QUICK AGGREGATION OF MARKOV CHAIN FUNCTIONALS VIA STOCHASTIC COMPLEMENTATION

Kutluyıl Doğançay and Vikram Krishnamurthy

Department of Electrical and Electronic Engineering
The University of Melbourne, Parkville, Victoria 3052, Australia
k.dogancay@ee.mu.oz.au

ABSTRACT

The paper presents a quick and simplified aggregation method for a large class of Markov chain functionals based on the concept of stochastic complementation. Aggregation results in a reduction in the number of Markov states by grouping them into a smaller number of aggregated states, thereby producing a considerable saving on computational complexity associated with maximum likelihood parameter and state estimation for hidden Markov models. The importance of the proposed aggregation method stems from the ease with which Markov chains with a large number of states can be aggregated. Three Markov chain functionals which have widespread use are considered to illustrate the application of our aggregation method.

1. INTRODUCTION

The computational complexity for maximum likelihood (ML) estimation of hidden Markov models (HMMs) [1] is proportional to the square of the number of Markov states, which rules out the implementation of ML estimators in many practical applications where the number of Markov states is large. One way of reducing the computational cost is to aggregate the Markov chain by grouping the states into a smaller number of aggregated states. In this paper we consider the use of *stochastic complementation* [2] for aggregating a certain class of Markov chain functionals which are shown to have a special "invariance" property making their aggregation trivial.

Apart from certain special cases where the Markov chain is "exactly lumpable" [3], aggregation generally results in an approximation to the original Markov chain. Stochastic complementation meets the requirement of exact *steady state* aggregation and, in addition, has several important properties: (i) the resulting aggregated matrix is stochastic and is irreducible if the original transition matrix is irreducible, (ii) for a certain class of Markov chains known as *nearly completely decomposable Markov chains* (NCDMCs) [2] stochastic complementation is an accurate approximation for the state probabilities at any time k.

The Simon-Ando theory for nearly completely decomposable chains [4] provides the theoretical basis for most aggregation techniques. In [5] the Simon-Ando theory has been applied to numerical aspects of queueing networks. The aggregation method developed in [5] is not exact in that it produces only an approximation to the stationary probability distribution of the unaggregated chain.

To begin with, we present our main aggregation result and state the conditions for its validity. Section 3 presents a number of Markov chain functionals which can be aggregated by our quick aggregation method. Section 4 demonstrates an application of the quick aggregation method to parameter estimation.

This work is supported by the Australian Research Council (ARC) and the Cooperative Research Centre for Sensor, Signal and Information Processing (CSSIP)

2. MAIN RESULT

Let us assume that $\{\boldsymbol{x}(k)\}$, $k \in \mathbb{Z}^+$, is an N-state, homogeneous, irreducible Markov chain with state space $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_N\}$ where $\boldsymbol{e}_i$ is the $N \times 1$ unit column vector of $\mathbb{R}^N$ with one in its ith entry and zero elsewhere. Let $\boldsymbol{P}$ denote the $N \times N$ transition probability matrix with its (i,j)th entry given by $p_{ij} = \Pr\{\boldsymbol{x}(k+1) = \boldsymbol{e}_j \mid \boldsymbol{x}(k) = \boldsymbol{e}_i\}$ and $\sum_{j=1}^{N} p_{i,j} = 1$, $\forall i \in \{1, \ldots, N\}$ (i.e. $\boldsymbol{P}$ is a stochastic matrix). Assume that $\boldsymbol{P}$ has the following K-level partition

$$\boldsymbol{P} = \begin{bmatrix} \boldsymbol{P}_{11} & \boldsymbol{P}_{12} & \cdots & \boldsymbol{P}_{1K} \\ \boldsymbol{P}_{21} & \boldsymbol{P}_{22} & \cdots & \boldsymbol{P}_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{P}_{K1} & \boldsymbol{P}_{K2} & \cdots & \boldsymbol{P}_{KK} \end{bmatrix}_{N \times N} \tag{1}$$

where all diagonal blocks $\boldsymbol{P}_{ii}$ are square matrices of size $N_i \times N_i$ such that $\sum_{i=1}^{K} N_i = N$. The following theorem formally states our main result.

Theorem 1. *Let $\boldsymbol{P}$ be an $N \times N$ irreducible stochastic matrix with a K-level partition as in (1). If the matrix partitions $\boldsymbol{P}_{ij}$ satisfy the following condition*

$$\boldsymbol{P}_{ij}\mathbf{1}_{N_j} = n_{ij}\mathbf{1}_{N_i} \quad \forall i,j \in \{1,\dots,K\} \tag{2}$$

where $0 < n_{ij} < 1$ and $\mathbf{1}_{N_i}$ is the $N_i \times 1$ column vector of ones, then the K-level aggregation of $\boldsymbol{P}$ is given by

$$\boldsymbol{C} = \begin{bmatrix} n_{11} & n_{12} & \cdots & n_{1K} \\ n_{21} & n_{22} & \cdots & n_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ n_{K1} & n_{K2} & \cdots & n_{KK} \end{bmatrix}$$

which is an exact aggregation of $\boldsymbol{P}$ in that the stationary (steady state) distribution of $\boldsymbol{P}$ can be recovered exactly from $\boldsymbol{C}$.

Proof. Stochastic complements of $\boldsymbol{P}_{ii}$ are given by [2] $\boldsymbol{S}_{ii} = \boldsymbol{P}_{ii} + \boldsymbol{P}_{i*}(\boldsymbol{I} - \boldsymbol{P}_i)^{-1}\boldsymbol{P}_{*i}$ where $\boldsymbol{P}_i$ is the principal block submatrix of $\boldsymbol{P}$ obtained by deleting the ith row and ith column of blocks from $\boldsymbol{P}$, $\boldsymbol{P}_{i*}$ is the ith row of blocks with $\boldsymbol{P}_{ii}$ removed, and $\boldsymbol{P}_{*i}$ is the ith column of blocks with $\boldsymbol{P}_{ii}$ removed.

The (i,j)th entry of the $K \times K$ aggregation matrix (or the coupling matrix) $\boldsymbol{C}$ is given by [2]

$$c_{ij} = \boldsymbol{s}_i^T \boldsymbol{P}_{ij}\mathbf{1}_{N_j}, \quad i,j \in \{1,\dots,K\} \tag{3}$$

where $\boldsymbol{s}_i = [s_{i1},\dots,s_{iN_i}]^T$ is the $N_i \times 1$ unique stationary distribution vector for the $N_i \times N_i$ stochastic complement $\boldsymbol{S}_{ii}$. The vector $\boldsymbol{s}_i$, which is the *Perron-Frobenius* eigenvector of $\boldsymbol{S}_{ii}$, is defined by $\boldsymbol{s}_i^T\boldsymbol{S}_{ii} = \boldsymbol{s}_i^T$, $s_{ij} > 0$ and $\sum_{j=1}^{N_i} s_{ij} = 1$. If (2) holds, (3) reduces to $c_{ij} = n_{ij}\boldsymbol{s}_i^T\mathbf{1}_{N_i} = n_{ij}$ since $\boldsymbol{s}_i^T\mathbf{1}_{N_i} = 1$. □

An important consequence of Theorem 1 is that if (2) holds, the aggregation matrix $\boldsymbol{C}$ can be obtained directly from $\boldsymbol{P}$ without having to compute stochastic complements $\boldsymbol{S}_{ii}$ and their stationary distribution vectors.

3. EXAMPLES OF MARKOV CHAIN FUNCTIONALS

We now apply Theorem 1 to three examples of Markov chain functionals for which (2) is readily satisfied. The examples are in no way artificial as they cover a wide range of applications in signal processing and telecommunications.

3.1. FIR Filtered Markov Chains

In many telecommunications problems, filtered Markov chains naturally arise as a result of channel coding and linear channel characteristics (see e.g. [6]).

Suppose that a K-state Markov chain $\boldsymbol{u}(k)$ with state space $\{\boldsymbol{e}_1,\dots,\boldsymbol{e}_K\}$ ($\boldsymbol{e}_i$ is the $K \times 1$ unit column vector), transition probability matrix

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1K} \\ a_{21} & a_{22} & \cdots & a_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ a_{K1} & a_{K2} & \cdots & a_{KK} \end{bmatrix}$$

and levels $\boldsymbol{g} = [g_1,\dots,g_K]^T$ is passed through an FIR filter of length L $H(z) = \sum_{i=0}^{L-1} h_i z^{-i}$. Denote the impulse response of $H(z)$ by $\boldsymbol{h} = [h_0,\dots,h_{L-1}]$. The process $\boldsymbol{X}(k) = [\boldsymbol{u}(k), \boldsymbol{u}(k-1),\dots,\boldsymbol{u}(k-L+1)]^T$ associated with the output of the FIR filter is also a Markov chain with $N = K^L$ states and state space $\{\boldsymbol{E}_1,\dots,\boldsymbol{E}_N\}$ where $\boldsymbol{E}_i = [\boldsymbol{e}_{i_1}, \boldsymbol{e}_{i_2},\dots,\boldsymbol{e}_{i_L}]_{K\times L}$ with $i \in \{1,\dots,N\}$ and $i_1,\dots,i_L \in \{1,\dots,K\}$. The transition probabilities of $\boldsymbol{X}(k)$ are

$$\begin{aligned} p_{ij} &= \Pr\{\boldsymbol{X}(k+1) = \boldsymbol{E}_j \mid \boldsymbol{X}(k) = \boldsymbol{E}_i\} \\ &= \Pr\{\boldsymbol{u}(k+1) = \boldsymbol{e}_{j_1},\dots,\boldsymbol{u}(k-L+2) = \boldsymbol{e}_{j_L} \mid \\ &\quad \boldsymbol{u}(k) = \boldsymbol{e}_{i_1},\dots,\boldsymbol{u}(k-L+1) = \boldsymbol{e}_{i_L}\} \\ &= \Pr\{\boldsymbol{u}(k+1) = \boldsymbol{e}_{j_1} \mid \boldsymbol{u}(k) = \boldsymbol{e}_{i_1}\}\delta(j_2 - i_1) \\ &\quad \cdots\delta(j_L - i_{L-1}) \\ &= a_{i_1 j_1}\delta(j_2 - i_1)\cdots\delta(j_L - i_{L-1}) \end{aligned} \tag{4}$$

for $i,j \in \{1,\dots,N\}$, where $\delta(\cdot)$ is the *Kronecker delta function.*

The levels of the filtered Markov chain are given by the functional

$$y(k) = \boldsymbol{h}^T\boldsymbol{X}^T(k)\boldsymbol{g} = \sum_{i=0}^{L-1} h_i\langle \boldsymbol{g}, \boldsymbol{u}(k-i)\rangle \tag{5}$$

where $\langle\cdot,\cdot\rangle$ denotes inner product.

The transition probability matrix $\boldsymbol{P}$ of $\boldsymbol{X}(k)$, whose entries are given by (4), can be written as follows after an appropriate permutation of the states:

$$\widetilde{\boldsymbol{P}} = \begin{bmatrix} \widetilde{\boldsymbol{P}}_{11} & \widetilde{\boldsymbol{P}}_{12} & \cdots & \widetilde{\boldsymbol{P}}_{1K} \\ \widetilde{\boldsymbol{P}}_{21} & \widetilde{\boldsymbol{P}}_{22} & \cdots & \widetilde{\boldsymbol{P}}_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ \widetilde{\boldsymbol{P}}_{K1} & \widetilde{\boldsymbol{P}}_{K2} & \cdots & \widetilde{\boldsymbol{P}}_{KK} \end{bmatrix}_{N\times N}$$

where the K-level partitions are $\widetilde{\boldsymbol{P}}_{ij} = [\mathbf{0}_{p\times q}, \boldsymbol{A}_{ij}, \mathbf{0}_{p\times r}]$ with $\mathbf{0}_{m\times n}$ denoting the $m \times n$ matrix of zeros, $p =$

K^{L-1}, $q = i-1$, $r = K^{L-1} - K^{L-2} - i + 1$, and

$$\boldsymbol{A}_{ij} = \begin{bmatrix} a_{ij}\mathbf{1}_K & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & a_{ij}\mathbf{1}_K \end{bmatrix}_{K^{L-1}\times K^{L-2}}, \quad L \geq 2.$$

It is easy to see that the blocks $\widetilde{\boldsymbol{P}}_{ij}$ satisfy (2) with $c_{ij} = a_{ij}$. Thus the K-level aggregation of $y(k)$ in (5) is given by the transition $\boldsymbol{C} = \boldsymbol{A}$, whether or not $\boldsymbol{A}$ itself can be partitioned in accordance with (2).

3.2. Binary Operations on Independent Markov Chains

In various signal processing problems such as pulse train de-interleaving and biological signal analysis [7], the signal of interest can often be modelled as the sum or product of independent Markov chains.

Let us assume that a given Markov chain is made up of L statistically independent K_i-state, homogeneous, irreducible, constituent Markov chains $\boldsymbol{u}^{(i)}(k)$, $i = 1, 2, \ldots, L$. The chain $\boldsymbol{u}^{(i)}(k)$ has state space $\{\boldsymbol{e}_1^{(i)}, \ldots, \boldsymbol{e}_{K_i}^{(i)}\}$, transition probability matrix $\boldsymbol{A}^{(i)}$ and levels $\boldsymbol{g}^{(i)} = [g_1^{(i)}, \ldots, g_{K_i}^{(i)}]^T$. Obviously, the process $\boldsymbol{x}(k) = [\boldsymbol{u}^{(1)}(k)^T, \ldots, \boldsymbol{u}^{(L)}(k)^T]$ associated with a binary operation on the chains is also a Markov chain with $N = \prod_{i=1}^{L} K_i$ states and the $N \times N$ transition probability matrix

$$\boldsymbol{P} = \boldsymbol{A}^{(1)} \otimes \boldsymbol{A}^{(2)} \otimes \cdots \otimes \boldsymbol{A}^{(L)} \tag{6}$$

where $\otimes$ denotes the *Kronecker product*.

The sum of chains $\boldsymbol{u}^{(i)}(k)$ is given by the functional $y_s(k) = \sum_{i=1}^{L} \langle \boldsymbol{g}^{(i)}, \boldsymbol{u}^{(i)}(k)\rangle$ and the product by $y_p(k) = \prod_{i=1}^{L} \left(\boldsymbol{g}^{(i)}\right)^T \boldsymbol{u}^{(i)}(k)$. We note that the transition probability matrix $\boldsymbol{P}$ in (6) associated with $y_s(k)$ and $y_p(k)$ can be rewritten as

$$\begin{aligned} \boldsymbol{P} &= \boldsymbol{A}^{(1)} \otimes \boldsymbol{B} \\ &= \begin{bmatrix} a_{11}^{(1)}\boldsymbol{B} & a_{12}^{(1)}\boldsymbol{B} & \cdots & a_{1K_1}^{(1)}\boldsymbol{B} \\ a_{21}^{(1)}\boldsymbol{B} & a_{22}^{(1)}\boldsymbol{B} & \cdots & a_{2K_1}^{(1)}\boldsymbol{B} \\ \vdots & \vdots & \ddots & \vdots \\ a_{K_11}^{(1)}\boldsymbol{B} & a_{K_12}^{(1)}\boldsymbol{B} & \cdots & a_{K_1K_1}^{(1)}\boldsymbol{B} \end{bmatrix} \end{aligned} \tag{7}$$

where $\boldsymbol{B} = \boldsymbol{A}^{(2)} \otimes \cdots \otimes \boldsymbol{A}^{(L)}$ is a stochastic matrix. The transition probability matrix in (7) can now be aggregated into K_1 levels with block partitions given by $\boldsymbol{P}_{ij} = a_{ij}^{(1)}\boldsymbol{B}$. In this case the condition (2) is satisfied with $n_{ij} = a_{ij}^{(1)}$ or $\boldsymbol{C} = \boldsymbol{A}^{(1)}$. The states of $\boldsymbol{x}(k)$ can be re-labelled so as to change the order of matrices $\boldsymbol{A}^{(i)}$ in the Kronecker product in (6).

3.3. Markov Modulated Markov Chains

Markov modulated Markov chains (MMMCs) arise in the context of binary time series. Specifically, MMMCs can be used to model 1-bit quantised Markov modulated autoregressive (AR) time series with applications in pulse train de-interleaving [8].

An MMMC is defined as a Markov chain whose transition probabilities are determined by the states of another independent Markov chain. Assume that $\boldsymbol{u}^{(1)}(k)$ is a K_1-state Markov chain with transition probability matrix $\boldsymbol{A}^{(1)}$ and $\boldsymbol{u}^{(2)}(k)$ is a K_2-state conditionally Markov chain (conditioned on $\boldsymbol{u}^{(1)}(k)$). In this case the MMMC is defined by $\boldsymbol{x}(k) = \begin{bmatrix}\boldsymbol{u}^{(1)}(k) \\ \boldsymbol{u}^{(2)}(k)\end{bmatrix}$. Using the usual state space definition, the transition probabilities of $\boldsymbol{x}(k)$ are given by

$$\begin{aligned} &\Pr\left\{\boldsymbol{x}(k+1) = \begin{bmatrix}\boldsymbol{e}_j^{(1)} \\ \boldsymbol{e}_n^{(2)}\end{bmatrix} \,\middle|\, \boldsymbol{x}(k) = \begin{bmatrix}\boldsymbol{e}_i^{(1)} \\ \boldsymbol{e}_m^{(2)}\end{bmatrix}\right\} \\ &\quad = \Pr\{\boldsymbol{u}^{(1)}(k+1) = \boldsymbol{e}_j \mid \boldsymbol{u}^{(1)}(k) = \boldsymbol{e}_i\} \\ &\qquad \times \Pr\{\boldsymbol{u}^{(2)}(k+1) = \boldsymbol{e}_n \mid \boldsymbol{u}^{(2)}(k) = \boldsymbol{e}_m, \\ &\qquad\quad \boldsymbol{u}^{(1)}(k+1) = \boldsymbol{e}_j\} \\ &\quad = a_{ij}^{(1)} v_{jmn} \end{aligned}$$

whence, upon appropriate labelling of the states, the transition probability matrix $\boldsymbol{P}$ of $\boldsymbol{x}(k)$ can be written as

$$\boldsymbol{P} = \begin{bmatrix} a_{11}^{(1)}\boldsymbol{V}_1 & a_{12}^{(1)}\boldsymbol{V}_2 & \cdots & a_{1K_1}^{(1)}\boldsymbol{V}_{K_1} \\ a_{21}^{(1)}\boldsymbol{V}_1 & a_{22}^{(1)}\boldsymbol{V}_2 & \cdots & a_{2K_1}^{(1)}\boldsymbol{V}_{K_1} \\ \vdots & \vdots & \ddots & \vdots \\ a_{K_11}^{(1)}\boldsymbol{V}_1 & a_{K_12}^{(1)}\boldsymbol{V}_2 & \cdots & a_{K_1K_1}^{(1)}\boldsymbol{V}_{K_1} \end{bmatrix}$$

where

$$\boldsymbol{V}_i = \begin{bmatrix} v_{i11} & v_{i12} & \cdots & v_{i1K_2} \\ v_{i21} & v_{i22} & \cdots & v_{i2K_2} \\ \vdots & \vdots & \ddots & \vdots \\ v_{iK_21} & v_{iK_22} & \cdots & v_{iK_2K_2} \end{bmatrix}, \quad i = 1, \ldots, K_1.$$

It is easy to show that $\boldsymbol{V}_i$ is a stochastic matrix. Then the K_1-level partitions of $\boldsymbol{P}$ given by $\boldsymbol{P}_{ij} = a_{ij}^{(1)}\boldsymbol{V}_j$ can be immediately seen to satisfy (2), resulting in the aggregated transition probability matrix $\boldsymbol{C} = \boldsymbol{A}^{(1)}$.

4. APPLICATION OF AGGREGATION TO NCDMCS

Suppose that a four-level NCDMC $\boldsymbol{u}(k)$, $k \in \mathbb{Z}^+$, with state space $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_4\}$ and levels $\boldsymbol{g} = [1, 2, 3, 4]^T$ is observed in zero-mean white Gaussian noise with variance $\sigma_n^2 = 0.5$. The transition probability matrix of the chain is

$$\boldsymbol{P} = \begin{bmatrix} 0.900 & 0.050 & 0.020 & 0.030 \\ 0.650 & 0.300 & 0.010 & 0.040 \\ 0.015 & 0.015 & 0.800 & 0.170 \\ 0.020 & 0.010 & 0.400 & 0.570 \end{bmatrix}.$$

The observations are given by $s(k) = \langle \boldsymbol{g}, \boldsymbol{u}(k) \rangle + n(k)$. The probability density function of the observations given the levels is

$$b_i\big(s(k)\big) = \frac{1}{\sqrt{2\pi\sigma_n^2}} \exp\left(-\frac{\big(s(k) - g_i\big)^2}{2\sigma_n^2}\right), \quad 1 \leq i \leq 4.$$

We wish to estimate the levels of the Markov chain and the noise variance σ_n^2. To do so, we firstly aggregate the filtered chain using (2)

$$\boldsymbol{C} = \begin{bmatrix} 0.950 & 0.050 \\ 0.030 & 0.970 \end{bmatrix}.$$

The probability density function of aggregated levels in noise is given by [9]

$$\widetilde{b}_i\big(s(k)\big) = \sum_{j \in \mathcal{S}_i} \pi_j b_j\big(s(k)\big)/\xi_i, \quad i = 1, 2$$

where $\mathcal{S}_1 = \{1, 2\}$, $\mathcal{S}_2 = \{3, 4\}$, and $\boldsymbol{\pi} = [\pi_1, \ldots, \pi_8]^T$ and $\boldsymbol{\xi} = [\xi_1, \xi_2]^T$ are the stationary distribution vectors for the transition probability matrix $\boldsymbol{P}$ and $\boldsymbol{C}$, respectively.

Given the observation sequence $\{s(1), \ldots, s(T)\}$, the likelihood for the aggregated chain can be computed using the forward part of the *forward-backward algorithm* [1]:

$$\alpha_1(j) = \xi_j \widetilde{b}_j\big(s(1)\big)$$
$$\alpha_{k+1}(j) = \left(\sum_{i=1}^{2} \alpha_k(i) c_{ij}\right) \widetilde{b}_j\big(s(k+1)\big)$$

for $j = 1, 2$ and $k = 1, \ldots, T-1$, where we used $\boldsymbol{\xi}$ as the initial probability distribution for $\boldsymbol{C}$.

The likelihood function is given by $L(\boldsymbol{\theta}) = \alpha_T(1) + \alpha_T(2)$ where $\boldsymbol{\theta} = [\boldsymbol{g}, \sigma_n^2]^T$. An ML estimate of $\boldsymbol{\theta}$ for the reduced chain is obtained from $\widehat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta}} L(\boldsymbol{\theta})$. In a simulation experiment, we have generated a 1000-point sequence $\{s(1), \ldots, s(1000)\}$ and computed $L(\boldsymbol{\theta})$. An off-line simplex maximisation algorithm yielded the estimates $\widehat{\boldsymbol{\theta}} = [1.0263, 2.2395, 3.0685, 3.8470, 0.5494]^T$.

5. CONCLUSION

We have presented a class of Markov chain functionals for which stochastic complementation based aggregation can be achieved in an effortless manner. The multitude and generality of such Markov chain functionals points to the potential benefit of the simplified aggregation method. As an application, the parameter estimation of an NCDMC in white Gaussian noise was considered.

6. REFERENCES

[1] L. R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proc. IEEE*, vol. 77, pp. 257–286, Feb. 1989.

[2] C. D. Meyer, "Stochastic complementation, uncoupling Markov chains, and the theory of nearly reducible systems," *SIAM Rev.*, vol. 31, pp. 240–272, June 1989.

[3] J. G. Kemeny and J. L. Snell, *Finite Markov chains.* Princeton, NJ: Van Nostrand, 1960.

[4] H. A. Simon and A. Ando, "Aggregation of variables in dynamic systems," *Econometrica*, vol. 29, pp. 111–138, 1961.

[5] P. J. Courtois, *Decomposability: Queueing and Computer System Applications.* New York: Academic Press, 1977.

[6] V. Krishnamurthy and L. B. White, "Blind equalization of FIR channels with Markov inputs," in *Proc. IFAC Int. Symposium on Adaptive Systems in Control and Signal Processing*, (Grenoble, France), pp. 633–638, July 1992.

[7] S.-H. Chung and R. A. Kennedy, "Coupled Markov chain model: characterization of membrane channel currents with multiple conductance sublevels as partially coupled elementary pores," *Math. Biosciences*, vol. 133, pp. 111–137, 1996.

[8] A. Logothetis and V. Krishnamurthy, "De-interleaving of quantized AR processes with amplitude information," submitted to *IEEE Trans. Signal Processing*, 1995.

[9] V. Krishnamurthy, "Adaptive estimation of hidden nearly completely decomposable Markov chains with applications in blind equalization," *Int. J. Adaptive Control and Signal Processing*, vol. 8, pp. 237–260, 1994.