An Information Theoretic Transmitter Enumerator for DS-CDMA Wireless Networks

Shahrokh Valaee

Abstract— **In this paper, a new information theoretic algorithm is proposed for signal enumeration in DS-CDMA networks. The approach is based on the predictive description length (PDL), which is the length of a predictive code of observations. The PDL cost is computed for the candidate models and is minimized to determine the number of signals. The proposed technique uses the maximum likelihood (ML) estimate of the correlation matrix. The only information used in the ML estimation of the correlation matrix is the multiplicity of the smallest eigenvalue, therefore the method is applicable to blind multiuser detection. The PDL algorithm has a signal-to-noise ratio resolution threshold that is smaller than that of the minimum description length (MDL). The proposed method can be used online and can be applied to time-varying and non-stationary systems.**

Keywords—**DS-CDMA,WLANs, 3G cellular networks, multiuser detection, signal enumeration, information theoretic techniques.**

I. Introduction

Most wireless networks, such as 3G cellular systems and IEEE 802.11 WLANs, use the DS-CDMA signalling. A DS-CDMA signal is formed by multiplying each data bit by the signature waveform of the modulating sequence. In practice, the signature waveforms of different users observed at each receiver are not orthogonal. In such systems, the performance of the conventional receivers—in terms of bit-error-rate (BER)—is very poor. To combat the degradation of performance, *multiuser detectors* [1] are usually used. A Multiuser detector should know the true number of signals.

In various applications, the true number of signals is not known at the receiver. Here, we present two such examples. In wireless cellular networks, the true number of signals is not known at mobile terminals. Therefore, if a blind multiuser detector is used, the true number of signals should be estimated. In the IEEE 802.11b standard, several WLANs can coexist in a common environment. In such cases, since each WLAN is a distinct network, it is very difficult to determine the total number of transmitters and communicate it to all wireless terminals. Several other applications can be found where the true number of transmitters is not know at the receiver. Therefore, effective *passive* signal enumeration techniques should be developed to detect the number of signals by observing the waveform of the received signal. In passive signal enumeration, transmitters are unaware of the detection process and no coordination between the transmitters and the receiver is assumed.

The author is with the Edward S. Rogers Department of Electrical and Computer Engineering, University of Toronto, 10 King's College Road, Toronto, Ontario, Canada M5S 3G4. Email: valaee@comm.utoronto.ca.

In this paper, we introduce a novel technique to enumerate DS-CDMA signals that are used in both cellular networks and WLANs. We use an information theoretic approach to detect the number of DS-CDMA signals. Recently, much attention has been given to information theoretic criteria [2]–[7]. A popular information theoretic technique is the *minimum description length* (MDL) that is based on minimizing the length of the code required to describe data. Codelength minimization is appropriate for model selection since the model, which best fits the data, is the one that gives the most information about it; having more information results in a smaller codelength.

In this paper, we propose the *predictive description length* (PDL) algorithm. The PDL criterion is the cumulative log-likelihood function of the observation vectors such that at each time instant, the *maximum likelihood* (ML) estimate of the parameter based on the past data is used in the probability distribution function. PDL achieves the shortest codelength for data relative to the generating model class and has a structure that is suitable for on-line tracking of time-varying systems.

II. DS-CDMA Signals

In DS-CDMA, the waveform received at a wireless terminal can be represented by

$$x(t) = \sum_{m=1}^{M} \alpha_m b_m(k) h_m(t - kT - \tau_m) + n(t) \qquad (1)$$

where M is the number of signals, α_m is the received signal amplitude, $b_m(k)$ is the kth transmitted data symbol of the mth user, T is the symbol interval, $h_m(t)$, $0 \le t \le T$ is the mth received signal waveform, τ_m is a random delay, and $n(t)$ is the additive noise. The transmitted data symbol $b_m(k)$ belongs to the set of equiprobable $\{\pm 1\}$ random variables. In DS-CDMA, the received signal waveforms are of the form

$$h_m(t) = \sum_{j=0}^{N-1} h_j^m p(t - jT_c) \qquad (2)$$

where $h_0^m, \ldots, h_{L-1}^m$ are the signature sequences of the mth user selected from the binary alphabet $\{\pm 1\}$, $p(t)$, $0 \le t \le T_c$ is the normalized chip signal, T_c is the chip period, and N is the total number of chip signals used for the transmission of a single data bit.

The receiver usually uses a matched filter for each chip signal and a sampler at the chip frequency to detect the sequence of ± 1 chips extended over the symbol interval T. The detected sequence is then used to estimate the transmitted data bit. If we represent the output of the matched filter for the jth chip by $x_j(k)$, we will have

$$
\begin{aligned}
x_j(k) &= \int_{kT+jT_c}^{kT+(j+1)T_c} x(t)p(t-kT-jT_c)\,dt \\
&= \sum_{m=1}^{M} \alpha_m b_m(k)\mu_m h_j^m + n_j(k) \qquad (3)
\end{aligned}
$$

where μ_m is a random number representing the uncertainty due to unknown τ_m, and $n_j(k)$ is the sampled noise component at the output of the matched filter. In vector form, (3) is represented by

$$\mathbf{x}_k = \sum_{m=1}^{M} \alpha_m \mu_m b_m(k)\mathbf{h}_m + \mathbf{n}_k \qquad (4)$$

where $\mathbf{x}_k$ is the $N \times 1$ observation vector at the kth time instant, $\mathbf{h}_m = [h_0^m, \ldots, h_{N-1}^m]^T$ is the $N \times 1$ signature waveform of the mth user, and $\mathbf{n}_k = [n_0(k), \ldots, n_{N-1}(k)]^T$ is the $N \times 1$ additive noise vector at the kth time instant; the superscript T denotes transposition. Now define

$$s_m(k) \triangleq \alpha_m \gamma_m b_m(k), \qquad (5)$$

and get

$$\mathbf{x}_k = \mathbf{H}\mathbf{s}_k + \mathbf{n}_k \qquad (6)$$

where $\mathbf{s}_k = [s_1(k), \ldots, s_M(k)]^T$ is the $M \times 1$ signal vector, and $\mathbf{H} = [\mathbf{h}_1, \ldots, \mathbf{h}_M]$ is the $N \times M$ matrix of signature waveforms.

Throughout this paper, we assume that the columns of the signature waveform matrix $\mathbf{H}$ are linearly independent. We further assume that the signal snapshots form an i.i.d. sequence of Gaussian random vectors with an unknown covariance matrix $\mathbf{S}^M$. The noise samples are assumed to be independent from the signal samples and form an i.i.d. sequence of Gaussian random vectors with an unknown covariance matrix $\sigma^2\mathbf{I}_N$ where $\mathbf{I}_N$ is the $N \times N$ identity matrix, and σ^2 is the variance of noise. With these assumptions, the observation vector will be a sample of the Gaussian process with zero mean and the correlation matrix

$$\mathbf{R}^M = \mathbf{H}\mathbf{S}^M\mathbf{H}^T + \sigma^2\mathbf{I}_N. \qquad (7)$$

It is possible to show that the observation vector can be decomposed into two orthogonal vectors in the signal and noise subspaces.The *signal subspace* is the subspace spanned by the column vectors of $\mathbf{H}$. If the signal correlation matrix $\mathbf{S}^M$ is full-rank, the signal subspace will coincide with the span of the eigenvectors of $\mathbf{R}^M$ corresponding to M largest eigenvalues. Note that the dimension of the signal subspace is M. The *noise subspace* is the orthogonal complement of the signal subspace. The dimension of the noise subspace is $N - M$. The objective of this paper is to estimate the dimension of the signal subspace, M, given that $M \in \mathcal{N} \triangleq \{0, 1, \ldots, N-1\}$.

A direct implication of signal and noise subspace decomposition technique is that for high signal-to-noise ratio (SNR), the eigenvalues of the sample correlation matrix, corresponding to signal components, are significantly larger than the noise eigenvalues. Furthermore, the noise eigenvalues of the true correlation matrix are identical and are equal to σ^2. These observations can be used to devise a simple signal enumeration technique by comparing the difference between consecutive eigenvalues. In the sequel, we will call this enumerator by EIG. We will show that the performance of this enumerator is inferior to that of MDL and PDL. An inherent problem of the EIG enumerator is that it cannot detect the true number of signals when $M = 0$.

III. Predictive Description Length

For any $m \in \mathcal{N}$, we construct an appropriate model of order m. Assume that each model m is represented by a conditional probability density function $f(\mathbf{x}|\phi^m)$ where $\mathbf{x}$ is the observation vector and ϕ^m is the corresponding parameter vector. The PDL cost of the observation vectors $\mathbf{x}_k, k = 1, \ldots, K$, for a model of order m is defined as

$$\mathrm{PDL}_m(K) = -\sum_{k=1}^{K} \log f(\mathbf{x}_k|\hat{\phi}_{k-1}^m) \qquad (8)$$

where $\hat{\phi}_{k-1}^m$ is the ML estimate of the parameter vector using the observations up to time $(k-1)$. The PDL principle is based on the predictive encoding of data and has its roots in the theory of stochastic complexity [8]. At each time instant, the parameter vector is estimated using the past observations. Therefore, the kth term, $-\log f(\mathbf{x}_k|\hat{\phi}_{k-1}^m)$, is indeed the codelength of the prediction error. In this paper, we will project the PDL metric onto a trellis structure and will use it to modify the PDL cost so as to fit it to time varying systems. The PDL cost is calculated for each model and the smallest one is selected as the best model, that is

$$\hat{M} = \arg\min_m \mathrm{PDL}_m(k). \qquad (9)$$

To set the initial point in the recursion (8) we collect N snapshots and form the sample correlation matrix using these snapshots. This sample correlation matrix is then used to estimate the parameter vector. The PDL cost is accumulated for all $i = N+1, N+2, \ldots, K$ to find the total code length. Therefore, our formulation of the PDL criterion is

$$\mathrm{PDL}_m(K) = -\sum_{k=N+1}^{K} \log f(\mathbf{x}_k|\hat{\phi}_{k-1}^m). \qquad (10)$$

In the following section, we derive the PDL cost to detect the number of linearly independent columns of $\mathbf{H}$ in (6).

For the model of order m, let the channel output signal at the kth time instant be expressed by

$$\mathbf{x}_k = \mathbf{H}_m\mathbf{s}_k^m + \mathbf{n}_k^m \qquad (11)$$

where $\mathbf{H}_m$ is the $N \times m$ matrix of signature waveforms, and $\mathbf{s}_k^m(t)$ is an $m \times 1$ signal vector, and $\mathbf{n}_k^m$ is the $N \times 1$ noise vector of model m. Assuming that the signal vector $\mathbf{s}_k^m$

and the noise vector $\mathbf{n}_k^m$ are independent, the correlation matrix of the observation vector $\mathbf{x}_k$ is

$$\mathbf{R}^m = \mathbf{H}_m \mathbf{S}^m \mathbf{H}_m^T + \sigma_m^2 \mathbf{I}_N \quad (12)$$

where $\mathbf{S}^m$ is the signal autocorrelation matrix and $\sigma_m^2 \mathbf{I}_N$ is the noise autocorrelation matrix. Here, we assume that the noise is white with the unknown variance σ_m^2 that depends on the selected model.

The conditional probability density function of the observation vector for model m is given by

$$f(\mathbf{x}|\mathbf{R}^m) = \frac{1}{(2\pi)^{\frac{N}{2}}\ |\mathbf{R}^m|^{\frac{1}{2}}} \exp\Big(-\frac{1}{2}\mathbf{x}^T[\mathbf{R}^m]^{-1}\mathbf{x}\Big). \quad (13)$$

From (13), the PDL for a model of order m at time instant $K \geq N+1$ is given by

$$\mathrm{PDL}_m(K) = \sum_{k=N+1}^{K} \Big(\log|\hat{\mathbf{R}}_{k-1}^m| + \mathbf{x}_k^T[\hat{\mathbf{R}}_{k-1}^m]^{-1}\mathbf{x}_k\Big) \quad (14)$$

where $\hat{\mathbf{R}}_{k-1}^m$ is the ML estimate of the correlation matrix for the model of order m using the observations up to time $(k-1)$; in (14), the constant terms that are independent of the selected model have been removed and the whole cost has been multiplied by 2. The PDL cost is computed for each model and the minimum is chosen to estimate M. Indeed, at each time instant $k \geq N+1$, the best model is selected from

$$\hat{M}_k = \arg\min_{m \in \mathcal{N}} \mathrm{PDL}_m(k). \quad (15)$$

In (15), we have implicitly assumed that the number of signals can change over time.

In the sequel, the sample correlation matrix is used to obtain the ML estimate of the true correlation matrix. The sample correlation matrix is defined with the recursion

$$\bar{\mathbf{R}}_k = (1-\delta)\bar{\mathbf{R}}_{k-1} + \delta \mathbf{x}_i \mathbf{x}_i^T \quad (16)$$

where $0 < \delta < 1$ is usually very small. This definition of the sample correlation matrix is very useful for nonstationary environments. In (16), by varying δ, we can have different weights for $\bar{\mathbf{R}}_{k-1}$ and $\mathbf{x}_k\mathbf{x}_k^T$. For large values of δ, the sample correlation matrix has a short memory and is sensitive to recent changes of the underlying statistics, and for small values of δ, it has a longer memory and sudden changes in the statistics of $\mathbf{x}_k\mathbf{x}_k^T$ are smoothed out.

Let $\bar{\lambda}_{k,j}, j = 1, \ldots, N$, be the eigenvalues of $\bar{\mathbf{R}}_{k-1}$ arranged in nonincreasing order, and $\bar{\mathbf{v}}_{k,j}$ be the corresponding eigenvectors. It is possible to show that the eigenvalues and the eigenvectors of the ML estimator $\hat{\mathbf{R}}_{k-1}^m$ are given by [9]

$$\hat{\lambda}_{k,j} = \begin{cases} \bar{\lambda}_{k,j} & \text{if } 1 \leq j \leq m, \\ \dfrac{1}{N-m} \displaystyle\sum_{\ell=m+1}^{N} \bar{\lambda}_{k,\ell} & \text{if } m+1 \leq j \leq N, \end{cases} \quad (17)$$

$$\hat{\mathbf{v}}_{k,j} = \bar{\mathbf{v}}_{k,j} \qquad \text{for } j = 1, \ldots, N. \quad (18)$$

These eigenvalues and eigenvectors are used to obtain the ML estimate of the correlation matrix $\hat{\mathbf{R}}_{k-1}^m$ as

$$\hat{\mathbf{R}}_k^m = \hat{\mathbf{V}}_k \hat{\mathbf{\Lambda}}_k \hat{\mathbf{V}}_k^T \quad (19)$$

with

$$\hat{\mathbf{V}}_k \triangleq [\bar{\mathbf{v}}_{k,1} \ldots \bar{\mathbf{v}}_{k,N}] \quad (20)$$

$$\hat{\mathbf{\Lambda}}_k \triangleq \begin{bmatrix} \mathrm{diag}(\bar{\lambda}_{k,1}, \ldots, \bar{\lambda}_{k,m}) & \mathbf{0}_{m,N-m} \\ \mathbf{0}_{N-m,m} & \hat{\sigma}_{m,k}^2 \mathbf{I}_{N-m} \end{bmatrix} \quad (21)$$

where $\mathrm{diag}(\bar{\lambda}_{k,1}, \ldots, \bar{\lambda}_{k,m})$ is a diagonal matrix with the diagonal entities given in the brackets,

$$\hat{\sigma}_{m,k}^2 = \frac{1}{N-m} \sum_{\ell=m+1}^{N} \bar{\lambda}_{k,\ell} \quad (22)$$

is the ML estimate of the noise variance σ_m^2, and $\mathbf{0}_{i,j}$ is an $i \times j$ all-zero matrix.

Using the cumulative structure of the PDL algorithm, we have

$$\mathrm{PDL}_m(k) = \mathrm{PDL}_m(k-1) + \ell_{m,k} \quad (23)$$

where $\ell_{m,k}$ is the description length of model m at time instant k and is given by

$$\begin{aligned} \ell_{m,k} &= \sum_{j=1}^{m} \log \bar{\lambda}_{k-1,j} + (N-m)\log \hat{\sigma}_{m,k-1}^2 \\ &\quad + \mathbf{x}_k^T[\hat{\mathbf{R}}_{k-1}^m]^{-1}\mathbf{x}_k \\ &= \sum_{j=1}^{m} \log \bar{\lambda}_{k-1,j} + (N-m)\log \hat{\sigma}_{m,k-1}^2 \\ &\quad + \sum_{j=1}^{m} \frac{\gamma_{k,j}}{\bar{\lambda}_{k-1,j}} + \frac{1}{\hat{\sigma}_{m,k-1}^2} \sum_{j=m+1}^{N} \gamma_{k,j} \end{aligned} \quad (24)$$

where the constant terms (independent from the model order m) have been removed, and $\gamma_{k,j}$ is found from

$$\gamma_{k,j} = \bar{\mathbf{v}}_{k-1,j}^T \mathbf{x}_k \mathbf{x}_k^T \bar{\mathbf{v}}_{k-1,j} - \bar{\lambda}_{k-1,j}. \quad (25)$$

The estimate of M is then obtained from

$$\hat{M}_k = \arg\min_{m} \Big\{\mathrm{PDL}_m(k-1) + \ell_{m,k}\Big\}. \quad (26)$$

The PDL criterion needs eigenvalue decomposition and inversion of the correlation matrix that both are computationally expensive. In this paper, we use the *first order perturbation* technique [10] to devise a recursive structure for PDL. Using a method similar to [10], it is possible to show that the eigenvalue decomposition of the sample correlation matrix is given[1] by

$$\bar{\lambda}_{k,j} = (1-\epsilon)\bar{\lambda}_{k-1,j} + \epsilon \bar{\mathbf{v}}_{k-1,j}^T \mathbf{x}_k \mathbf{x}_k^T \bar{\mathbf{v}}_{k-1,j} \quad (27)$$

$$\bar{\mathbf{v}}_{k,j} = \bar{\mathbf{v}}_{k-1,j} + \epsilon \sum_{i=1}^{N} b_{ij} \bar{\mathbf{v}}_{k-1,j} \quad (28)$$

$$b_{ij} = \frac{\bar{\mathbf{v}}_{k-1,j}^T \mathbf{x}_k \mathbf{x}_k^T \bar{\mathbf{v}}_{k-1,j}}{\max\{0.01\bar{\lambda}_{1,j}, \bar{\lambda}_{k-1,j} - \bar{\lambda}_{k-1,i}\}}. \quad (29)$$

[1] Due to space limitation, the details of the derivations are not presented here.

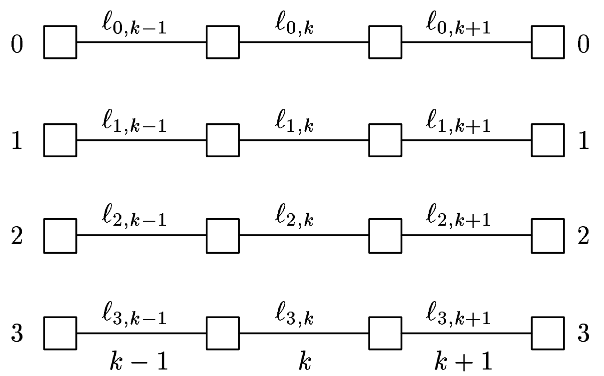

Fig. 1. The trellis representation of the PDL algorithm for an example with four models $\{0,1,2,3\}$. A node at the mth row of the trellis corresponds to a model with m signals.

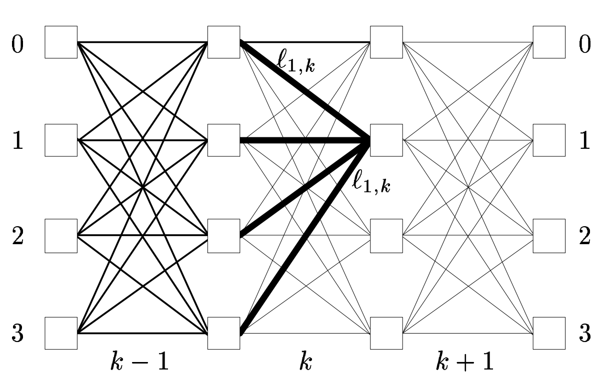

Fig. 2. The trellis representation of the PDL algorithm for an example with four models $\{0,1,2,3\}$. A node at the mth row of the trellis corresponds to a model with m signals. All lines arriving at a single node m (an example is illustrated in heavy lines) have the same weight, $\ell_{m,k}$.

The PDL algorithm proceeds by solving (27)-(29) and using the results in (24) and (23). It has been shown that the first order perturbation can significantly improve the computational cost of the eigenvalue decomposition of the sample correlation matrix [10].

The MDL cost can also be obtained with a similar approach. In this paper, we compare our results to the following formulation of the MDL algorithm that was originally proposed in [2] and has been widely used in order statistics (see for instance [11])

$$\mathrm{MDL}_m(K) = K \log\Big(\frac{(\hat{\sigma}^2_{m,K})^{N-m}}{\prod_{j=m+1}^{N} \bar{\lambda}_{K,j}}\Big) + \frac{m}{2}(2N-m)\log K. \tag{30}$$

The first term measures the multiplicity of the smallest eigenvalue and the second term is the penalty factor for over-modelling.

IV. Trellis Representation

The PDL algorithm can also be represented on a trellis structure as illustrated in Fig. 1. Each node on the trellis corresponds to a fixed model. Therefore, the total number of nodes at each stage is N—corresponding to the models $\{0, 1, \ldots, N-1\}$. The weight of each line arriving at a node m at time instant k, is given by the PDL cost of the corresponding model at that time instant, $\ell_{m,k}$. The total PDL cost for a given model at each time instant k is obtained by causal filtering of the sequence $\{\ell_{m,k}\}$. Our formulation of the PDL in (23) finds the average of all $\ell_{m,k}, k = N+1, \ldots, K$. With the trellis representation of the PDL algorithm, it is possible to define other filtering methods on the sequence $\{\ell_{m,k}\}$. For instance, the PDL cost may be defined as

$$\mathrm{PDL}_m(k) = \alpha \mathrm{PDL}_m(k-1) + (1-\alpha)\ell_{m,k}. \tag{31}$$

This formulation provides a filtering scheme that uses an exponential weighting factor α to emphasize the recent values of $\ell_{m,k}$ more than the values in the past. This approach is particulary useful in nonstationary environments.

The trellis representation of the PDL algorithm is very useful for the cases where the number of sources is time varying. In Fig. 2, we have extended the trellis to allow transition between different models. On this trellis, all lines arriving at a node have identical weights. Such as before, various filtering techniques can be used to define the PDL cost. The filtering methods may allow aggregation of cost along crossing lines. These filters are of two-dimensional nature and accumulate the PDL cost over model orders (straight lines in the trellis) as well as across models (cross lines in the trellis). We do not investigate the two-dimensional filtering methods in this paper and defer them to a future work.

The trellis representation of the PDL algorithm sheds light on the way this method handles time-varying number of signals. Note in Fig. 2 that a very small $\ell_{m,k}$ will tend to move the minimum PDL cost path in the direction of $\ell_{m,k}$ (node m). This suggests that comparing the PDL cost functions of different models can detect the changes in the number of signals. Indeed, we have used this fact in our simulation studies to locate changes in the number of signals by investigating $\mathrm{PDL}_m(k) - \mathrm{PDL}_{m+1}(k)$. We have shown that the difference between the PDL costs has a knee-point at the vicinity of the change point.

V. Simulation Results

In this section, we present the simulation results. Consider a DS-CDMA system using the 31-bit Gold codes. We assume that 4 signals are presently active in a Rayleigh fading channel. The received signal is the superposition of the signals of all users. The received signal is collected over a window of size 100. The signal-to-noise ratio (SNR) is defined as the ratio of the power of any one of the signals to the power of noise. Noise is assumed to be a white Gaussian process.

We performed 100 independent runs and found the eigenvalues of the sample correlation matrix in each run. Fig. 3 illustrates the eigenvalues of the sample correlation matrix averaged over 100 runs. Note that the eigenvalues can be decomposed into signal and noise eigenvalues. The large eigenvalues (in this example the first 4 eigenvalues) are the signal eigenvalues. In this example, the distance between the 4th and the 5th eigenvalues is larger than the distance

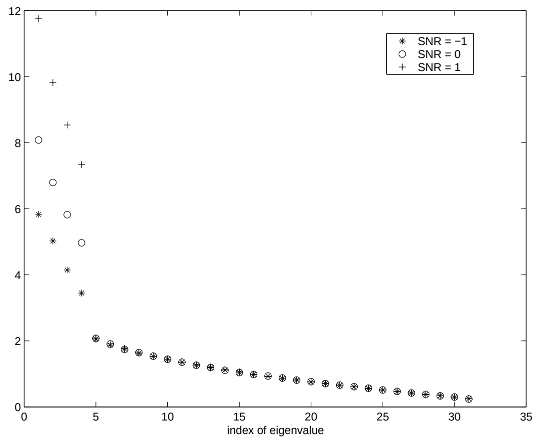

Fig. 3. The eigenvalues of the sample correlation matrix averaged for 100 independent runs. The results correspond to 3 different SNRs.

		SNR (dB)					
	m	-3	-2	-1	0	1	2
EIG	1	47	28	18	9	4	1
	2	25	18	15	3	0	0
	3	17	19	13	6	1	1
	4	10	35	54	82	95	98
	5	1	0	0	0	0	0
MDL	1	100	95	10	0	0	0
	2	0	4	13	0	0	0
	3	0	1	36	3	0	0
	4	0	0	41	97	100	100
	5	0	0	0	0	0	0
PDL	1	80	12	0	0	0	0
	2	13	21	0	0	0	0
	3	6	35	6	0	0	0
	4	1	32	94	100	100	100
	5	0	0	0	0	0	0

TABLE I

THE NUMBER OF DS-CDMA SIGNALS DETECTED BY THE EIG, MDL, AND PDL ALGORITHMS FOR 100 INDEPENDENT RUNS.

between the other consecutive eigenvalues. Therefore, a simple detector (denoted as EIG) selects the number of signals by comparing the difference between the consecutive eigenvalues. We will compare our proposed detector to EIG in the sequel.

The MDL and PDL techniques were used to estimate the number of signals. We have performed 100 independent runs and presented the results in Table I. In this table, three detectors have been compared. Each row in the table shows the number of times that the detector selected the corresponding model for the given SNR. Note that PDL outperforms the MDL and EIG methods. As the SNR decreases, the signal and noise eigenvalues approach each other and cannot be easily separated. In such cases, none of the methods can properly estimate the true number of DS-CDMA signals.

The recursive structure of the PDL algorithm can be very useful in non-stationary environments. In this section, we study two cases at which the number of sources changes inside the window of observation. We consider a window size of 300 samples with SNR $=$ 3dB. In the first example, we assume that the number of signals is 3 at the beginning of the window and changes to 4 at $k = 100$. We compute the PDL cost for 20 independent runs and find the average of these runs. The PDL cost is computed for all $1 \leq k \leq 300$. At each time instant, we estimate the number of signals by locating the minimum PDL cost. The number of detected signals has been shown as a function of time in Fig. 4. Note that the number of detected signals changes from 3 to 4 in the vicinity of the 100th time instant.

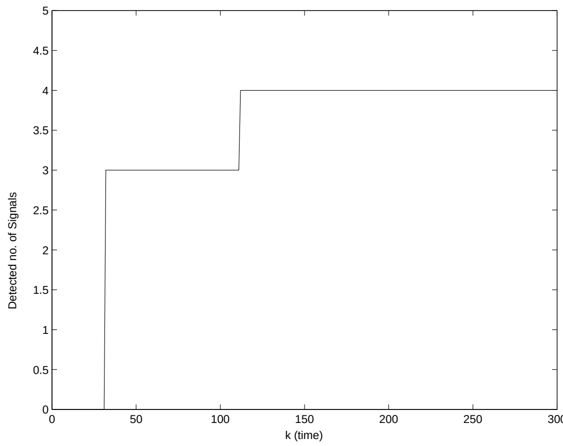

Fig. 4. The estimated number of sources as a function of the observation window. The true number of signals changes from 3 to 4 at $k = 100$.

We have also computed $\mathrm{PDL}_m(k) - \mathrm{PDL}_{m+1}(k)$, the difference between the PDL costs of consecutive models for all time instants inside the window of observation. The results have been illustrated in Fig. 5 for three different values of $m = 1, 2, 3$. Note an abrupt change in the slope of $\mathrm{PDL}_3(k) - \mathrm{PDL}_4(k)$ in the vicinity of $k = 100$. The change indicates that the underlying model which was used for $k = 1, \ldots, 100$ is not valid for the rest of the window. This figure can be used to locate the change point inside the window of observation. Since MDL operates on a batch of data, it cannot locate the change.

We also study the case at which the number of signals changes from 4 to 3 at the 100th time instant. Such as before, the PDL cost is averaged over 20 independent runs. The results have been shown in Fig. 6 and Fig. 7. Note that the number of signals has not been detected properly over the window $k = 101, \ldots, 300$. The reason for this malfunctioning is that the number of prominent eigenvalues of the sample correlation matrix is 4 even after the actual number of signals is reduced to 3. Fig. 7 shows the difference between the PDL costs of consecutive model orders. Notice the abrupt change in the slope of the curve $\mathrm{PDL}_3(k) - \mathrm{PDL}_4(k)$. The change indicates that the underlying model is not valid and the sample correlation matrix should be reset and recomputed using the recent data.

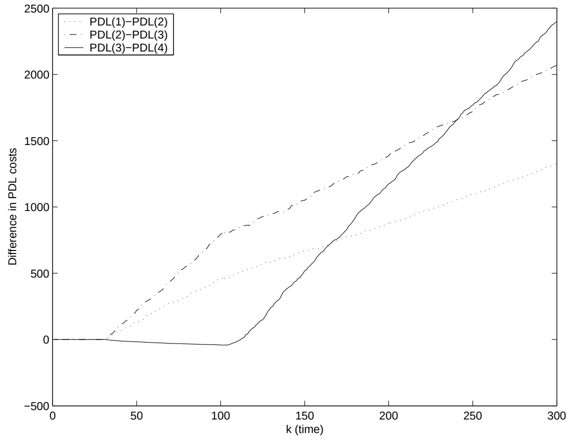

Fig. 5. The difference between the PDL terms of a model of order m and the corresponding terms of the model of order $m+1$. The true number of signals changes from 3 to 4 at $k = 100$.

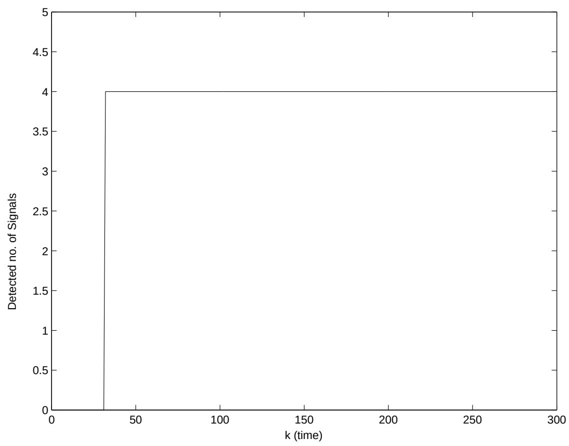

Fig. 6. The estimated number of sources as a function of the observation window. The true number of signals changes from 4 to 3 at $k = 100$.

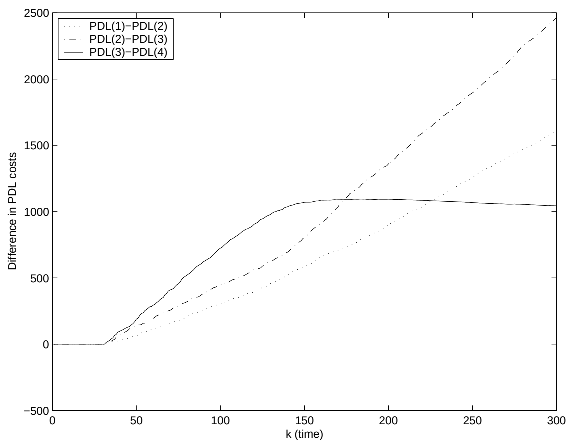

Fig. 7. The difference between the PDL terms of a model of order m and the corresponding terms of the model of order $m+1$. The true number of signals changes from 4 to 3 at $k = 100$.

VI. Summary

In this paper, we have introduced a new information theoretic method to estimate the number of signals in DS-CDMA networks. Our approach is based on the predictive description length (PDL). PDL is the length of a predictive code that encodes the observed data. We use the code-length as a metric that describes the observation vector. The best model is the one that gives the smallest code length. The PDL cost is computed for all candidate models and the one with the smallest cost is selected as the best-fit model.

The proposed method is based on the ML estimate of the correlation matrix. To apply our technique we do not need the signature waveform of DS-CDMA signals and only use the multiplicity of the smallest eigenvalue of the correlation matrix. Therefore, this technique can be used in blind multiuser detection at which the signature waveform of signals is not known. The simulation results show that the performance of the PDL algorithm is better than that of the MDL and EIG methods.

References

[1] S. Verdu, *Multiuser Detection.* Cambridge, U.K.: Cambridge Univ. Press, 1998.

[2] M. Wax and T. Kailath, “Detection of signals by information theoretic criteria,” *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 33, pp. 387–392, April 1985.

[3] T. E. Shrimpton and S. Schell, “Source enumeration using a signal-selective information theoretic criterion,” *Military Communications Conference (MILCOM)*, vol. 3, pp. 1092–1097, 1997.

[4] K. C. Tan, M. Zhang, and M. Wax, “Signal enumeration and direction-of-arrival estimation for coherent signals in correlated noise with unknown parameters,” *Int. Conf. on Signal Processing (ICSP)*, vol. 1, pp. 339–342, 1998.

[5] A. Barron, J. Rissanen, and B. Yu, “The minimum description length principle in coding and modeling,” *IEEE Trans. Information Theory*, vol. 44, pp. 2743 –2760, Oct. 1998.

[6] M. H. Hansen and B. Yu, “Model selection and the principle of minimum description length,” *Jour. American Stat. Assoc.*, vol. 96, pp. 746–774, June 2001.

[7] P. J. Green and D. P. Taylor, “Dynamic signal enumeration algorithm for smart antennas,” *IEEE Trans. Signal Processing*, vol. 50, pp. 1307–1314, June 2002.

[8] J. Rissanen, “Fisher information and stochastic complexity,” *IEEE Trans. Information Theory*, vol. 42, pp. 40–47, Jan. 1996.

[9] T. W. Anderson, “Asymptotic theory for principal component analysis,” *Ann. Math. Stat.*, vol. 34, pp. 122–148, 1963.

[10] B. Champagne, “Adaptive eigendecomposition of data covariance matrices based on first-order perturbations,” *IEEE Trans. Signal Processing*, vol. 42, pp. 2758–2770, Oct. 1994.

[11] E. Fishler and H. Messer, “On the use of order statistics for improved detection of signals by the MDL criterion,” *IEEE Trans. Signal Processing*, vol. 48, pp. 2242–2247, August 2000.