Multiuser Detection with Particle Filtering

Jianqiu Zhang, Yufei Huang, and Petar M. Djurić
Dept. of ECE, State Univ. of New York at Stony Brook, Stony Brook, NY, 11794
e-mail: {jizzhang, yfhuang, djuric}@ece.sunysb.edu

ABSTRACT
In this paper, we apply particle filtering to multiuser detection (MUD) in synchronous CDMA systems with
perfect channel state information (CSI). To apply particle filtering to MUD, first we find a factored representation
of the posterior distribution. We show that the whitened matched filter (WFM) output allows such represen-
tation. No system approximations are made and the method is “soft” in nature. As a result, a near op-
timum performance is observed in simulations both in the equal power case and the near-far case. Since the
computational complexity of the method is not exponential with the number of users, the method may have
potential for industrial application.

1 Introduction
When MUD was introduced in the eighties, it quickly
received a great deal of attention due to its potential
for reducing the effects of multiple access interference
(MAI) and thereby for increasing the capacity of CDMA
systems. Since optimum maximum likelihood MUD is
exponential in complexity, numerous approximate detec-
tors were developed for MUD. In general, successive/iterative detectors [3], such as the decorrelating
decision-feedback detector, perform better than their corresponding “one-pass” detectors like, for example,
the decorrelating detector. These successive/iterative
detectors usually form interim hard decisions for later
stages and are, thus, prone to error propagation. Conse-
quently, the performance of these conventional detectors
is not near optimum.

In [10], a Markov chain Monte Carlo method, the
Gibbs sampler, was used for CDMA MUD. It was im-
plemented in a Bayesian framework, and it was demon-
strated that it could provide near optimum performance.
It is “soft” in nature, that is, the method allows for ex-
change of “extrinsic” information in an iterative (turbo)
joint MUD and channel decoding. However, the Gibbs
sampler has inherent drawbacks. It is hard to deter-
hine when the underlying Markov chain converges, and
sometimes, the Gibbs sampler gets stuck at a local op-
timum. As a result, when we experimented with it on
MUD, there was an error floor in its performance.

In this paper, we aim to develop a Bayesian-based so-
lution that can outperform conventional detectors and
overcome the shortcomings of the Gibbs sampler. To
that end we would like to keep the capability of pro-
viding “extrinsic” information while performing MUD.
We propose a particle filtering approach, a methodology
that has reemerged recently in the fields of engineering
[4], econometrics [8], and statistics [5]. However, to ap-
ply particle filtering, one needs to find a dynamic state
space model representation of the system which in our
problem is not obvious. We show that the whitened
matched filter (WFM) output, which is used in decor-
relating decision feedback detectors, allows such a represen-
tation. Note that our work is different from that in
[6], where the binary data of all users in a symbol
interval are considered as a super symbol. As a result,
the sample space grows exponentially with the number
of users. By contrast, our algorithm samples one user at
a time in the binary space and can handle large num-
ber of users.

2 System description
Consider a synchronous CDMA system with a chip rate
(processing gain) $C$ and $K$ users. Let $T$ denote the
symbol duration and $s_k(t)$ the normalized deterministic
signature waveform assigned to the $k$th user. Here $t \in
[0,T]$ and $k \in \{1,\ldots,K\}$. Let $b_k \in \{-1, +1\}$ be the
bit transmitted by the $k$th user, $a_k$ the channel state
information of the $k$th user, and $\sigma_n(t)$ the received zero
mean complex white Gaussian noise with variance $\sigma^2$.

We can express the received signal $y(t)$ as

$$y(t) = \sum_{k=1}^{K} a_k b_k s_k(t) + \sigma_n(t) \quad t \in [0,T]. \quad (1)$$

The received signal is a superposition of $K$ antipodally
modulated synchronous signature waveforms plus noise.
The cross-correlation between the signature waveforms

---

*This work was supported by the National Science Foundation under Awards CCR-0000320 and CCR-0008267.
of the $i$th and the $j$th users is defined according to
\[
R_{ij} = s_i(t) s_j(t) dt
\]

where $R_{ij}$ is the $ij$th element of the cross-correlation matrix $R$.

In CDMA systems, we often work with the matched filter output,
\[
y_k = < y(t), s_k(t) > = \int_0^T y(t) s_k(t) dt.
\]

The set of matched filter outputs $y = [y_1, y_2, \cdots, y_K]^T$ (T stands for vector or matrix transposition) can be represented in a vector-matrix form according to
\[
y = RA \hat{b} + n
\]

where $A = \text{diag} \{ a_1, a_2, \cdots, a_K \}$ is the diagonal matrix of the channel state information, $b = [b_1, b_2, \cdots, b_K]^T$ is the user symbol vector, and $n$ is a complex-valued Gaussian vector with independent real and imaginary components and covariance matrix equal to $\sigma^2 R$.

The cross-correlation matrix is positive definite, and we can factor it by Cholesky decomposition. There exists a unique lower triangular matrix $F$ such that $R = F^T F$. If we apply $F^{-1} F = (F^T)^{-1}$ to the matched filter output, we obtain
\[
\tilde{y} = (F^T)^{-1} y = F^{-T} F A \hat{b} + F^{-T} n = F \hat{a} + \tilde{n}
\]
or equivalently
\[
\tilde{y} = F \hat{a} + \tilde{n}
\]

where $B = \text{diag} \{ b_1, b_2, \cdots, b_K \}$ is the user symbol matrix, and $a$ is the $K \times 1$ vector of channel state information. The covariance matrix of $\tilde{n}$ becomes
\[
E (\tilde{n} \tilde{n}^T) = \sigma^2 F^T R F^{-1} = \sigma^2 I
\]

where $I$ is the identity matrix. Since the noise becomes i.i.d. white Gaussian, $\tilde{y}$ is called the whitened matched filter (WMF) output. The expression in (6) can be written in component-wise form as
\[
\tilde{y}_k = \sum_{l=1}^K F_{k,l} a_l + \tilde{n}_k, \quad k = 1, 2, \cdots, K.
\]

In the sequel, let $\tilde{y}_1:k = \{ \tilde{y}_1, \cdots , \tilde{y}_k \}$, and define $b_{1:k}$ and $a_{1:k}$ similarly. The objective is to detect $b_{1:k}$ given $\tilde{y}_1:k$.

3 Particle filtering MUD with perfect CSI

All the prior information and information from the data is combined in the posterior distribution $p(b_{1:k} | \tilde{y}_{1:k})$. Direct evaluation of the posterior is impossible for large $K$ due to its high dimensionality. An alternative is to obtain samples (trajectories) $\{ b_{1:j} \}_{j=1}^J$ and weights $\{ w_{1:j} \}_{j=1}^J$ associated with the trajectories, where $j$ is the trajectory index, and using them to compute desired estimates. For example, we can approximate the expected value of any function of $b_{1:k}$, $f(b_{1:k})$, by
\[
E(f(b_{1:k})) \approx \frac{1}{J} \sum_{j=1}^J w_{1:j} f(b_{1:j} | \tilde{y}_{1:k}).
\]

If the trajectories are drawn from the posterior distribution itself, all the trajectories have equal weights. Otherwise, if they are drawn from a proposal importance function $\pi (b_{1:k} | \tilde{y}_{1:k})$, the weights are evaluated according to
\[
w_{1:j} = \frac{p(b_{1:j} | \tilde{y}_{1:k})}{\pi (b_{1:j} | \tilde{y}_{1:k})}, \quad \forall j.
\]

By the law of large numbers, under certain conditions, the approximation in (9) reaches the true value when the number of samples $J$ approaches infinity.

The posterior distribution can be factored according to
\[
p(b_{1:k} | \tilde{y}_{1:k}) \propto p(\tilde{y}_{1:k} | b_{1:k}) p(b_k) p(b_{1:k-1} | \tilde{y}_{1:k-1})
\]

and if we choose the proposal distribution in the form of $\pi (b_{1:k} | \tilde{y}_{1:k}) = \pi (b_{1:k-1} | \tilde{y}_{1:k-1}) \pi (b_k | b_{1:k-1}, \tilde{y}_{1:k})$, we can form the trajectories recursively. Suppose that from previous iterations we have generated the trajectories $\{ b_{1:j-1} \}_{j=1}^J$ from $\pi (b_{1:k-1} | \tilde{y}_{1:k-1})$ with weights $\{ w_{k-1} \}_{j=1}^J$. If we draw particles $b_{1:k}$ from the importance proposal distribution $\pi(b_k | b_{1:k-1}, \tilde{y}_{1:k}) = p(b_k | b_{1:k-1}, \tilde{y}_{1:k})$ and append them to $b_{1:k}$, the expanded trajectory $b_{1:k}$ can be weighted with respect to $p(b_{1:k} | \tilde{y}_{1:k})$ according to
\[
w_{1:k} = \frac{p(b_{1:k} | \tilde{y}_{1:k})}{p(b_{1:k-1} | \tilde{y}_{1:k-1}) \pi(b_k | b_{1:k-1}, \tilde{y}_{1:k})} \frac{p(b_{1:k} | \tilde{y}_{1:k-1})}{p(b_{1:k-1} | \tilde{y}_{1:k-1})} \pi(b_k | b_{1:k-1}, \tilde{y}_{1:k})
\]

or equivalently
\[
w_{1:k} = \frac{p(b_{1:k} | \tilde{y}_{1:k})}{p(b_k | b_{1:k-1}, \tilde{y}_{1:k-1}) \pi(b_k | b_{1:k-1}, \tilde{y}_{1:k})} \pi(b_k | b_{1:k-1}, \tilde{y}_{1:k})
\]

where $w_{k-1} = \frac{p(b_{1:k-1} | \tilde{y}_{1:k-1})}{\pi(b_{1:k-1} | \tilde{y}_{1:k-1})}$ is obtained from the previous iteration. In deriving the above equation, we utilized the fact that $p(b_{1:k} | \tilde{y}_{1:k-1}, \tilde{y}_{1:k-1}) = p(b_{1:k})$, i.e., $b_k$ is independent of other users and previous observations. We have also ignored the term $p(\tilde{y}_{1:k} | b_{1:k-1})$ because it is the same for all trajectories. The importance proposal distribution used here is referred to as optimal because it takes into account all the previous particles and all available observations, and as a result produces weights...
with minimal variance conditional on $b_{1:k-1}^{(j)}$ and $\bar{y}_{1:k}$ [2],

the proposal distribution can be evaluated according to

$$
\begin{align*}
\pi(b_k | b_{1:k-1}^{(j)}, \bar{y}_{1:k}) &= p(b_k | b_{1:k-1}^{(j)}, \bar{y}_{1:k}) \\
&\propto p(b_k | b_{1:k-1}^{(j)}, \bar{y}_{1:k-1}) p(b_k | b_{1:k-1}, \bar{y}_{1:k-1}) \\
&= p(b_k | b_{1:k-1}^{(j)}) p(b_k). 
\end{align*}
$$

(13)

Note that the weight is proportional to the sum of the proposal densities in (13). In our case, since $b_k \in \{-1, 1\}$, there are only two proposal densities to evaluate and the weight can be easily obtained.

In summary, the algorithm proceeds as follows:

1. Draw a particle $b_k^{(j)}$ from the proposal distribution (13).

2. Append $b_k^{(j)}$ to $b_{1:k-1}^{(j)}$ and obtain the new trajectory $b_{1:k}^{(j)}$.

3. Evaluate the weight of the $j$th trajectory using (12).

When the algorithm is completed with the last user $K$, we have trajectories and weights $\{b_{1:K}^{(j)}, w_{1:K}^{(j)}, j = 1\}$ that can approximate $p(b_{1:K} | \bar{y}_{1:K})$, the desired posterior distribution. This process of recursively obtaining the particles $b_k^{(j)}$ is called particle filtering.

From the generated particles and their weights, we can obtain various types of estimates. For example, the marginalized posterior distribution can be approximated by

$$
p(b_k | \bar{y}_1:K) \simeq \frac{1}{\sum_{j=1}^{J} w_{1:K}^{(j)}} \sum_{j=1}^{J} w_{1:K}^{(j)} \delta(b_k - b_k^{(j)})
$$

(14)

where $\delta(\cdot)$ is the Dirac delta function. If the adopted estimate of $b_k$ is the one that maximizes the marginalized posterior distribution, we have

$$
b_k = \arg \max_{b_k} p(b_k | \bar{y}_1:K)
$$

\begin{align}
&\simeq \arg \max_{b_k} \sum_{j=1}^{J} w_{1:K}^{(j)} \sum_{j=1}^{J} w_{1:K}^{(j)} \delta(b_k - b_k^{(j)}).
\end{align}

(15)

If $b_k = [b_k^{(1)}, \ldots, b_k^{(J)}]^T$ and $w_K = [w_1^{(j)}, \ldots, w_K^{(j)}]^T$, (15) can be simplified as

$$
\hat{b} = \text{sign}(b_K^T w_K).
$$

(16)

Note that “extrinsic” information can also be derived from (14).

An important aspect of the particle filtering process is the need for resampling. Namely, after several steps, some weights of the samples become trivial and stop contributing to the overall estimates. In the literature of particle filtering, resampling is used so that samples with negligible weights are replaced by those from high distribution areas of the desired posterior distribution. There are many strategies for resampling, and we use the residual resampling procedure as described in [1].

The complexity of the algorithm is $O(KJ)$, i.e., proportional to the product of the number of particles and number of users. If the number of particles is fixed, then the complexity is only linear with respect to the number of users.

4 Simulations

We simulated a synchronous CDMA system with $K = 15$ users with equal powers and a chip rate of $C = 30$. The spreading codes were generated randomly and the same spreading code was used in all experiments. Residual resampling was performed after every 5 users. The results of the performance comparison with other popular CDMA multiuser detectors are presented in Figure 1. The performance curves in the figure were obtained by averaging the Bit Error Rates (BERs) of all 15 users.

The detectors used in the comparison include the three-stage successive cancellation detector with decorrelating first stage (3-stage) [9], the detector based on Gibbs sampling [10], and the decorrelating decision feedback detector (DDF) [3]. For the Gibbs sampler, we experimented with two scenarios with different burn-in periods (the periods until convergence). In the first case, 100 samples were generated of which the first 50 samples were discarded (Gibbs-50). In the second case, 150 samples were drawn, and the first 100 samples were discarded (Gibbs-100). As a reference, we used the breadth-first tree-search algorithm [7], which is optimal, to provide a lower bound for the detectors.

From the results, we see that the particle filtering provides near-optimum performance. We used two particle filtering detectors, one with 50 particles for each user (PF-50) and another with 100 particles (PF-100). It appears that the performance gain by increasing the number of particles from 50 to 100 is only marginal.

The figure also shows that at high SNRs, the Gibbs sampler gets trapped at some local optimum and that it takes long time for the algorithm to converge to the global optimum. Consequently, with a limited burn-in period, the Gibbs sampler exhibits an error floor. In comparison, the particle filtering based detector does not have this problem.

We also investigated the performance of these detectors in a near-far scenario. In our experiment, the targeted user (the first user), had an SNR of 9 dB, and the signal strength of the remaining 14 users was identical. In comparison with the power of the targeted user, the power of the remaining users, $E_b / E_1$, varied from $-10$ dB to 10 dB. In Figure 2, we plotted the BER of the first user as a function of $E_b / E_1$. It is clear that particle filtering almost always outperforms the 3-stage detector and although it performs worse than the Gibbs sampler with weaker interferers, it is more consistent than
5 Conclusions

In this paper, we used the WMF output to derive a factored representation of the posterior distribution function for the MUD problem, and the representation allowed for application of particle filtering. The proposed method provides consistently better performance than several existing detectors which have been tested. The complexity of the method is linear in the number of users and therefore, it has great potential in practice. The method, however, needs further study which may lead to additional improvements of its performance. For example, a thorough examination of the detector’s performance as a function of the number of particles and number of users is necessary, and other strategies for resampling should be explored.

References


