Multiuser Detection of Synchronous Code-Division Multiple-Access Signals by Perfect Sampling

Yufei Huang and Petar M. Djurić, *Senior Member, IEEE*

Abstract—Code-division multiple-access (CDMA) is a multiplexing technique that shows significant advantages over analog and conventional time-division multiple access (TDMA) systems. This technology has become a driving force behind the rapidly advancing communications industry. In order to recover the transmitted signal at the receiver in a CDMA system, demodulating techniques are engaged, where a prominent role is played by the multiuser detector. In this paper, we introduce a class of novel Bayesian multiuser detectors that are constructed by employing perfect sampling algorithms: the sandwiched CFTP and the Gibbs coupler. We show that the detector based on the sandwiched CFTP can be applied to systems with negative cross-correlations, whereas the Gibbs coupler detector can be used without restrictions. A salient feature of the proposed detectors is the use of exact (perfect) samples from posterior distributions. This feature provides them with several advantages over detectors based on the Gibbs sampler. Simulation results on systems with and without near–far resistance demonstrate improved performance of the proposed detectors over some other popular detectors. In the end, we also discuss some important computational issues of the proposed detectors.

Index Terms—Code-division multiple-access, coupling from the past, Gibbs coupler, Gibbs sampler, Markov chain Monte Carlo, multiuser detection, perfect sampling.

I. INTRODUCTION

CODE-DIVISION multiple-access (CDMA) [1]–[3] is a multiplexing technique that enables multiple users to access a common channel simultaneously. In a CDMA system, each user is assigned a unique signature waveform, and the data message of the user is then spread by modulating the signature waveform. Since the modulated signal has a much wider bandwidth than that required for simple point-to-point communications, a CDMA system is also referred to as spread spectrum system. CDMA systems show significant advantages over analog and conventional time-division multiple access (TDMA) systems, including increased capacity, enhanced privacy and security, and reduced effects of multipath fading. As a result, the technology of CDMA has become a driving force behind the rapidly advancing communications industry.

Recovery of transmitted signals at the receiver is achieved by demodulation techniques. A widely used technique called the single-user matched filter regards the CDMA channel as a single-user channel and considers the detection problem of each user individually. As is well known, in the absence of user interference, the single-user matched filter is optimal in the sense of minimizing the bit-error rate. However, this is no longer true in CDMA systems where multiuser interference (MUI) is present. The degraded performance of the single-user matched filter is particularly emphasized in the presence of near–far effects. To overcome these disadvantages of the single-user matched filter, multiuser detectors have been developed [1], [4].

Multiuser detectors regard the MUI as additional information rather than noise. Thus, better performance over single-user detectors is achieved by processing this information. The optimum multiuser detector is, however, a combinatorial optimization problem. Although it can be always solved by exhaustive search [1], the computational complexity of exhaustive search increases exponentially with the number of users and, therefore, makes it infeasible for systems with usual capacity. Due to the complexity of the optimum detector, much effort has been devoted to finding suboptimal linear detectors that can achieve certain balance between performance and computational complexity [5]. An example is the decorrelating detector, or decorrelator, which is computationally simple and exhibits relatively good near–far resistance. However, there is still a large margin left between the performance of the linear detectors and the optimum detector.

In recent years, with the advent of powerful computers, much attention has been given to multiuser detectors using nonlinear methods. Among them are the detectors that employ the technique of successive interference cancellation [6]. They include the multistage interference cancellation algorithm [7] and the decorrelating decision-feedback scheme [8]. In addition, there are detectors using the expectation-maximization (EM) algorithm [9], genetic algorithms [10], neural networks [11], and sampling-based methods.

Sampling-based methods, especially Markov chain Monte Carlo algorithms (MCMCs) [12], have been intensively studied in statistics during the past decade. These methods have demonstrated excellent performance in solving high-dimensional optimization and integration problems. Among them, the Gibbs sampler [13] and the Metropolis–Hastings algorithm [14] are the most popular algorithms. In the past few years, they have penetrated the signal processing and communications community, and as a result, many applications of MCMC methods in signal processing and communications can be found in the literature. In a recent paper, a method based on Gibbs sampling for joint multiuser channel parameter estimation and signal detection has been proposed [15].

MCMC sampling, however, may generate samples from desired distributions only approximately. Additionally, the

generated samples are dependent, which may render bias and larger variance in Monte Carlo computations. In 1996, Propp and Wilson proposed a perfect sampling algorithm called coupling from the past (CFTP) [16], which completely resolves the drawbacks of MCMC in that it generates i.i.d. samples exactly from a desired distribution. Their work has drawn much interest, and as a result, further progress on perfect sampling has been reported [17]–[19].

In this paper, we present a class of new multiuser detectors that are developed under the Bayesian framework and implemented by perfect sampling algorithms. We first discuss the multiuser detector by the Gibbs sampler. Using the Gibbs sampler as a building block, we show how the sandwiched CFTP can be applied to a system with negative cross-correlations between signature waveforms. Furthermore, we employ a new perfect sampling algorithm called the Gibbs coupler and demonstrate how a detailed scheme can be constructed for detection problems in general systems. Overall, a salient feature of the proposed detectors is the use of exact samples from the posterior distributions. This advantage improves the performance of the new detector over the one based on Gibbs sampling.

The paper is organized as follows. The problem of optimum multiuser detection is addressed in Section II. A background of perfect sampling is provided in Section III, where the Gibbs sampling algorithm, CFTP, and the Gibbs coupler are introduced. In Section IV, the implementation of the Gibbs coupler on multiuser detection of synchronized CDMA signals is carefully studied. Simulation results are presented in Section V.

II. OPTIMUM MULTIUSER DETECTION

A K-user synchronous CDMA system with white Gaussian noise can be modeled as

$$y(\tau) = \sum_{k=1}^{K} A_k b_k s_k(\tau) + n(\tau), \ \tau \in [0, T] \qquad (1)$$

where

$y(\tau)$	received signal;
$s_k(\tau)$	antipodal signature waveform of the kth user;
A_k	amplitude of the kth user's signal;
$b_k \in \{-1, 1\}$	bit transmitted by the kth user;
$n(\tau)$	additive white Gaussian noise with zero mean and variance σ^2;
T	symbol duration.

The main interest in multiuser detection is to detect correctly the transmitted symbols by each user. In a real system, the amplitudes of the users' signals and the noise variance are also unknown to the receiver, and they can be estimated from pilot signals. The estimation of these parameters is not of concern in this paper; see [20]–[22] for detailed discussions of this issue. Here, we assume that all the parameters except the b_ks are known, and our objective is to estimate $\mathbf{b}^T = \begin{bmatrix} b_1 & b_2 & \cdots & b_K \end{bmatrix}$. The conventional single-user matched filter makes the decision on each user's symbol separately, and the estimate of the kth user is expressed as

$$\hat{b}_k = \text{sgn}(y_k) \qquad (2)$$

where $y_k = \int_0^T s_k(\tau) y(t) d\tau$ is the kth matched filter output. Note that the single-user matched filter does not exploit any information about the users' correlations and regards the users' interference simply as noise. Therefore, the single-user matched filter is not optimal in the presence of multiuser interference, which exists in CDMA systems. To achieve optimum detection, one must employ a multiuser detection strategy. From a Bayesian perspective, the optimum decision is made using the posterior distribution of $\mathbf{b}$. Since nothing is known *a priori* about $\mathbf{b}$, a noninformative prior is chosen for $\mathbf{b}$, and the *a posteriori* distribution of $\mathbf{b}$ can be expressed as

$$p(\mathbf{b}|y(\tau)) \ \tau \in [0, T]$$

$$\propto \exp\left(-\frac{1}{2\sigma^2} \int_0^T \left[y(\tau) - \sum_{k=1}^{K} A_k b_k s_k(\tau)\right]^2 d\tau\right)$$

$$\propto \exp\left(\frac{1}{2\sigma^2}\left(2\sum_{k=1}^{K} A_k y_k b_k - \sum_{k=1}^{K}\sum_{l=1}^{K} A_k A_l \rho_{kl} b_k b_l\right)\right) \qquad (3)$$

where $\rho_{kl} = \int_0^T s_k(\tau) s_l(t) d\tau$ represents the crosscorrelation between the kth and the lth signature waveform. Under the Bayesian paradigm, we adopt the *maximum a posteriori* (MAP) detector [23] and the *marginalized posterior mode* (MPM) detector [23], [24]. The MAP detector $\hat{\mathbf{b}}_{MAP}$ is the set of symbols that maximizes the posterior distribution (3), i.e.,

$$\hat{\mathbf{b}}_{MAP} = \arg\left\{\max_{\mathbf{b} \in \{-1, 1\}^K} p(\mathbf{b}|y(\tau))\right\}, \tau \in [0, T] \qquad (4)$$

and the MPM detector calculates the marginalized posterior mode, i.e.,

$$\left(\hat{\mathbf{b}}_{MPM}\right)_k = \text{sgn}\left(\sum_{\mathbf{b} \in \{-1, 1\}^K} b_k p(\mathbf{b}|y(\tau))\right), \tau \in [0, T] \qquad (5)$$

for all $k = 1, 2, \ldots K$, where $(\hat{\mathbf{b}}_{MPM})_k$ denotes the kth element of $\hat{\mathbf{b}}_{MPM}$. Clearly, the MAP detection is basically a combinatorial optimization problem, and the solution can be obtained by exhaustive search. Exhaustive search over a variable space with large dimension K, however, is computationally prohibitively expensive because it involves the calculation of 2^K terms. Similarly, the MPM detector $\hat{\mathbf{b}}_{MPM}$ requires a multidimensional summation that again involves the calculation of 2^K terms. The overwhelming computational requirements of the optimum solution has been driving the research on linear and nonlinear suboptimal solutions with lower complexity. In recent years, more attention has been given to nonlinear detectors, which usually provide better performance than linear detectors. Among the nonlinear detectors, the ones based on MCMC algorithms are of particular interest here.

III. BACKGROUND ON PERFECT SAMPLING ALGORITHMS

A. Markov Chain Monte Carlo

In Bayesian analysis, it is very important to optimize or integrate over the posterior distribution $p(\mathbf{x}|\mathbf{y})$, where $\mathbf{x}^T = \begin{bmatrix} x_1 & x_2 & \cdots & x_N \end{bmatrix}$ represents a vector of N unknown

variables, and $\mathbf{y}$ denotes the observation data. For practical problems, the dimension of the unknowns N is usually large. Then, the required high-dimensional optimization and integration impose a great challenge to conventional numerical techniques, and quite often, these techniques fail to provide satisfactory solutions. Nonetheless, Monte Carlo methods, especially those based on Markov Chain Monte Carlo (MCMC) sampling, have shown excellent performance. Their development has provided the impetus for the recent boom of Bayesian applications in statistical signal processing.

In the multiuser detection problem, our interest lies in employing Monte Carlo methodology to solving the integrals (summation) and maximization of the posterior distributions. Unlike the conventional numerical methods that are deterministic in nature, the Monte Carlo methods adopt a stochastic approach. To illuminate the general idea of Monte Carlo integration and optimization, suppose that we are interested in finding the solutions to the following problems:

$$J_O = \arg \max_{\mathbf{x}} p(\mathbf{x}|\mathbf{y}) \tag{6}$$

and

$$J_I = \int_{\mathbf{x}} f(\mathbf{x}) p(\mathbf{x}|\mathbf{y}) d\mathbf{x} \tag{7}$$

where $f(\mathbf{x})$ is some function of $\mathbf{x}$. The key for Monte Carlo solutions is the generation of random samples $\{\mathbf{x}^{(t)}\}_{t=1}^M$ from the posterior distributions $p(\mathbf{x}|\mathbf{y})$, where M denotes the number of generated samples. The solution J_O can be approximated by

$$J_O \approx \arg \left\{ \max_{\mathbf{x} \in \{\mathbf{x}^{(t)}\}_{t=1}^M} p(\mathbf{x}|\mathbf{y}) \right\} \tag{8}$$

and it converges to the true value as $M \to \infty$. The approach is called *stochastic exploration* [25] since it tries to explore the whole variable space of the unknowns in a random fashion. This method is more effective when the variable space is discrete and finite. Another Monte Carlo optimization approach is referred to as *stochastic approximation* [25], and it uses randomly drawn samples to approximate the posterior distributions. In this paper, we use the stochastic exploration approach to define the MAP detector. The Monte Carlo approximation of J_I can be obtained from

$$J_I \approx \frac{1}{M} \sum_{t=1}^M f\left(\mathbf{x}^{(t)}\right). \tag{9}$$

It is shown that by the strong law of large numbers, this approximation converges to J_I almost surely and in mean square as $M \to \infty$ [26], [27]. In general, if the variable space of $\mathbf{x}$ is continuous, the Monte Carlo approach approximates the original high-dimensional optimizations and integrations by much simpler discrete optimizations and summations. Further, when the variable space of $\mathbf{x}$ is discrete, the computational complexity of the Monte Carlo optimization is a function of the sample size M, which is usually far smaller than the size of the discrete variable space.

Very often, however, direct sampling from $p(\mathbf{x}|\mathbf{y})$ is extremely difficult. By contrast, MCMC methods can obtain samples from such distributions by constructing a homogeneous Markov chain on the support of $\mathbf{x}$, where the equilibrium distribution of the chain is $p(\mathbf{x}|\mathbf{y})$ [14], [26]–[29]. The chain starts from some initial state, but it takes some transition time that is believed to be long enough for the chain to have converged. This transition time is called "burn-in" period.

One main drawback of MCMC sampling is that the burn-in period is unknown. Thus, the generated samples can only be considered to approximately follow the stationary distribution. In addition, MCMC methods usually produce highly correlated samples. The correlation between the samples makes the stochastic exploration less efficient since if the chain is trapped in some local maximum, it tends to stay around the local maximum for a long time before it gets out. Although much work was devoted to overcome these problems, there was not much success until a complete solution was proposed by Propp and Wilson [16]. Their method is known as perfect sampling and is called *coupling from the past* (CFTP). CFTP is briefly discussed in Section III-C.

B. Gibbs Sampler

Among all the MCMC algorithms, the Gibbs sampler seems to be the most popular one. The Gibbs sampler was first introduced by Geman and Geman [13] in their study of image restoration. With Gibbs sampling, we usually avoid drawing all the variables at the same time and, instead, sample them a few at a time only. Thereby, a high-dimensional sampling is converted to several lower dimensional samplings, which essentially simplifies the original problem. The Gibbs sampler has the unique feature that it accepts all the proposed samples from the intermediate distributions. To illustrate the algorithm in some detail, suppose that we want to sample from the posterior distribution $p(\mathbf{x}|\mathbf{y})$. Given some arbitrary starting value $\mathbf{x}^{(0)}$, the Gibbs sampler obtains samples according to the following iterative procedure.

```
At the tth iteration
  Sample  x₁⁽ᵗ⁾ from  p(x₁|x₂⁽ᵗ⁻¹⁾, x₃⁽ᵗ⁻¹⁾, ..., xₙ⁽ᵗ⁻¹⁾, y)
  Sample  x₂⁽ᵗ⁾ from  p(x₂|x₁⁽ᵗ⁾, x₃⁽ᵗ⁻¹⁾, ..., xₙ⁽ᵗ⁻¹⁾, y)
  ⋮  ⋮
  Sample  xₙ⁽ᵗ⁾ from  p(xₙ|x₁⁽ᵗ⁾, x₂⁽ᵗ⁾, ..., xₙ₋₁⁽ᵗ⁾, y)
```

Apparently, for its implementation, the full conditional distributions $p(x_i|\mathbf{x}_{-i}, \mathbf{y})$, for $i = 1, 2, \ldots, N$ are required, where $\mathbf{x}_{-i}$ denotes all the components of $\mathbf{x}$ except x_i. In addition, these full conditional distributions should be easy for sampling. The joint distribution of the obtained sequence $\mathbf{x}^{(t)}$ is shown to converge to $p(\mathbf{x}|\mathbf{y})$ as $t \to \infty$ [13], [30]. For more information on Gibbs sampling, see [26], [27], [29], and [31].

C. Perfect Sampling Algorithms

Perfect sampling algorithms refer to algorithms that can generate exact samples from a desired distribution by running Markov chains. The first perfect sampling algorithm was proposed by Propp and Wilson and is known as CFTP [16]. CFTP was originally designed to generate samples from discrete state spaces, but later, its concept was extended to accommodate sampling from continuous state spaces [18], [19]. Our interest in this paper includes algorithms on state spaces of the form $\{-1, 1\}^N$, where N is an integer. For further references regarding perfect sampling on both discrete and continuous state spaces, see [32]–[34].

Now, we explain the CFTP algorithm. Suppose that the desired discrete state space $\mathcal{S}$ is of size $M = |\mathcal{S}|$. The basic idea

of CFTP is to initiate M Markov chains at every possible state in the state space $\mathcal{S}$ from some time $-T$ in the past and run them to time 0. It is noted that all the Markov chains should have the desired distribution as their stationary distribution and at any instant of time t, the same random seed $R^{(t)}$ and updating function $\Phi\left(\cdot, R^{(t)}\right)$ are applied to every chain to determine their new states. Suppose that there comes a time $\bar{t}$ when all the chains have reached the same state $\bar{x}$. Now, if we restart all the M chains from the infinite past but keep the same random seeds $R^{(t)}$ and update functions $\Phi\left(\cdot, R^{(t)}\right)$ for the transitions during $-T$ to $\bar{t}$, these chains will coalesce to $\bar{x}$ and at the latest at $\bar{t}$. The reason is the following. When the chains that started from the infinite past reach time $-T$, they only occupy a subset of the M states. Since the M chains that originally started from $-T$ with random seeds $R^{(t)}$ and update functions $\Phi\left(\cdot, R^{(t)}\right)$ coalesce to $\bar{x}$ at time $\bar{t}$, so must the subset of the M chains started from $-T$ with the random seeds $R^{(t)}$ and update functions $\Phi\left(\cdot, R^{(t)}\right)$. The coalescence at $\bar{t}$ implies that the effect of the M initial states has actually worn off. Apparently, since the chains that have started in the infinite past have coalesced, $\bar{x}$ is a steady state that comes from the desired distribution. To avoid bias, coalescence is always examined at time $t = 0$. Therefore, the objective is to find the starting time $-T$ from which all the M chains will coalesce by time 0. With this objective, the CFTP algorithm is implemented by the following iterative scheme:

```
CFTP(T)
  t ← -T,  S_t ← S
  while t < 0
    t ← t + 1
    S_t ← φ(S_{t-1}, U_t)
  if |S_t| = 1 then
    return(S_0)
  else
    CFTP(T' with T' > T).
```

We further elaborate on the CFTP algorithm with an example. In Fig. 1, we display a realization of a CFTP algorithm with the trajectories of the Markov chains. The variable space of interest contains eight different states, each of which is represented by a vector of three binary symbols. These states are listed on the vertical axis of the figure. The CFTP trial contains four iterations, each of which is depicted in a sub-figure. In each iteration, eight Markov chains are started (one from each state). A perfect sample was obtained when the algorithm was restarted at time -4.

It has been shown in [16] that CFTP produces a perfect sample in finite time with probability 1. For problems with large state spaces, the implementation of CFTP often becomes prohibitive due to the heavy computational burden in tracing all the chains. Propp and Wilson pointed out that a practically efficient simulation can be accomplished when the designed Markov chain is *monotonic* [16], [33]. A monotonic Markov chain has an updating function that preserves the partial order $\preceq$ on its state space, that is, $\Phi(x, R) \preceq \Phi(y, R)$ for all R whenever $x \preceq y$. According to the partial order, a maximal and a minimal state can be determined on the state space. Then, at any instant of time, the monotonicity will cause all the chains
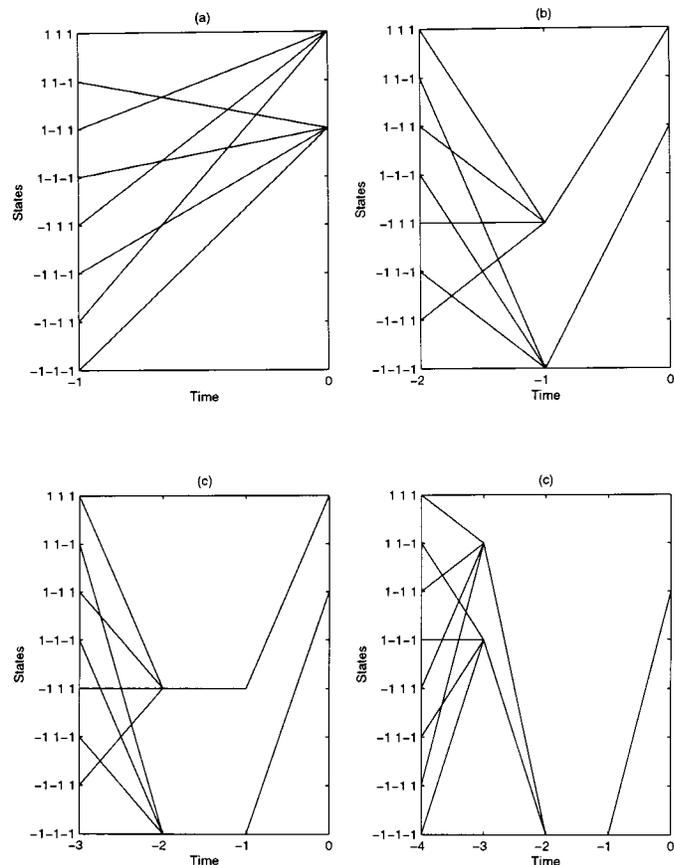

Fig. 1. Realization of the CFTP algorithm. In (a), (b), and (c), the chains did not coalesce by $t = 0$. In (d), a perfect sample was generated when the algorithm started at time -4.

starting from different states to be *sandwiched* between the two paths that started from the extreme states. Obviously, when these two extreme paths coalesce at time 0, all the other paths coalesce into the same state as well. Therefore, an efficient CFTP can be carried out by a sandwiched algorithm where only chains from the two extreme states are traced and examined for coalescence by time 0. The sandwiching effect can also be seen in Fig. 1, where the Markov chains are monotonic. It is obvious that the chains are sandwiched between the two chains initiated from [1 1 1] and $\begin{bmatrix} -1 & -1 & -1 \end{bmatrix}$. Successful attempts at applying the sandwiched CFTP have been made in analyses of binary images based on the Ising model [16], [33].

D. Gibbs Coupler

In the previous section, we have indicated that efficient CFTP algorithms exist only for monotonic Markov chains. However, in many cases, monotonic Markov chains are either difficult to construct or even do not exist for the considered problem. In [35] and [36], component-updating perfect sampling schemes have been proposed. A prominent feature of the algorithms is that they are computationally much more efficient than the CFTP when the chains are neither monotonic nor antimonotonic. Here, we refer to these schemes as the Gibbs coupler because they combine features of both the CFTP and the Gibbs sampler. Like the Gibbs sampler, the Gibbs coupler is designed for problems with large number of variables, and for its implementation, full

conditional distributions are needed. However, in contrast to the Gibbs sampler, the Gibbs coupler generates i.i.d. samples exactly from desired distributions. Now, we explain the Gibbs coupler in more detail. Let $\boldsymbol{S}^{(t)} = \{S_1^{(t)}, S_2^{(t)}, \ldots, S_N^{(t)}\}$ denote the support of $\mathbf{x}$ at time t, where $S_i^{(t)},[1]$ $i = 1, 2, \ldots, N$ represents the support of component x_i at time t. Then, at any transition t, the Gibbs coupler only examines the proposals of the components and records in $S_i^{(t)}$ $i = 1, 2, \ldots, N$ the possible values that the ith component could take from running the Gibbs sampler given $\mathbf{x}_{-i}^{(t)} \in S_{-i}^t$. In the case, when the variable space is binary, $S_i^{(t)}$ can only be $\{-1\}$, $\{1\}$, or $\{-1, 1\}$. Finally, at time $t = 0$, if $\boldsymbol{S}^{(0)}$ is shrunk to a singleton, then the unique state is a perfect sample from the desired distributions. We outline the Gibbs coupler algorithm with the following chart:

```
Gibbs coupler(T):
 t ← −T
 while t < 0
  t ← t + 1,  i ← 0
  while i <= N
   update S_i^(t) using
   p(x_i|S_1^(t),...,S_{i-1}^(t),S_{i+1}^(t-1),...,S_N^(t-1))
 if size of all S_i^(0) for i  =  1,2,...,N is
equal to 1 then
  return (S^(0))
 else
  Gibbs coupler (T' with T' > T).
```

We can see that the overall framework of the algorithm still follows that of the CFTP. This framework actually guarantees that unbiased samples from the stationary distribution are obtained if coalescence occurs. However, unlike the CFTP, the main coupling method of the Gibbs coupler is component based.

IV. MULTIUSER DETECTION BY PERFECT SAMPLING

A. Gibbs Sampling Multiuser Detector

Prior to the discussion of multiuser detection by perfect sampling, we discuss multiuser detection by Gibbs sampling. The Gibbs sampling scheme described in this section will be used to construct the Markov Chains when applying perfect sampling algorithms. As has been indicated in Section III-A, our objective is to draw samples from the posterior distribution, and once the desired number of samples is acquired, the proposed detectors are easily obtained. To implement the Gibbs sampler, we need the full conditional distributions derived from the posterior distribution. For our multiuser detection problem, they can be directly obtained from (3) as

$$
p(b_i = 1 | \mathbf{b}_{-i}), y(\tau), \tau \in [0, T]
$$

$$
\propto \exp\left(\frac{1}{2\sigma^2}\left(2A_i y_i - 2\sum_{k=1, k\neq i}^{K} A_i A_k \rho_{ki} b_k\right)\right)
$$

$$
= \left[1 + \exp\left(-\frac{2}{\sigma^2} z_i\right)\right]^{-1} \tag{10}
$$

[1] $S_i^{(t)}$ is also called a *summary state* in [37] because it summarizes the state of x_i.

where $i = 1, 2, \ldots K$, and

$$
z_i = A_i y_i - \sum_{k=1, k\neq i}^{K} A_i A_k \rho_{ki} b_k. \tag{11}
$$

The distributions described by (10) are Bernoulli distributions, and drawing a sample from them is analogous to a coin-tossing experiment, i.e.,

$$
b_i = \begin{cases} 1, & \text{if } U_i \leq p(b_i = 1 | \mathbf{b}_{-i}, y(\tau)), \tau \in [0, T] \\ -1, & \text{otherwise} \end{cases} \tag{12}
$$

where U_i is a random number from $U(0, 1)$. In fact, (12) is also the updating function of the Markov chain. Now, once we know how to draw samples from the full conditional distributions, the Gibbs sampling from (3) is simple and follows the standard iterative procedure described in Section III-B. One interesting observation here is that z_i has exactly the same expression as that of the decision statistics in a multistage detector. The only difference is that in the Gibbs sampling detector, the b_ks are the most recent samples, whereas in a multistage detector, they are the estimates obtained from the previous stage. In both cases, (11) can be considered to be a process of removing the interference of other users from the signal of the desired user. A multistage detector uses it to make a hard decision after interference cancellation, whereas the Gibbs sampler does not make any decision directly with z_i. Instead, z_i affects the posterior probabilities of the symbols that the ith user could have transmitted. After several iterations, a final decision, say, using the MAP estimator on b_i, is made. Therefore, we view the Gibbs sampler as a stochastic version of the multistage detector. Intuitively, we find the Gibbs sampler scheme more versatile than the multistage detector, although in some cases, it might not be as efficient.

One important parameter of the Gibbs sampler is the length of burn-in period since only samples after the burn-in period are considered to be samples from the true distribution. In practice, the burn-in period is usually estimated heuristically [15]. Quite often, however, such *ad hoc* estimates can be poor. The ultimate remedy to this problem is perfect sampling.

B. Multiuser Detection by the Sandwiched CFTP

It was stated before that the sandwiched CFTP is only applicable when a monotonic Markov chain exists. To apply the sandwiched CFTP, we first need to choose an algorithm to construct the Markov chains. From the discussion in the last section, the Gibbs sampler is a natural choice. Next, we need to examine if monotonicity exists on the updating function (12) of the Gibbs sampler. It turns out that monotonicity associated with (12) holds only for a system with negative and equal cross-correlations, i.e., for $\rho_{ij} = \rho < 0 \; \forall i \neq j$ and with the assumption that all the A_ks are positive. This can be seen if we impose a partial order on $\mathbf{b}$ such that $\mathbf{b} \preceq \hat{\mathbf{b}}$ if $b_i = -1$ whenever $\hat{b}_i = -1$. Then, probability (10) becomes an increasing function of $\mathbf{b}_i$ on such a partial order. As a result, it can be deduced that $\Phi(\mathbf{b}, U) \preceq \Phi(\hat{\mathbf{b}}, U)$ if $\mathbf{b} \preceq \hat{\mathbf{b}}$, which complies with the definition of monotonicity. Under this partial order, the two extreme states can easily be determined as $\mathbf{b}_{\max}^T = \begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix}$ and $\mathbf{b}_{\min}^T = \begin{bmatrix} -1 & -1 & \cdots & -1 \end{bmatrix}$. Then, the perfect samples are generated by following the scheme described in Section III-C. There, only two chains are checked for coalescence: the chains initiated from the two extreme states $\mathbf{b}_{\max}$ and $\mathbf{b}_{\min}$.

We note that sandwiched algorithms can always be used, regardless of the existence of monotonicity on variable spaces. They can be employed whenever there exist two chains that always sandwich the remaining chains during the chain propagation. By sandwiching, we mean that two of the chains always have the largest and the smallest probability of generating a 1 by the Gibbs sampling transition. We observe that with the Gibbs sampler, the chains from the all 1 states and all −1 states sandwich the probabilities of all the other chains in cases of unequal negative cross-correlation. This implies that in our problem, we can relax the condition of equal negative crosscorrelations and allow for unequal negative cross-correlations. Hence, we conclude that the sandwiched CFTP is applicable to multiuser systems with negative crosscorrelations and that the two sandwiching chains are the ones initiated from the all 1 and all −1 states.

C. Multiuser Detection by Gibbs Coupling

In the last section, it was pointed out that the sandwiched CFTP is possible for the special case of negative cross-correlations. In general, however, the ρ_{ij}s can be either positive or negative. Therefore, the sandwiched CFTP algorithm is not appropriate in a general situation. To find a more general perfect sampling algorithm, we turn our attention to the Gibbs coupler.

In [35] and [36], detailed Gibbs coupler algorithms are constructed for problems modeled by Markov random fields where neighboring properties can be used to facilitate the computation. However, these algorithms are not suited for the multiuser detection problem because here, no such neighboring properties are assumed. Hence, a special algorithm is needed to apply the Gibbs coupler to multiuser detection. We show in the sequel how it can be designed.

From the overall framework of the Gibbs coupler described in Section III-D, we can see that one critical issue is the efficient determination of the support content $\mathcal{S}_i^{(t)}$ for all is at every time instant t. Efficient updates can be achieved by introducing *sandwich distributions* at every update. At any instant of time t, *sandwich distributions* $L_i^{(t)}(\cdot)$ and $U_i^{(t)}(\cdot)$ are defined by

$$L_i^{(t)}(x_i = 1) \leq p(x_i = 1|\mathbf{y}) \leq U_i^{(t)}(x_i = 1), \quad i = 1, 2, \ldots, N \tag{13}$$

with $\mathbf{y} \in \mathcal{S}_{-i}^{(t)}$, where $\mathcal{S}_{-i}^{(t)} = \left\{ \mathcal{S}_1^{(t)}, \mathcal{S}_2^{(t)}, \ldots, \mathcal{S}_{i-1}^{(t)}, \mathcal{S}_{i+1}^{(t)}, \ldots, \mathcal{S}_N^{(t)} \right\}$ denotes the collection of supports of $\mathbf{x}_{-i}$ at time t with the individual component supports at time t being $\{-1\}$, $\{1\}$, or $\{-1, 1\}$. We notice that the definition (13) indicates that the two sandwich distributions bound all the probabilities of $x_i = 1$ for every Markov chain in between. Therefore, if the same random

number for the ith update at time t, $R_i^{(t)}$ is less than or equal to $L_i^{(t)}(x_i = 1)$ (greater than or equal to $U_i^{(t)}(x_i = 1)$), $x_i^{(t)}$ can only be equal to 1 (−1), which will be the case with all the other Markov chains. On the other hand, if $R_i^{(t)}$ is between $L_i^{(t)}(x_i = 1)$ and $U_i^{(t)}(x_i = 1)$, the value of $x_i^{(t)}$ will be uncertain. In this case, we leave the support $\mathcal{S}_i^{(t)}$ as $\{-1, 1\}$. Thus, the update of the support $S_i^{(t)}$ can be formulated as

$$\mathcal{S}_i^{(t)} = \Phi\left(\mathcal{S}_{-i}^{(t)}, R_i^{(t)}\right) = \begin{cases} \{1\}, & \text{if } R_i^{(t)} \leq L_i^{(t)}(x_i = 1) \\ \{-1\}, & \text{if } R_i^{(t)} \geq U_i^{(t)}(x_i = 1) \\ \{-1, 1\}, & \text{otherwise} \end{cases} \tag{14}$$

It is obvious that the choice of sandwich distributions will affect the rate of coalescense. If the distributions are chosen according to

$$L_i^{(t)}(x_i = 1) = \min_{\mathbf{x}_{-i}^{(t)} \in \mathcal{S}_{-i}^{(t)}} \left\{ p\left(x_i = 1|\mathbf{x}_{-i}^{(t)}\right) \right\} \tag{15}$$

and

$$U_i^{(t)}(x_i = 1) = \max_{\mathbf{x}_{-i}^{(t)} \in \mathcal{S}_{-i}^{(t)}} \left\{ p\left(x_i = 1|\mathbf{x}_{-i}^{(t)}\right) \right\} \tag{16}$$

they achieve the largest probability of coalescense, and hence, their use leads to fastest coalescense. We also notice that for as long as the calculation of the sandwich distributions is straightforward, the Gibbs coupler is computationally much more efficient than the CFTP, especially for high-dimensional state spaces. This is because the Gibbs coupler algorithm only needs to calculate the two sandwich distributions for every transition, and the CFTP must calculate the transition probabilities of all the Markov chains.

The key of the implementation of the Gibbs coupler in our multiuser detection problem is the determination of the sandwich distributions on the full conditional distributions. Recall that the full conditional distributions are defined by (10). We notice that the maximum and the minimum in (10) with respect to $\mathbf{b}_{-i}$ can easily be determined only by checking the sign of $A_i A_k \rho_{ki}$. Thus, at time t, the sandwiched distributions defined by (15) and (16) are readily derived as (17) and (18), shown at the bottom of the page, where $\beta_k = A_i A_k \rho_{ki}$, $\mathbf{I}_{i1}^{(t)} \subset \{1, 2, \ldots, i-1, i+1, \ldots, K\}$ contains the indices of the elements in $\{b_k^{(t)}\}_{k=1, k \neq i}^K$ that have not coalesced at time t, and $\mathbf{I}_{i2}^{(t)} \subset \{1, 2, \ldots, i-1, i+1, \ldots, K\}$ contaiins the indices of the remaining elements in $\{b_k^{(t)}\}_{k=1, k \neq i}^K$ (the ones that have coalesced at time t). Then, at any time t, the algorithm updates the support of the ith component according to (14), and the coalesced state at time 0 is recorded as a perfect sample from the posterior distribution (3).

$$L_i^{(t)}(b_i = 1) = \left[1 + \exp\left(\frac{2}{\sigma^2} \left(-A_i y_i + \sum_{k \in \mathbf{I}_{i1}^{(t)}} |\beta_k| + \sum_{k \in \mathbf{I}_{i2}^{(t)}} \beta_k b_k^{(t)} \right) \right) \right]^{-1} \tag{17}$$

and

$$U_i^{(t)}(b_i = 1) = \left[1 + \exp\left(\frac{2}{\sigma^2} \left(-A_i y_i - \sum_{k \in \mathbf{I}_{i1}^{(t)}} |\beta_i| - \sum_{k \in \mathbf{I}_{i2}^{(t)}} \beta_k b_k^{(t)} \right) \right) \right]^{-1} \tag{18}$$

From the above discussion, we can see that the Gibbs coupler is actually equivalent to the sandwiched CFTP in the case of negative cross-correlation. In that case, the sandwich distributions of the Gibbs coupler are always equal to the transition distributions of the two sandwiching chains in the sandwiched CFTP. Hence, the Gibbs coupler may be viewed as a general perfect algorithm for multiuser detection.

Note that the Gibbs coupler scheme introduced here is computationally very simple and much more efficient than the CFTP. However, as with other perfect sampling algorithms, the coalescing time of the Gibbs coupler is a random variable, and except for systems with negative cross-correlations where the Gibbs coupler is equivalent to the sandwiched CFTP, the convergence rate of the Gibbs coupler is usually smaller than that of the CFTP. In general, the Gibbs coupler has a relatively fast convergence if the Gibbs samplers in the corresponding CFTP algorithms mix rapidly. In [35], for problems with Markov random fields, it is argued that fast coalescence of the Gibbs coupler is expected if the interaction of the field is weak. Analogously, in multiuser detection, fast coalescence is expected when there are small cross-correlations in the system.

D. Calculating the Proposed Detectors

In the previous sections, we discussed the methods for drawing perfect samples from posterior distributions. Next, we show several ways of computing the output of Bayesian detectors.

First of all, we need to collect a certain number of samples for the Monte Carlo approximation that we suppose is equal to M. There are two ways of obtaining the samples. If i.i.d. samples are desired, each sample is drawn by an independent Gibbs coupler. Although, in general, these samples will provide more accurate results, in many cases, the computational requirement to obtain them might be overwhelming. Another way of drawing samples is to use the Gibbs coupler only as a gauge to detect convergence (burn-in) of the Gibbs sampler. More specifically, we first obtain a perfect sample from the Gibbs coupler, and then, we switch to the Gibbs sampler using the obtained perfect sample, as an initial state. Note that the subsequent M samples generated by the Gibbs sampler are also exact samples from the posterior distributions, but they are correlated. Obviously, this approach represents a tradeoff between performance and computational intensity.

Once the M samples are acquired, we can compute both the MAP and MPM detectors according to the Monte Carlo approximations. To compute the MAP detector, the posterior probability of each obtained sample is calculated, and the MAP detector is the one that yields the largest posterior probability. On the other hand, to compute the MPM detector, say, for the ith user, one only needs to consider the sample value of b_i in each sample, and the MPM detector of b_i is set to

$$\left(\hat{\mathbf{b}}_{MPM}\right)_i = \text{sgn}\left(\frac{1}{M}\sum_{t=1}^{M} b_i^{(t)}\right) \qquad (19)$$

where $b_i^{(t)}$ is the tth sample value of b_i.

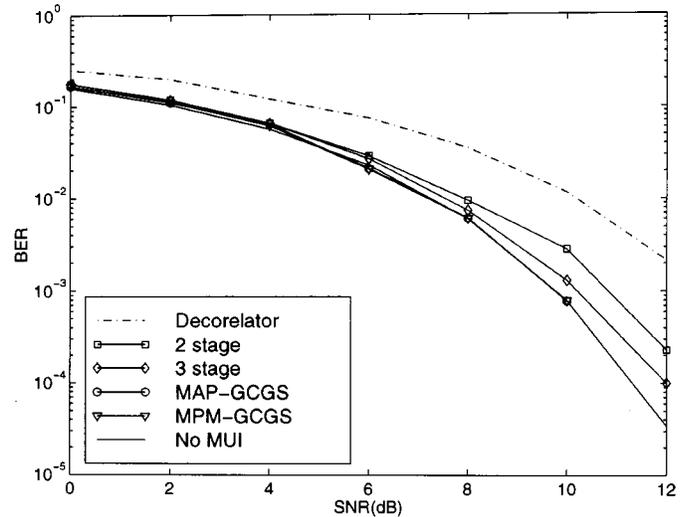With the two approaches of drawing samples and the two criteria, we can define four detectors.

Fig. 2. BERs of the proposed detectors of the first user as functions of SNR. There are 31 users with equal power.

1) MAP detector using the Gibbs coupler (MAP-GC);
2) MAP detector using the Gibbs coupler and Gibbs sampler (MAP-GCGS);
3) MPM detector based on the Gibbs coupler (MPM-GC);
4) MPM detector based on the Gibbs coupler and Gibbs sampler (MAP-GCGS).

V. SIMULATION RESULTS

We conducted several experiments that demonstrate the performance of the proposed Bayesian detectors and compare it with the performance of other existing techniques.

The first two experiments were designed to study the proposed detectors on systems with negative cross-correlations. In these experiments, a 31-bit Gold sequence was used as the spreading code of a 31-user system. As was indicated in Section IV-B, monotonic Markov chains can be constructed for these systems, and the Gibbs coupler scheme is equivalent to the sandwiched CFTP algorithm. In the first experiment, the users had equal power. Bit error rates (BERs) of the MAP-GCGS and MPM-GCGS detectors were examined under different SNRs. In formulating the decisions, the detectors used 300 samples. The results of the first user are illustrated in Fig. 2. The results of the the decorrelating detector and the multistage (two-stage and three-stage) detectors with decorrelating first stage are also presented. The theoretically attainable performance in the absence of multiuser interference is plotted as a lower bound. To compute the BER at a tested SNR, Monte Carlo trials were performed, where the number of trials was precomputed by assuring that there would be at least 200 errors among the trials with no MUI. From the plot, we see that the curves corresponding to the MAP-GCGS and the MPM-GCGS detectors almost overlap, which indicates similarity in the performance of these detectors. In addition, the BERs of the two detectors are almost the same as the BER of the single user, especially at high SNRs. Compared with other tested detectors, the proposed detectors always perform better, and the performance improvement is especially evident at higher SNRs.

TABLE I
AVERAGE COALESCING TIME T OF THE GIBBS COUPLER AT EACH TESTED
SNR IN THE FIRST EXPERIMENT

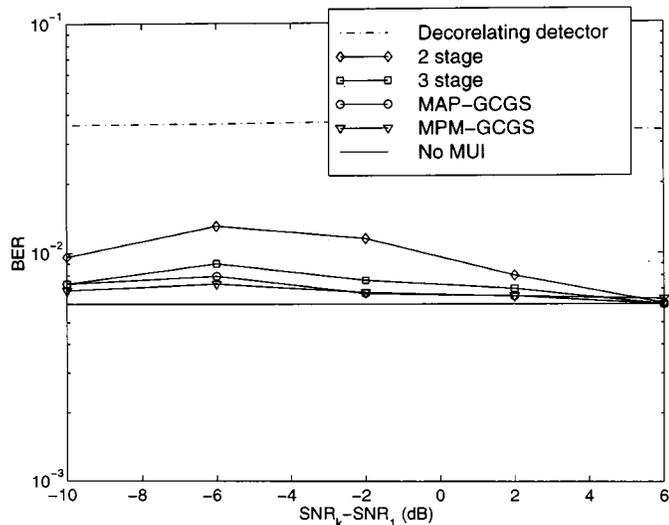| dB	0	2	4	6	8	10
T	3.731	4.064	3.8427	3.3212	2.8548	2.8376

Fig. 3. BERs of the proposed detectors on the first user as functions of power of the interference users.' There are 31 users, and SNR of the first user is 8 dB.
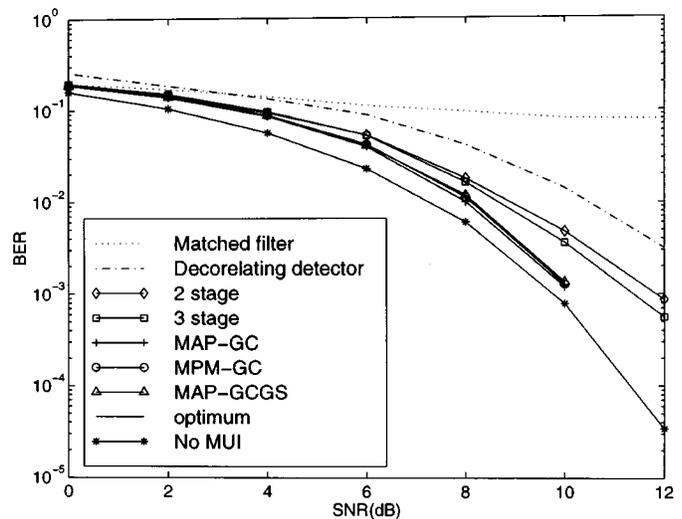

Fig. 4. BERs of the proposed detectors of the first user as functions of the SNR. There are four users with equal power. The cross-correlation matrix of the signature waveform is given by (20).

TABLE II
AVERAGE COALESCING TIME T OF THE GIBBS COUPLER AT EACH TESTED
SNR IN THE EXPERIMENT WITH A FOUR-USER-EQUAL-POWER SYSTEM

dB	0	2	4	6	8	10
T	2.3123	2.5738	3.1	5.8346	18.2608	52.5285

In Table I, we have listed the average coalescing times (expressed in steps) of the Gibbs coupler at each SNR. Surprisingly, in all the cases, these times are very small. Compared with the time needed to generate the desired 300 samples that follow the first perfect sample, the computation for detection of coalescence is considerably smaller. Therefore, in this experiment, the use of the Gibbs coupler is preferred over the Gibbs sampler because for a small price, we can obtain perfect samples from the posterior distributions.

In the second experiment, we studied the near–far effect in the above 31-user system. This time, the SNR of the first user was fixed at 8 dB, and the strength of the remaining 30 users was allowed to vary from −10 dB below to 6 dB above that of the first user. Again, the BERs of the MAP-GCGS and the MEM-GCGS detectors were examined. The results are depicted in Fig. 3 along with those of the decorrelator and the multistage detectors. Since the decorrelator is near–far resistant, we see that the BERs of the decorrelator do not change with the change of interference strength. The BERs of the decorrelator, however, are too large and too far away from the bound. Notice that the other tested detectors dramatically outperform the decorrelator. What is more, they all approach the single-user lower bound when the powers of the interference users are at 6 dB. The two Bayesian detectors have lower BERs than the two multistage detectors throughout the tested region. Their performance gain over the multistage detectors is maximal when the interference users have a power of −6 dB below that of the first user.

In the next two experiments, we tested the proposed detectors in a scenario that would allow for gaining insight into asynchronous and bandwidth efficient systems. We adopted

the system setting used in [8, ex. 2]. To be specific, there were four users in the system, and the cross-correlation matrix of the corresponding signature waveforms was given by

$$\boldsymbol{\rho} = \frac{1}{7} \begin{bmatrix} 7 & -1 & 3 & 3 \\ -1 & 7 & 3 & -1 \\ 3 & 3 & 7 & -1 \\ 3 & -1 & -1 & 7 \end{bmatrix}. \tag{20}$$

Throughout these two experiments, 11 samples were recorded for the Monte Carlo computation. First, there were four users with equal power. We evaluated the BERs of MAP-GC, MPM-GC, and MAP-GS of the first user for various SNRs. The results are shown in Fig. 4. In the figure, we also plotted the curve corresponding to the optimum MAP detector, which was computed by exhaustive search. The results demonstrate that the BER of the MAP-GC detector is slightly better than that of the MPM-GC and MAP-GCGS detectors and that it almost overlaps with that of the optimum detector. Again, even though the multistage detectors are clearly better than the decorrelator, their performance is inferior to that of the proposed detectors at high SNRs. For example, we observe that for the same noise level, the proposed detectors have a gain in signal power of 0.5 dB at BER of 10^{-2} and 1.5 dB around BER of 10^{-3}. Similarly, we also recorded the average convergence times of the Gibbs coupler at various SNRs, and they are presented in Table II. It is interesting to notice that unlike in the first experiment, the time T increases with the increase of SNR. To explain the effect, we need to consider the relationship between the coalescing times and the users' cross-correlations, as well as the effect of the noise on this relationship. First, recall that the Gibbs coupler has slower coalescing times on systems with higher
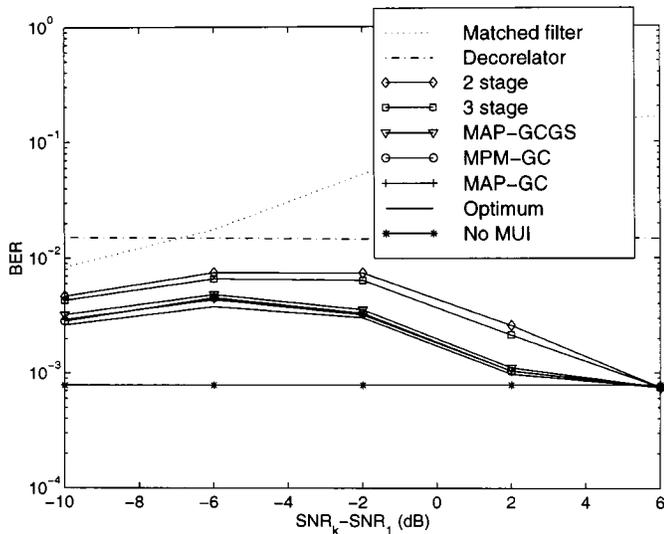
Fig. 5. BERs of the proposed detectors of the first user as functions of the SNR difference between the interference users and the first user. There are four users in the system, and the SNR of the first user is fixed at 10 dB. The cross-correlation matrix of the signature waveform is given by (20).
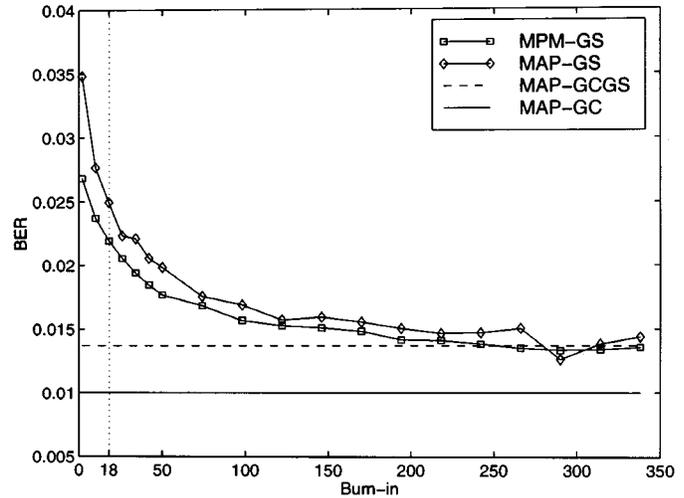

Fig. 6. BERs of the detectors obtained by the Gibbs sampler as functions of burn-in period. There are four equal-power users at 8 dB. The cross-correlation matrix of the signature waveform is given by (20).

cross-correlations between the users' signature waveforms. Second, note that the noise would make the cross-correlations less effective. This means that with the increase of noise level, the detectors tend to underestimate the cross-correlations of the system. In the extreme, when the noise level is very high, the detectors would simply consider the multiusers' signals as independent. In this experiment, when the SNR is low, the detectors simply see low cross-correlations in the system, and therefore, coalescing is relatively fast. With the increase of SNR, the high cross-correlations of the users' waveforms become more effective in slowing down the coalescence.

In the next experiment, we tested the near–far effect on this four-user system. Again, we fixed the SNR of the first user, this time at 10 dB, and let the SNRs of the remaining users vary. The experimental results are illustrated in Fig. 5. We see that with the increase of interference strength, the proposed detectors and the multiuser detectors have better performance. In addition, they are all able to achieve the single user bound when the SNR of the interference is 6 dB above that of the first user. It is clear that the multistage detectors improve significantly over the matched filter and the decorrelator. In addition, the proposed detectors provide further reduction of BERs, and they almost perform as the optimum detector.

We then wanted to compare the raw Gibbs sampler detectors with our proposed detectors. The objective of this experiment was to demonstrate and further stress the advantages of the perfect sampling solutions over the Gibbs sampler from both computational and performance perspectives. In the experiment, we continued to use the above four-user system setting, but we focused on a specific scenario of equal-power users with SNR fixed at 8 dB. We applied the Gibbs sampler on the system and computed the MAP and the MPM detectors using 11 samples after burn-in. Note that the burn-in period is vaguely defined and that Markov chains converge gradually. For convenience, we denoted the two detectors by "MAP-GS" and "MPM-GS," respectively. In Fig. 6, we plotted the BERs of the MAP-GS and

the MPM-GS detectors as functions of the "burn-in" period. For comparison, we also displayed the BERs of the MAP-GCGS and the MPM-GCGS. Several observations can be made from the figure. First of all, the BERs of both the MAP-GS and the MPM-GS detectors approach that of the MAP-GCGS detector as the burn-in period increases. The reason is that with the increase of the burn-in, the subsequent samples are closer to the true posterior distribution, which in turn results in better accuracy in the calculation of the MAP and the MPM detectors. Second, it takes both detectors almost 300 burn-in iterations to approach the performance of the MAP-GCGS detector. Clearly, this time is much longer than the average coalescing time of the Gibbs coupler that is only about 18 iterations in duration. Can we then claim that the actual convergence time of the Gibbs sampler is longer than the coalescing time of the Gibbs coupler? Our answer is no. Rather, we see an equivalence between the convergence time of the Gibbs sampler and the coalescing time of the Gibbs coupler. We elaborate on this point in the following. Since Monte Carlo trials are run to compute a BER, intuitively, we think that there is a strong correlation between the changes in BER and the actual number of converged trials. To be specific, the more trials converged, the better the BER of the detector. Now, from the average coalescing time of the Gibbs coupler, we conjecture that there was a considerable amount of trials that converged by the 18th iteration. As a result, there would be a big improvement in BER if the detectors used generated samples after the 18th iteration. We find that our conjecture agrees with the scenario shown in Fig. 6, where big improvements in BERs of the two Gibbs sampler detectors indeed happen before and around the 18th burn-in iteration. After 18 iterations of burn-in, the BER curves decrease rather gradually and stretch on until about 300 iterations, when the change becomes close to steady. This indicates that there was also a great number of trials that converged after 18 iterations and that by the 300th iteration, almost all the trials converged. Even though we cannot prove our claim, it is quite reasonable to think that the actual mean convergence time of the Gibbs sampler is not 300 iterations and that it is comparable with the mean coalescing time of the Gibbs cou-

pler. Then, why does it take 300 burn-in steps for all the Gibbs sampler trials to approach the performance of the MAP-GCGS detector? The reason is the way the *ad hoc* burn-in is used in the experiment. There, as well as in practice, once an *ad hoc* burn-in period is determined, it is applied to all the Gibbs sampler trials. These trials are independent of each other, and some of them may converge faster, whereas others converge slowly, and the time of convergence is longer than the applied burn-in period. However, to achieve an equivalent performance as with converged samples, this "one-size-fits-all" burn-in period must be at least as long as the burn-in period of the trial with slowest convergence. We have indicated that in our experiment, 300 is the approximate duration of the burn-in period of the trials with slowest convergence. Consequently, the two Gibbs sampler detectors require a burn-in period of 300 iterations for each trial to approach the performance of the Gibbs coupler detector. As has been stressed throughout this paper, the determination of the burn-in period in Gibbs sampling with certainty is an unsolved problem. With *ad hoc* approaches, we will either overestimate or underestimate the burn-in period. With an underestimated burn-in period, the performance of the corresponding Monte Carlo approximation is degraded. On the other hand, as is demonstrated in the experiment, to guarantee good performance, we will have to choose a much overestimated burn-in interval for most of the trials. Apparently, this entails a waste of computation for most of the trials. In this respect, the perfect sampling algorithms like the Gibbs coupler have clear advantage over Gibbs sampling because they are able to determine the coalescence for each trial separately. Therefore, they not only attain better performance but may also be computationally more efficient than the Gibbs sampler.

Now, we turn to the third observation. We see that the two detectors based on Gibbs sampling cannot approach the BER of the MAP-GC detector, which has a smaller BER than that of the MAP-GCGS detector. The difference between the two detectors based on Gibbs coupling is that the MAP-GC detector uses independent samples, and the MAP-GCGS detector operates with correlated samples that are generated from a single Markov chain. The problem of using samples from a single Markov chain is that once the chain enters a high density region, it tends to stay there for a long period before it moves out of the region. As a result, the collected samples could all come from the same high-density region. If this high-density region is near a local optimum but not within the maximum density region as desired, the subsequent Monte Carlo approximation for the MAP or the MPM detectors would be biased. One solution to the problem is to increase the sample size so that the number of samples can be large enough to include the maximum density region. Apparently, in our case, the detectors based on Gibbs sampling used more than 11 samples to match the performance of the MAP-GC. This observation suggests that detectors based on Gibbs coupling can achieve the same performance as the detectors based on Gibbs sampling with fewer samples, which is clearly another advantage.

In the last experiment, we examine the effect of the cross-correlation on the coalescing time. Since the coalescing time of the Gibbs coupler is a random variable and very difficult to analyze, we studied it through examples. In the experiment, we fixed the
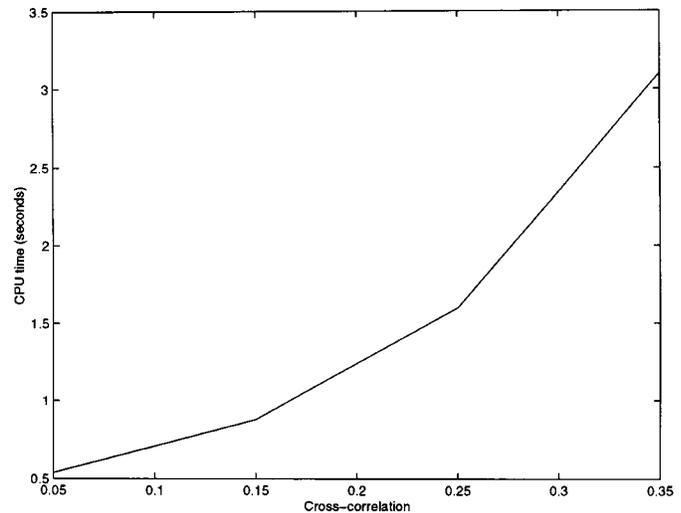

Fig. 7. Plot of CPU time for computing the MAP-GC detector as a function of different crosscorrelations. There are five users.

number of users to $K = 5$, and set equal cross-correlation between the waveforms of the users' signatures. We recorded the mean CPU time for computing the output of the MAP-GC detector as a function of increase cross-correlations. The results are displayed in Fig. 7. We notice that the CPU time increases rapidly with the increase of cross-correlations. This result implies that the detectors computed by the Gibbs coupler are more fit for systems with small cross-correlations.

VI. CONCLUSION

We applied perfect sampling algorithms including the sandwiched CFTP and the Gibbs coupler to compute the MAP and MPM detectors for synchronous CDMA signals. We introduced several methods for drawing samples and discussed the corresponding detectors. Numerical simulation results for equal-power and near-far effect scenarios on a 31-user system demonstrated better performance of the proposed detectors over the decorrelator and multistage detectors. Simulations also showed better performance of the proposed detectors in potentially asynchronous channels. An experiment was also conducted to illustrate several advantages of Gibbs coupling over Gibbs sampling when used for multiuser detection. Finally, we briefly addressed computational issues of the proposed methods. We indicated that the computational time of the Gibbs coupler algorithm increases rapidly with the cross-correlations between the waveforms of the users' signatures.

REFERENCES

[1] S. Verdú, *Multiuser Detection*. New York: Cambridge Univ. Press, 1998.
[2] T. S. Rappaport, *Wireless Communications: Principles and Practice*. Upper Saddle River, NJ: Prentice-Hall, 1996.
[3] A. J. Viterbi, *CDMA: Principles of Spread Spectrum Communications*. Reading, MA: Addison-Wesley, 1995.

[4] M. K. Tsatsanis and G. B. Giannakis, "Optimal decorrelating receivers for DS-CDMA systems: A signal processing framework," *IEEE Trans. Signal Processing*, vol. 44, pp. 3044–3055, Dec. 1996.

[5] R. Lupas and S. Verdu, "Linear multiuser detectors for synchronous code-division multiple-access channels," *IEEE Trans. Inform. Theory*, vol. 35, pp. 123–136, Jan. 1989.

[6] P. Patel and J. Holtzman, "Analysis of a simple successive interference cancellation scheme in a DS/CDMA system," *IEEE J. Select. Areas Commun.*, vol. 12, pp. 796–807, June 1994.

[7] M. K. Varanasi and B. Aazhang, "Near-optimum detection in synchronous code-division multiple-access systems," *IEEE Trans. Commun.*, vol. 39, pp. 725–736, May 1991.

[8] A. Duel-Hallen, "Decorrelating decision-feedback multiuser detector for synchronous code-division multiple-access channel," *IEEE Trans. Commun.*, vol. 41, pp. 285–290, Feb. 1993.

[9] M. J. Borran and M. Nasiri-Kenari, "An efficient detection technique for synchronous CDMA communication systems based on the expectation maximization algorithm," *IEEE Trans. Veh. Technol.*, vol. 49, pp. 1663–1668, Sept. 2000.

[10] C. Ergun and K. Hacioglu, "Multiuser detection using a genetic algorithm in CDMA communications systems," *IEEE Trans. Commun.*, vol. 48, pp. 1374–1383, Aug. 2000.

[11] S. H. Yoon and S. S. Rao, "Annealed neural network based multiuser detector in code division multiple access communications," *Proc. Inst. Elect. Eng. Commun.*, vol. 147, no. 1, pp. 57–62, Feb. 2000.

[12] W. R. Gilks, S. Richardson, and D. J. Spiegelhalter, "Introducing Markov chain Monte Carlo," in *Markov Chain Monte Carlo in Practice*. London, U.K.: Chapman & Hall, 1996.

[13] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-6, pp. 721–741, 1984.

[14] W. K. Hastings, "Monte Carlo sampling methods using Markov chains and their applications," *Biometrika*, vol. 57, pp. 97–109, 1970.

[15] X. Wang and R. Chen, "Adaptive Bayesian multiuser detection for synchronous CDMA with Gaussian and impulsive noise," *IEEE Trans. Signal Processing*, vol. 47, pp. 2013–2027, July 2000.

[16] J. G. Propp and D. B. Wilson, "Exact sampling with coupled Markov chains and applications to statistical mechanics," in *Random Structures Algor.*, 1996, vol. 9, pp. 223–252.

[17] W. S. Kendall, "Perfect simulation for the area-interaction point process," in *Probability Toward 2000*, C. Heyde and L. Accardi, Eds. New York: Springer-Verlag, 1998, pp. 218–234.

[18] D. J. Murdoch and P. J. Green, "Exact sampling form a continuous state space," *Scand. J. Statist.*, vol. 25, pp. 483–502, 1998.

[19] P. J. Green and D. J. Murdoch, "Exact sampling for Bayesian inference: Toward general purpose algorithms," in *Bayesian Statistics*, A. P. Dawid, J. M. Bernardo, J. O. Gerger, and A. F. M. Smith, Eds. Oxford, U.K.: Oxford Univ. Press, 1999, vol. 6, pp. 301–321.

[20] J. G. Proakis, *Digital Communication*, 3rd ed. New York: McGraw-Hill, 1995.

[21] T. S. Rappaport, Ed., *Smart Antennas: Adaptive Arrays, Algorithms and Wireless Position Location*. Piscataway, NJ: IEEE, 1998.

[22] J. C. Liberti Jr. and T. S. Rappaport, *Smart Antennas for Wireless Communications: IS-95 and Third Generation CDMA Applications*. Englewood Cliffs, NJ: Prentice-Hall, 1999.

[23] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[24] G. Winkler, *Image Analysis, Random Fields and Dynamic Monte Carlo Methods*. Berlin, Germany: Springer-Verlag, 1995.

[25] C. P. Robert and G. Casella, *Monte Carlo Statistical Methods*. New York: Springer, 1999.

[26] M. A. Tanner, *Tools for Statistical Inference*. New York: Springer Verlag, 1996.

[27] S. Richardson, W. R. Gilks, and D. J. Spiegelhalter, *Markov Chain Monte Carlo in Practice*. London, U.K.: Chapman & Hall, 1996.

[28] S. P. Brooks, "Markov chain Monte Carlo method and its application," *Statistician*, vol. 47, no. 1, pp. 69–100, 1998.

[29] A. E. Gelfand and A. F. M. Smith, "Sampling-based approaches to calculating marginal densities," *J. Amer. Statist. Assoc.*, vol. 85, pp. 398–409, 1990.

[30] W. H. Wong, J. Liu, and A. Kong, "Covariance structure and convergence rate of the Gibbs sampler with various scans," *J. R. Statist. Soc. B*, vol. 57, pp. 157–169, 1995.

[31] G. Casella and E. I. George, "Explaining the Gibbs sampler," *Amer. Statist.*, vol. 46, pp. 167–174, 1992.

[32] J. G. Propp and D. B. Wilson, "Coupling from the paset: A user's guide," in *Series in Discrete Mathematics and Theoretical Computer Science*, D. Aldous and J. G. Propp, Eds: Amer. Math. Soc., 1998, vol. 41, pp. 181–192.

[33] M. Fismen, "Exact Simulation Using Markov Chains," Dept. Math., Norwegian Univ. Sci. Technol., Oslo, Tech. Rep. 6/98, 1997.

[34] E. Thönnes. (2000) A Primer on Perfect Simulation. [Online]. Available: http://dimacs.rutgers.edu/~dbwilson/exact

[35] O. Häggström and K. Nelander, "On exact simulation of Markov random fields using coupling from the past," *Scand. J. Statist.*, vol. 26, no. 3, pp. 395–411, 1999.

[36] M. Huber, "Exact sampling and approximate counting techniques," in *Proc. 30th Annu. ACM Symp. Theory Comput.*, 1998, pp. 31–40.

[37] A. M. Childs, R. B. Patterson, and D. J. C. MacKay, "Exact sampling from nonattractive distributions using summary states," *Phys. Rev. E*, vol. 63, 2001.

Yufei Huang was born in Shanghai, China, in 1973. He received the B.S. degree in applied electronics from Northwestern Polytechnic University, Xi'an, China, in 1995 and the M.S. and Ph.D. degrees in electrical engineering from the State University of New York (SUNY) at Stony Brook in 1997 and 2001, respectively.

He is now a Post-Doctoral Researcher with the Department of Electrical and Computer Engineering, SUNY Stony Brook. His current research interests are in the theory of Monte Carlo methods and their applications to array processing and wireless communications.

Petar M. Djurić (SM'99) received the B.S. and M.S. degrees in electrical engineering from the University of Belgrade, Belgrade, Yugoslavia, in 1981 and 1986, respectively, and the Ph.D. degree in electrical engineering from the University of Rhode Island, Kingston, in 1990.

From 1981 to 1986, he was Research Associate with the Institute of Nuclear Sciences, Belgrade. Since 1990, he has been with the State University of New York (SUNY) at Stony Brook, where he is Professor with the Department of Electrical and Computer Engineering. He works in the area of statistical signal processing, and his primary interests are in the theory of modeling, detection, estimation, and time series analysis and its application to a wide variety of disciplines, including telecommunications, biomedicine, and power engineering.

Prof. Djurić has served on numerous Technical Committees for the IEEE and SPIE and has been invited to lecture at universities in the United States and overseas. He was Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING and is currently Treasurer of the IEEE Signal Processing Conference Board. He is also Vice Chair of the IEEE Signal Processing Society Committee on Signal Processing Theory and Methods and a Member of the American Statistical Association and the International Society for Bayesian Analysis.