A High-Efficiency Bayesian Receiver for Multiuser Communications

Zhi Tian

Abstract—An efficient sequential Monte Carlo inference technique is proposed for MAP symbol detection and channel estimation in multiuser wireless communications. Intelligent dynamic sample allocation strategy is designed to optimize the simulation efficiency, expressed as the probability of correct selection of the MAP decisions within a given computing budget.

Keywords: multiuser detection, Bayesian inference, Monte Carlo simulations, simulation efficiency

I. INTRODUCTION

Stochastic Bayesian detection has recently emerged as a competitive receiver design paradigm for wireless communications applications [1] [2]. It uses sequential Monte Carlo simulations to perform Bayesian inference of a probabilistically modeled communication system so as to obtain the maximum a posterior (MAP) symbol detection and/or channel estimation results. The Monte Carlo sampling approach has great intuitive appeal with sound theoretical footing, but tends to be slow to converge. It is imperative to develop efficient sampling strategies in order for this class of methods to find more practical applications.

For a multi-dimensional optimization problem such as encountered in symbol detection for multiuser communications, it has been observed [3] that the random point set generated by a posterior distribution often shows a clustered point structure in a high dimensional sample space. When a randomly drawn sample falls outside the critical cluster, it has little contribution in the final estimation, therefore becomes an ineffective sample. If there are too many ineffective samples, the Monte Carlo procedure becomes inefficient. In practice, it is easy to generate samples from the prior, but difficult to generate samples from the posterior distribution. The potentially huge waste of samples in unlikely spaces explains the slow convergence of Monte Carlo methods. The key to speed up the convergence of the sampling process lies in quickly finding the proper “information-packed” sample space.

In this paper, we develop an adaptive sampling method in which the sample allocation process is optimized for efficient MAP detection of finite alphabet data. Different from other popular sampling techniques such as importance sampling [4] and Markov Chain Monte Carlo methods [5], the underlying philosophy is to obtain good estimates through intelligent sample allocation while the value of an estimate is still poor. A dynamic simulation scheme is designed where previous samples are used to guide the future sampling strategy in the quest to optimize the simulation efficiency, expressed as the probability of correct selection of the true MAP estimates in the limit of time. By focusing on the “important” space, we are able to sort out the top hypotheses quickly, without concerning about the accuracy of the estimated posterior probability densities of the target variables. It is then demonstrated that the optimized adaptive sampling (OAS) method can be effectively applied to wireless communication systems for high-efficiency symbol detection and channel estimation. Its salient convergence behavior, along with the near-optimal performance inherent to Bayesian estimation, enhances the role of stochastic Bayesian inference in solving computationally demanding blind multiuser detection problems.

II. PROBLEM STATEMENT

Consider a multiple-access digital communication system with K users, employing normalized modulation waveforms h_1, h_2, \ldots, h_K through a channel with additive white Gaussian noise. The transmitted data symbols b_k, k = 1, \ldots, K, take values from a g-ary finite alphabet set \mathcal{B} = \{\beta_1, \ldots, \beta_g\}. The received baseband signal samples can be modeled as

\begin{equation}
\mathbf{x}(n) = \sum_{k=1}^{K} A_k \mathbf{h}_k b_k(n) + \mathbf{w}(n), \quad n = 1, \ldots, N,
\end{equation}

where N is the block size of the data symbols per user, A_k is the amplitude of the kth user, and \mathbf{w}(n) is a zero-mean white noise vector with variance \sigma_w^2. It is assumed that the receiver has the knowledge of the L-dimensional signature vectors \mathbf{h}_k of all active users. Equation (1) can be equivalently written as

\begin{equation}
\mathbf{x}(n) = \mathbf{H A b}(n) + \mathbf{w}(n) = \mathbf{H b}(n) + \mathbf{w}(n),
\end{equation}

where \mathbf{H} \triangleq [\mathbf{h}_1 \mathbf{h}_2 \cdots \mathbf{h}_K], \quad \mathbf{a} \triangleq [A_1 A_2 \cdots A_K]^T, \quad \mathbf{A} \triangleq \text{diag}\{\mathbf{a}\}, \quad \text{and} \quad \mathbf{b}(n) \triangleq [b_1(n), b_2(n), \ldots, b_K(n)]^T, \quad \mathbf{B}(n) \triangleq \text{diag}\{\mathbf{b}(n)\}, \quad \text{for} \quad n = 1, \ldots, N.

Denote the data observations by \mathbf{X} \triangleq [\mathbf{x}(1) \mathbf{x}(2) \cdots \mathbf{x}(N)]. A simulation-based optimization problem for detecting b_k(n) can be defined in the maximum a posterior sense by:

\begin{equation}
\hat{b}_k(n) = \max_{\beta_j \in \mathcal{B}} \mu_{k,j}(n) \triangleq P[b_k(n) = \beta_j | \mathbf{X}] = E_{\mathbf{m}} \{\lambda_{k,j}(m)\}, \quad k = 1, \ldots, K; \quad n = 1, \ldots, N,
\end{equation}

where the performance criterion \mu_{k,j}(n) is the expectation of the sample performance \lambda_{k,j}^{(m)}(n) \triangleq P\{\hat{b}_k(n) = \beta_j | \mathbf{X}\}, assuming no knowledge of the channel amplitudes \{A_k\}_{k=1}^K and the noise level \sigma_w^2. \lambda_{k,j}^{(m)}(n) is an ergodic random sample of \hat{b}_k(n) generated from the finite alphabet set \mathcal{B} at the mth simulation.

The standard approach is to estimate \mu_{k,j}(n) by the sample mean performance measure

\begin{equation}
\bar{\mu}_{k,j}(n) \triangleq \frac{1}{M_{k,n,j}} \sum_{m=1}^{M_{k,n,j}} \lambda_{k,j}^{(m)}(n),
\end{equation}

where \( M_{k,n,j} \) is the number of simulations in which \( \hat{b}_k(n) = \beta_j \). The total number of simulations is \( M = \sum_{n=1}^{N} M_{k,n,j} \forall k, n \). An efficient sampling strategy will control the sample allotment numbers \( M_{k,n,j} \) in an effective manner so that the correct MAP detection in (3) can be reached by minimizing over (4) with the least computing budget M.
III. Algorithm Description

For clarity, we first describe a high-efficiency optimized adaptive sampling (OAS) method for symbol detection, assuming known channel parameters \( \mathbf{a} \) and \( \sigma_w^2 \). This scheme is then combined with Bayesian channel estimation [1] for blind optimum multiuser detection.

A. Adaptive Sampling with Optimum Sample Allocation

A key question to expedite the search for the maximum a posteriori decision is: where are the most “effective” samples located in a huge sample space based on the feedback from available observations? We propose an optimized adaptive sampling method to answer this question. It iterates the sample allotment to the maximum a posteriori estimates through an optimization formulation. The solutions to this formulation give guidance on the sample generating process by indicating where samples should be distributed so as to be more effective. The OAS method is documented in detail in [6], and summarized below.

With the knowledge of the channel information \( \mathbf{a} \) and \( \sigma_w^2 \), it can be readily shown that

\[
\lambda_{k,m}^{(n)}(n) \propto \exp\left(-\frac{1}{2\sigma_w^2} \| x(n) - H\mathbf{A}_{k,m}^{(m)}(n) \|^2 \right). \tag{5}
\]

Based on these simulation performance measures, the observed MAP solution for \( b_k(n) \) is given by \( \beta_l = \arg \max_{\beta_j \in B} \mu_{k,j}(n) \). The convergence of a sampling method towards the exact inference solution on \( b_k(n) \) is indicated by the probability of \( \beta_l \) being the optimal solution to (3), denoted by the probability of correct selection (\( P_S \)):

\[
P_S \triangleq \Pr(\beta_l \text{ maximizes } \Pr(b_k(n) = \beta_l|X)).
\]

To optimize the convergence rate, we wish to choose \( \{M_{k,n,j}\}_{j=1}^{q} \) such that \( P_S \) is maximized, subject to a limited computing budget \( M \),

\[
\max_{\{M_{k,n,j}\}_{j=1}^{q}} \quad P_S \left( \lambda_{k,m}^{(n)}(n) \right)_{m=1}^{M_{k,n,1}}, \ldots, \left( \lambda_{k,m}^{(n)}(n) \right)_{m=1}^{M_{k,n,q}} \quad \text{subject to} \quad \sum_{j=1}^{q} M_{k,n,j} = M.
\]

To find an explicit expression for \( P_S \) as a function of \( M_{k,n,j} \), we define a random variable \( \tilde{\mu}_{k,j}(n) \) whose posterior distribution is the posterior distribution of \( \mu_{k,j}(n) \). Assuming a non-informative prior distribution on \( \mu_{k,j}(n) \), it can be shown that the posterior distribution of \( \mu_{k,j}(n) \) is [7]

\[
\tilde{\mu}_{k,j}(n) \sim \mathcal{N}\left( \tilde{\mu}_{k,j}(n), \frac{\sigma^2_{k,n,j}}{M_{k,n,j}} \right), \tag{7}
\]

where \( \sigma^2_{k,n,j} \) is the sample performance variance for \( b_k(n) \),

\[
\sigma^2_{k,n,j} \triangleq \frac{1}{M_{k,n,j}-1} \sum_{m=1}^{M_{k,n,j}} \left( \lambda_{k,m}^{(n)}(n) - \mu_{k,j}(n) \right)^2. \tag{6}
\]

A lower bound for the probability of correct selection is given by [6] [8]

\[
P_S = \Pr \left( \bigcup_{j=1}^{q} \left( \mu_{k,j}(n) - \tilde{\mu}_{k,j}(n) > 0 \right) \right), \tag{8}
\]

\[
\geq 1 - \sum_{j=1}^{q} \Pr \left( \mu_{k,j}(n) < \tilde{\mu}_{k,j}(n) \right), \tag{9}
\]

\[
= 1 - \sum_{j=1}^{q} Q \left( \frac{\delta_{k,n,(l,j)}}{\sigma_{k,n,(l,j)}} \right) \triangleq P_S, \tag{10}
\]

where \( Q(\cdot) \) is the error function, \( \delta_{k,n,(l,j)} \triangleq \tilde{\mu}_{k,j}(n) - \mu_{k,j}(n) > 0 \) and \( \sigma^2_{k,n,(l,j)} \triangleq \frac{\sigma^2_{k,n,l}}{N_{k,n,l}} + \frac{\sigma^2_{k,n,l}}{N_{k,n,j}} + \frac{\sigma^2_{k,n,l}}{N_{k,n,l}} \). The lower bound \( P_S \) denotes the approximate probability of correction selection and can be computed quickly from samples \( \{\lambda_{k,m}^{(n)}(n)\}_{m=1}^{M} \) in a closed form.

Substituting \( P_S \) into the objective function in (6), the optimal sample allocation rule asymptotically becomes (as \( N \rightarrow \infty \))[6] [8]

\[
M_{k,n,l} = \sigma_{k,n,l} \sqrt{\sum_{j=1}^{q} \frac{M_{k,n,j}}{\sigma_{k,n,j}}}, \tag{11}
\]

\[
M_{k,n,i} = \left( \frac{\sigma_{k,n,i}}{\sigma_{k,n,j}} \frac{\delta_{k,n,(l,j)}}{\delta_{k,n,(l,i)}} \right)^2, \quad i \neq j, l. \tag{12}
\]

For the binary signaling case of \( q = 2 \), the optimum strategy is simplified to \( M_{k,n,1} = \frac{\sigma_{k,n,1}}{\sigma_{k,n,2}} M_{k,n,2} \).

Based on the above sample allocation rule, a sequential sampling algorithm is developed for efficient Bayesian inference [6]. The so-called optimized adaptive sampling algorithm involves two recursive steps: (i) identify the index \( l \) of the current best estimate, and update the sample means \( \tilde{\mu}_{k,n,j} \) variances \( \sigma^2_{k,n,j} \) and \( \delta_{k,n,(l,j)} \) for \( j = 1, \ldots, q \); (ii) decide the dynamic sample allotment numbers \( M_{k,n,j} \) for every block of new samples, using (11) and (12). The process iterates until the total simulations \( M \) is exhausted, or \( \beta_l \) is stabilized. Such an adaptive sampling process aggressively biases samples towards making the most significant contribution to the MAP solutions. It is expected to converge faster than other Monte Carlo methods that emphasize the accuracy of the posterior densities of all hypotheses.

B. Bayesian Multiuser Detector in Gaussian Noise

For blind multiuser detection, we use a Bayesian receiver scheme similar to that in [1] for concurrent channel parameter estimation, but replace the Gibbs-sampler-based symbol detection therein by the OAS method, resulting in a different inference and detection structure. To make the optimized sample allocation process more convenient for inference in communication systems, we make a slight modification to the OAS method. Instead of deterministically allocating discrete sample numbers, we draw random samples \( b_k(n) \) based on instantaneous probability densities \( \{\pi_{k,j}(n)\}_{j=1}^{q} \), which are periodically updated with every \( \Delta M \) increment of simulations via \( \sigma^2_{k,j}(n) \propto M_{k,n,j} \). The overall detection algorithm is described below, in which \( \mathbf{B}^\dagger \) denotes the symbol estimates available at
the current iteration.

For $m = 1 : M$

1. Draw random samples of $\mathbf{a}^{(m)}$ from the conditional probability $p(\mathbf{a}|\mathbf{B}^c, \mathbf{X})$ given by [1]

$$p(\mathbf{a}|\mathbf{B}^c, \mathbf{X}) \propto N(\mathbf{a}; \mathbf{\Sigma}_{a}) I_{\{\mathbf{a} > 0\}}$$

(13)

with

$$\mathbf{\Sigma}_{a}^{-1} = \mathbf{\Sigma}_{0}^{-1} + \frac{1}{\sigma_a^2} \sum_{n=1}^{N} \mathbf{B}^c(n) \mathbf{H}^H \mathbf{B}^c(n)$$

and

$$\mathbf{a} \sim \mathcal{N}(\mathbf{a}_{0}; \mathbf{\Sigma}_{a})$$

$$\mathbf{x} \sim \mathcal{N}(\mathbf{B}^c(n) \mathbf{H}^H \mathbf{x}(n); \mathbf{\Sigma}_{a})$$

2. Draw data samples $\mathbf{b}_{k}^{(m)}(n)$ from the instantaneous densities $p_{k,j}(n)$, and compute the sample performance measures $\lambda_{k,j}^{(m)}(n)$ from (3), $k = 1, \ldots, K, n = 1, \ldots, N$.

3. If the iteration increment reaches $\Delta M$, compute a new set of sample allocation numbers $\{M_{k,n,j}\}$ using (11) and (12) for the $q$-ary signaling case or the simplified version for the binary case, and update the instantaneous data densities according to $p_{k,j}(n) \propto M_{k,n,j}$, $j = 1, \ldots, q$. Meanwhile, without additional computation, $\mathbf{b}_{k}^c(n)$ is updated by $\mathbf{b}_{k}^c(n)$ that maximizes the current sample means $\left\{\mathbf{b}_{k,j}(n)\right\}_{j=1}^{q}$.

This OAS-based Bayesian multiuser detector is similar to those with MCMC inference in that current samples are biased by the information drawn from previous samples. However, it has several advantages over the Gibbs-sampler-based detection [1]:

- No “burn-in” time for convergence. As indicated in (4), all samples are equally used in Bayesian inference.
- The use of sequential symbol estimation in step 3 instead of sequential sample generation reduces the costly updating of the sample distributions at every iteration.
- The use of the second-order statistic $\sigma_{k,n,j}^2$ in sample allocation makes it more “intelligent” than other adaptive sampling approaches where only the first-order statistic is involved.

IV. SIMULATIONS

Consider a five-user synchronous CDMA channel with a processing gain of $L = 10$. The system setup is the same as that in [1], where the user-spreading waveform matrix $\mathbf{H}$ and the non-informative prior distributions of the channel parameters $\mathbf{a}$ and $\sigma_a^2$ are specified. The data block size of each user is $N = 256$.

To illustrate the convergence behavior of the proposed adaptive sampling method, we plot the estimates of the parameters $b_3(50)$ and $b_5(250)$, along with the results from a Gibbs sampler, in Figure 1(a). The marginal posterior distributions of the unknown parameters $A_3$ and $A_4$ in the steady state are illustrated by the corresponding histograms, shown in Figure 1(b). The histograms are based on 500 samples collected after the initial 50 iterations. It is shown that the OAS algorithm converges very fast. The detection performance of the proposed blind adaptive multiuser detector is illustrated in Figure 2, in which all users have the same signal-to-noise ratio ($E_s/N_0$). The results of the conventional matched filter (MF) and the decorrelating detector (DD) [9] on the same user are also presented. The theoretical performance in the absence of multiple access interference is plotted as a lower bound (LB). It is illustrated that the BER of the OAS-based Bayesian MAP detector is very close to the lower bound. The conventional and decorrelating detectors, on the other hand, perform poorly in the high SNR region.

Although the simulation scenario described here is simple, the proposed blind Bayesian detection algorithm itself is generic, and may accommodate various system setups such as coded CDMA with any channel codes, e.g., block codes, convolutional codes, and Turbo codes.

V. SUMMARY

This paper presents a highly efficient Bayesian inference procedure for MAP detection using simulations. It allocates samples in a way that optimally improves an asymptotic approximation to the probability of correct selection. The resulting optimized adaptive sampling algorithm is applied to multiuser communication systems for blind optimum symbol detection. Such a receiver design is attractive to a wide range of wireless applications by virtue of its efficiency, optimality, and intuitive system modeling capacity.

REFERENCES


Fig. 1. Samples and histograms: $\{A_k\}_{k=1}^5 = \{-4, -2, 0, 2, 4\} dB$.

Fig. 2. Bit error rate performance of various multiuser detectors.