transactions on Information Theory*, 45(2):641–657, March 1999.

[8] Ralf R. Müller. Multiuser receivers for randomly spread signals: Fundamental limits with and without decision–feedback. *Submitted to IEEE Transactions on Information Theory*, June 1998.

[9] Ralf R. Müller. *Power and Bandwidth Efficiency of Multiuser Systems with Random Spreading*. Shaker–Verlag, Aachen, Germany, 1999.

[10] L.R. Bahl, J. Cocke, F. Jelinek, and J. Raviv. Optimal decoding of linear codes for minimizing symbol error rate. *IEEE Transactions on Information Theory*, IT–20(3):284–287, March 1974.

[11] T. Miyajima, T. Hasegawa, and M. Haneishi. On the multiuser detection using a neural network in code–division multiple–access. *IEICE Transactions on Communications*, E76–B:961–968, August 1993.

[12] George I. Kechriotis and Elias S. Manolakos. Hopfield neural network implementation of the optimal CDMA multiuser detector. *IEEE Transactions on Neural Networks*, 7(1):131–141, January 1996.

[13] Dariush Divsalar, Marvin K. Simon, and Dan Raphaeli. Improved parallel interference cancellation for CDMA. *IEEE Transactions on Communications*, 46(2):258–268, February 1998.

[14] Alex J. Grant and Paul D. Alexander. Random sequence multisets for synchronous code–division multiple–access channels. *IEEE Transactions on Information Theory*, 44(7):2832–2836, November 1998.

[15] Sergio Verdú and Shlomo Shamai (Shitz). Spectral efficiency of CDMA with random spreading. *IEEE Transactions on Information Theory*, 45(2):622–640, March 1999.

[16] David Tse and Sergio Verdú. Optimum multiuser asymptotic efficiency of CDMA with random spreading. In *IT Workshop on Detection, Estimation, Classification and Imaging*, Santa Fe, NM, February 1999.

[17] Sergio Verdú. *Multiuser Detection*. Cambridge University Press, New York, 1998.

[18] Andrew J. Viterbi. When not to spread spectrum — a sequel. *IEEE Communications Magazine*, 23(4):12–17, April 1985.

[19] Ralf Müller. Demodulation von CDMA–Signalen bei Modellierung der Teilnehmerinterferenzen als farbiges Rauschen. Master's thesis, Lehrstuhl für Nachrichtentechnik, Universität Erlangen–Nürnberg, Erlangen, Germany, January 1996.

[20] Peter Schramm and Ralf R. Müller. Spectral efficiency of CDMA systems with linear MMSE interference suppression. *IEEE Transactions on Communications*, 47(5):722–731, May 1999.

[21] Comments and Corrections. Erratum to "Spectral efficiency of CDMA systems with linear MMSE interference suppression“. *IEEE Transactions on Communications*, 47(9):1443, September 1999.

[22] A.G. Burr. Performance of linear separation of CDMA signals with FEC coding. In *Proc. of IEEE International Symposium on Information Theory (ISIT)*, page 354, Ulm, Germany, June/July 1997.

[23] Shlomo Shamai (Shitz) and Sergio Verdú. The effect of frequency–flat fading on the spectral efficiency of CDMA. *Submitted to IEEE Transactions on Information Theory*, 1999.

[24] Stephen V. Hanly and David N.C. Tse. Multi–access fading channels: Part II: Delay–limited capacities. *IEEE Transactions on Information Theory*, 44(7):2816–2831, November 1998.