Co-channel interference reduction in Rayleigh fading channels

Alex Grant

Institute for Telecommunications Research
University of South Australia

Abstract

Interference can be a significant problem in loosely organized networks such as ad-hoc radio networks. One technique for dealing with interference is to attempt joint detection of all signals arriving at a node. Iterative algorithms based on the turbo-principle yield significant performance improvements when applied to this problem. This paper concentrates on performance analysis of iterative interference cancelation for forward-error correction coded signals in the presence of Rayleigh fading. The variance transfer analysis of Alexander, Grant & Reed (Europ. Trans. Telecomm. 1998) is generalized for unequal user powers and codes. A surprising result is that the received power variation due to fading can assist convergence.

1 Introduction and System Model

The ultimate goal of an ad-hoc packet radio network is to provide flexible connectivity between terminals, in the absence of established network infrastructure. Ideally, global coordination of transmissions is minimized and data processing is performed in a distributed fashion. Mutual interference arising from simultaneous transmissions is a limiting factor for such networks. The results of [1] indicate that transport capacity at best scales only as the square-root of the network population, resulting in a vanishing per-node capacity. Later work [2] showed how linear scaling can be recovered through the use of multiuser coding strategies. One practical strategy is to use multiuser detection at each network node. This decentralized method yields useful capacity increases for random access channels [3]. Many multiuser detection algorithms have been reported in the literature. This paper focuses on iterative multiuser decoding using the soft interference canceler [4].

Email address: Alex.Grant@unisa.edu.au (Alex Grant).

1 This work was supported by the Australian Government under ARC Grant DP0344856.
and the main contribution is to extend the convergence analysis of [4] to cope with heterogeneous systems in which the various transmitters use different forward error correction codes and are received at different power levels. As a result, the effect of Rayleigh fading channels on receiver performance is determined.

Figure 1 shows a forward error correction coded linear multiple-access channel with \( K \) transmitters, or “users” sharing a common channel. User \( k = 1, 2, \ldots, K \) encodes their binary information sequence \( b_k[l] \) as follows. First, the sequence is encoded using a rate \( R_k \) code \( C_k \), producing a coded binary sequence \( d_k[l] \), \( l = 1, 2, \ldots, L \). Each coded sequence is permuted with an interleaver \( \pi_k \) and is subsequently mapped onto the antipodal constellation \( \{-1, +1\} \) via \( M \), giving sequences of modulated code symbols \( x_k[\pi_k(l)] = M(d_k[l]) \). At time \( i = 1, \ldots, L \), each user transmits the vector \( s_k[i]w_k[i]x_k[i] \), which is the multiplication of \( x_k[i] \) with an \( N \)-dimensional unit energy binary modulation waveform \( s_k[i] \in \{-1/\sqrt{N}, 1/\sqrt{N}\}^N \). In the case of direct-spread code-division multiple-access, \( N \) is often referred to as the processing gain. A useful analytical model adopted here is to let each element of \( s_k[i] \) be i.i.d. over users and time. The coefficients \( w_k \) represent the transmit amplitude and may also model multiplicative fading (although not strictly a function of the transmitter). Frequency selective fading will not be considered in this paper. For conceptual and notational clarity, users are symbol synchronous. More generally, \( s_k[i] \) may represent any complex vector modulation (including multiple antennas or multiple carriers). Extension to higher order constellations is also straightforward.

The baseband received vector \( r[i] \in \mathbb{R}^N \) is

\[
r[i] = S[i]W[i]x[i] + n[i],
\]

where \( S[i] = (s_1[i], s_2[i], \ldots, s_K[i]) \), is an \( N \times K \) matrix with the modulation sequence for user \( k \) as column \( k \) and \( W[i] = \text{diag}(w_1[i], \ldots, w_K[i]) \). The length \( K \) vector \( x[i] \) has elements \( x_k[i] \in \{-1, +1\} \) and \( n[i] \in \mathbb{R}^N \) is a sampled i.i.d. Gaussian noise process with covariance \( \sigma^2 I \). Define \( y[i] = S[i]^\dagger r[i] \) and \( \beta = K/N \). For compact notation, the time index will be suppressed resulting in \( r = SWx + n \), bearing in mind that this hides the action of the interleavers.

![Fig. 1. Linear multiple access channel.](image-url)
2 Iterative Interference Cancelation

Maximum likelihood (ML) decoding of forward error correction (FEC) coded linear multiple-access systems was first described in [5]. It was shown that brute force computation of the jointly ML codeword sequences for $K$ users has complexity $O(D^K\kappa)$ for $D$-ary modulation and constraint length $\kappa$ codes.

Figure 2 shows the “canonical” iterative multiuser decoder [6–9]. This decoder treats the FEC codes as an “outer code” and the interdependency introduced by the linear multiple access channel as an “inner code” (see Figure 1). The decoder iterates between a-posteriori probability (APP) computation for the inner code (ignoring the time-domain constraints due to FEC encoding) and individual APP decoding of each user’s FEC code (ignoring inter-user dependency). Remembering to keep track of re-orderings due to the interleavers, the APP channel decoder computes extrinsic marginal probabilities

$$q_k(d) = \frac{1}{p_k(d)} \left[ \alpha \sum_{d_i d_k = d} P(y|d_1, \ldots, d_K) \prod_{i=1}^{K} p_k(d_i) \right]$$  \hspace{1cm} (2)

where $\alpha$ normalizes the term inside the square brackets and $P(\cdot|\cdot)$ is the channel transition probability (incorporating modulation). The domain of summation is prohibitively complex, $O(D^K\kappa)$ for $D$-ary modulation.

One low complexity alternative is to replace the inner channel APP decoder with a linear filter. Figure 3 shows the general framework. First, convert from input probabilities to signal estimates via $\hat{x}_k = U(p_k(d))$. Typically $U$ computes conditional means according to the $p_k$. Second, apply an interference suppressing linear transformation $\tilde{x} = F(r, \hat{x})$, where $F$ may vary with time and iteration number. Finally, convert back to extrinsic probabilities $q_k(d) = T(\tilde{x}_k)/p_k(d)$, typically invoking a Gaussian assumption on $\tilde{x}$.

One of the simplest linear interference suppression methods is soft interference cancelation, introduced in [10,4]. In the framework of Figure 3, soft symbols are generated from the extrinsic probabilities,

$$\hat{x}_k = U(p_k(d)) = \sum_d M(d)p_k(d).$$  \hspace{1cm} (3)

Each user cancels the effect of the soft symbols from the received vector to produce an updated signal estimate

$$\tilde{x}_k = s_k^T \left( r - \sum_{i \neq k} s_i w_i \hat{x}_i \right).$$  \hspace{1cm} (4)

This implies perfect receiver knowledge of the modulation vectors and amplitudes. Extensions of this algorithm have been described for unknown flat and frequency
selective fading channels e.g. [11]. Under a Gaussian assumption, these signals are converted to probabilities via

\[ T(\tilde{x}_k, d) = \alpha p_k(d) \exp \left(-\frac{(M(d) - \tilde{x}_k)}{2v}\right) \]

where \( v \) is the variance of the residual noise plus interference, which must be estimated. Finally, the prior probabilities are removed to produce extrinsic probabilities for input to the bank of single user APP decoders,

\[ q_k(d) = \frac{\exp \left(-\frac{(M(d) - \tilde{x}_k)}{2v}\right)}{\sum_{d'} p_k(d') \exp \left(-\frac{(M(d') - \tilde{x}_k)}{2v}\right)} \] (5)


### 3 Variance Transfer Analysis

In [4] a convergence analysis of iterative soft interference cancelation was given. This analysis is based on evolution of error variance through the system components and is similar in concept to the by now well known EXIT chart analysis first appearing in [12]. Variance transfer analysis works by scalar parametrization of the iterative process using the average error variance on $\tilde{x}_k$ (output of canceler) and $\hat{x}_k$ (output of decoder). Under the assumption of i.i.d. modulation vectors, equal received powers, and use of the same code for each user, the linearity of the cancelation process yields a simple characteristic. Write the canceler input as $\hat{x} = x + z_D$ where $\text{var } z_D = v_D$ and the canceler output as $\tilde{x} = x + z_{IC}$, where $\text{var } z_{IC} = v_{IC}$. In [4] it was shown that

$$v_{IC}(v_D) = \frac{K - 1}{N} v_D + \sigma^2 \rightarrow \beta v_D + \sigma^2. \quad (6)$$

where the convergence is as $K, N \to \infty$ with constant $\beta$. Simple expressions like (6) motivate the use of error variance, rather than mutual information. The decoder characteristic, $v_D(v_{IC})$ is typically obtained by monte-carlo simulation. The main goal of this paper is to extend (6) to heterogeneous systems, in which the transmitters use different encoders and are received at different power levels. Familiarity with the analysis of [4] will be assumed.

Let each code $C_k$ have a variance transfer curve $f_k(v)$, which is the error variance at the output of the APP decoder for input noise variance $v$ and unit symbol energy. From this unit power curve, the characteristic for power $w_k^2$ is easily found as $f_k(v/w_k^2)$. Let $S$ have i.i.d. zero mean entries with variance $1/N$. Assume that $K$ is large enough so that the initial noise variance $v^{(0)}$ for each user $k$ is independent of $k$ (the superscript keeps track of the iteration number). Let user $k$ be a specific user of interest. At the output of APP decoder $k$ we have $\hat{x}_k = x_k + z_k$ where $\text{var } z_k = f_k(v^{(0)}/w_k^2)$ and these $z_k$ are assumed to be mutually independent and also independent of $S$. These independence assumptions is of course not strictly true, but are good in the limit of large block length and independent interleaving of each user. The accuracy of the results obtained show that this assumption is acceptable.

Collect the extrinsic estimates into a vector $\hat{x}_k = (\hat{x}_1, \ldots, \hat{x}_{k-1}, 0, \hat{x}_{k+1}, \ldots, \hat{x}_K)^t$. Similarly, let $z_k = (z_1, \ldots, z_{k-1}, 0, z_{k+1}, \ldots, z_K)^t$. The canceled signal for user $k$ is

$$s_{k}^t(r - SW\hat{x}_k) = s_k^t w_k x_k + s_k^t (SWz + n)$$

where

$$s_k^t w_k x_k + s_k^t (SWz + n)$$

represents the signal and noise terms.
which has signal power $w_k^2$ and noise power

$$v_k^{(1)} = \sigma^2 + \text{var} \sum_{j \neq k} s_k^j s_j w_j z_j$$

$$= \sigma^2 + \sum_{j \neq k} \text{var}(s_k^j s_j w_j z_j) \quad \text{independence of the } z_j$$

$$= \sigma^2 + \sum_{j \neq k} w_j^2 \text{var}(s_k^j s_j) \text{var}(z_j) \quad \text{independence of } z_k, S$$

$$= \sigma^2 + \frac{1}{N} \sum_{j \neq k} w_j^2 f_j(v(0)/w_j^2).$$

For large enough $K$ we therefore have at iteration $n$

$$v^{(n+1)} = \sigma^2 + \beta \sum_{k=1}^{K} \frac{w_k^2}{K} f_k(v^{(n)}/w_k^2) = \sigma^2 + \beta \bar{f}(v^{(n)})$$

where

$$\bar{f}(v) \triangleq \sum_{k=1}^{K} \frac{w_k^2}{K} f_k(v/w_k^2)$$

is defined as the system average variance transfer curve. In the case of identical power levels and codes, the existing result (6) is recovered. The next section considers some specific examples to illustrate this analysis.

### 4 Examples

The average transfer curve may be simplified in certain circumstances, such as a finite number of user classes. For example two codes $f_1$ and $f_2$ and unit powers,

$$\bar{f}(v) = \lambda f_1(v) + (1 - \lambda) f_2(v)$$

where there are $\lambda K$ users in class one and $(1 - \lambda) K$ users in class two.

Figure 4 shows the result of mixing the $(2, 1)$ repeat code and the maximal free distance 4 state convolutional code. The component code characteristics (obtained by individual monte-carlo simulation) are shown with dashed curves. The mixed code characteristic (14) is the solid curve and the canceler characteristic (6) is the solid straight line. Figure 4(a) shows how the variance transfer curve, due to $\lambda = 0.25$, results in convergence (an open tunnel) at $\beta = 2.2$. Neither of the component codes result in convergence at this load. Figure 4(b) shows a $\lambda = 0.5$ combination, again allowing convergence at a higher load than either of the components. In this case, a price is paid in increased final error performance.

Another interesting scenario is where there is only one code, but two or more power
classes. For example, with two power classes,

$$\bar{f}(v) = \lambda w_1^2 f(v/w_1^2) + (1 - \lambda) w_2^2 f(v/w_2^2).$$

(15)

The additional factor of $w_k$ inside the argument of the $f_k$ gives an extra degree of freedom in shaping the average code curve.

Figure 5(a) shows the average variance transfer curve for the 4 state convolutional code, with two power classes. One class has power $w_1^2 = 2$ and the second class is varied between $w_2^2 = 0.1$ and $w_2^2 = 1$ in steps of 0.1. The class sizes are chosen so that the total average power is fixed at 1. It is apparent that a wide variety of characteristics are possible, even using only two power classes. A similar idea has already been used in [13] to optimize spectral efficiency (they parameterize using multiuser efficiency). Figure 5(b) shows the actual simulated trajectory (zig-zag line) for the 4 state code with $w_1^2 = 2.5$ and $w_2^2 = 10$ (class sizes are chosen such that $E_b/N_0 = 10$ dB). Plotted for reference is the standard $w^2 = 5$ ($E_b/N_0 = 10$ dB) curve (dashed). There are $K = 18$ users and $N = 8$ dimensions (a total of 1.125 bits/dim). In the absence of power mixing, the receiver would not converge at all. Naturally, allowing even more power classes results in greater flexibility and provides the possibility to optimize the convergence, see [13].

Now consider quasi-static Rayleigh fading, where the $w_k$ are modeled as independent zero mean, unit variance Gaussian random variables. This situation may arise in packet data transfer in an ad hoc network where the length of the packet is short compared to the channel coherence time. Rather than enforcing strict power control to optimize power classes, it is interesting to see what kind of characteristic arises naturally. Under the usual large systems assumption, the average variance curve
becomes an expectation, namely

\[ \bar{f}(v) = \sum_{k=1}^{K} \frac{w_k^2}{K} f_k(v/w_k^2) \rightarrow E \left[ w^2 f\left(\frac{v}{w^2}\right) \right], \]  

(16)

where the expectation is with respect to \( w \), a zero mean, unit variance Gaussian random variable. Figure 6(a) shows the result (applied to the 4 state convolutional code). It appears that the Rayleigh fading power profile is particularly helpful. The supportable load is \( \beta = 4 \), which is more than double than that of AWGN with no fading. Note however that the resulting bit error rates are those obtained over the corresponding single-user fading channels. These results invoke a large systems assumption, so how accurate are they for reasonable sized systems? Figure 6 shows two examples. Figure 6(b) compares the faded (solid lines) and non-faded (dashed lines) trajectories for the 4 state convolutional code, \( N = 16 \) dimensions and \( K = 25 \) users at \( E_b/N_0 = 10 \) dB. In the absence of fading, the receiver requires 7 iterations to converge, compared to only 3 iterations in the presence of fading. Clearly the fading power profile is more conducive to iterative decoding. Figure 6(c) shows a more highly loaded system, with \( K = 50 \), confirming the accuracy of the analysis.

Given a set of target received power levels (perhaps optimized for some finite number of user classes), is it interesting to consider the possibility of exploiting the natural power variation due to fading, assigning naturally strong users to high power classes and naturally weaker users to low power classes. At first sight this would appear to be beneficial, but a simple analysis of this idea reveals that in fact the assignment of users to classes is irrelevant in terms of total required transmit power. Fading provides no advantage or disadvantage in terms of meeting specified receive power requirements.
5 Conclusion

The analysis of [4] has been extended to heterogeneous systems in which the received powers and the transmitters’ codes do not have to be the same. A simple formula has been found for the variance transfer characteristic of the soft interference canceler. Computer simulations have verified the accuracy of this analysis. Application of this analysis shows how allocating different codes or transmission powers to different user groups can result in greatly improved convergence behavior. Of particular interest is the finding that the received power profile resulting from a flat Rayleigh fading channel is particularly compatible with the iterative receiver. In that scenario, spectral efficiencies of up to two bits per dimension can be achieved. These results are favorable for ad-hoc wireless networks, where natural variation in received power may assist with interference cancelation and improve network performance.
References


