ASYMPTOTIC DETECTION PERFORMANCE OF TYPE-BASED MULTIPLE ACCESS IN SENSOR NETWORKS

Gökhan Mergen

Qualcomm Inc.,
675 Campbell Technology Parkway,
Campbell, CA 95008.

Vidyut Naware, Lang Tong[*]

School of Electrical and Computer Engineering,
Cornell University,
Ithaca, NY 14850.

ABSTRACT

The problem of communicating sensor readings over a *multiaccess* channel for detecting a target is considered. A natural way of communication in target detection is to let sensors *simultaneously* transmit one of two predetermined frequency tones indicating whether the target is detected or not. Recently, this scheme has been generalized to consider non-binary sensor observations by letting sensors simultaneously transmit *orthogonal waveforms* depending on the value of their observations—Type-Based Multiple Access (TBMA). TBMA was shown to be *asymptotically optimal* in terms of detection-error probability under the idealistic assumptions that the sensor channel gains are identical, and the sensor data are conditionally independent and identically distributed (i.i.d.). In this paper, TBMA is analyzed in a more general framework by considering non-i.i.d. data and non-identical channel gains. An asymptotically optimal detector is proposed and its error-exponents for detection probabilities are characterized using tools from *large deviations theory*. Numerical simulations are used to demonstrate that the error exponents provide reasonably accurate estimates of the performance of TBMA.

1. INTRODUCTION

We consider the problem of media access communication between sensor nodes and a *fusion center*. We consider a group of n sensors transmiting their data to the fusion center over a multiaccess channel (MAC). An important step in formulating this problem is modelling the sensor data. In general, sensors observe real-valued data. For practical purposes, however, the observations are generally *quantized* before communication. In this paper, we do not deal with how the quantization is done, and assume that the data $X_i \in \{1, \cdots, k\}$ of sensor i is *already* quantized to k possible levels. In target detection, sensors may quantize their data to two levels indicating whether target is detected or not. In parameter estimation, X_i's may model quantized measurements.

We use notation $\theta \in \mathbb{R}$ to denote the parameter to be detected. The parameter can be discrete in case of target detection (*i.e.,* , $\theta \in \{0, 1\}$ indicating the existence of target). Sensor data are statistically correlated, since nearby sensors tend to have correlated observations. A simple model for sensor data incorporates the *conditionally i.i.d.* assumption:

$$X_1, \cdots, X_n \overset{\text{i.i.d.}}{\sim} p_\theta \quad \text{given } \theta, \tag{1}$$

i.e., given the parameter θ, the sensor data $X_1, \cdots, X_n$ are conditionally i.i.d. according to a probability mass function (pmf) $p_\theta = (p_\theta(1), \cdots, p_\theta(k))$.

An interpretation of the conditionally i.i.d. assumption is that each sensor observes the same parameter θ, but with *i.i.d. observation noise*. The conditionally i.i.d. assumption is applicable in some scenarios, while in some others it is not. For example, it does not hold if the node observations have varying degrees of reliability. Furthermore, conditionally i.i.d. assumes that the sensed area is *uniform*, *i.e.,* the parameter *does not vary* in the observed area. In case the sensor observations come from a wide area with heterogenous parameter values, the model (1) needs to be generalized.

1.1. Type-Based Multiple Access

We deal with the transmission of sensor data $X_1, \cdots, X_n$ over a multiaccess channel. It is assumed sensor i has channel gain $h_i \in \mathbb{R}$,[1] which does not vary during the course of transmission. In this paper, we shall be primarily interested in the following scheme, which will be called Type-Based Multiple Access (TBMA).

[*]This work was supported in part by the Multidisciplinary University Research Initiative (MURI) under the Office of Naval Research Contract N00014-00-1-0564, by the Army Research Laboratory CTA on Communication and Networks under Grant DAAD19-01-2-0011, and by the National Science Foundation under Contract CCR-0330059.

[1]The results of this paper can be generalized to complex-valued channel gains with minor changes.

Let $s_1, \cdots, s_k$ be a set of k predetermined orthonormal waveforms. In the TBMA scheme, sensor i transmits the waveform s_{X_i} corresponding to its observation X_i with a certain energy E, *i.e.,* it transmits $\sqrt{E}s_{X_i}$. Due to the additive nature of wireless medium, the fusion center receives

$$z = \sum_{i=1}^{n} h_i \sqrt{E} s_{X_i} + w, \tag{2}$$

where w is white Gaussian channel noise with $\sigma^2/2$ power spectral density.

The motivation for TBMA arises from the special case that the sensor data are conditionally i.i.d. and sensor channel gains are identical (say, $h_i = 1, \forall i$). In this case, the received signal becomes

$$z = \sum_{j=1}^{k} \sqrt{E} N_j s_j + w, \tag{3}$$

where $N_j = \sum_{i=1}^{n} 1(X_i = j)$ is the number of sensors that observe symbol j. After matched filtering by $s_1, \cdots, s_k$, it can be seen that z contains a noisy version of the *histogram* of sensor observations. The basic idea in TBMA is to detect the target (or estimate the parameter) from this noisy histogram.

1.2. Related Work and Our Contribution

Estimation/detection over multiaccess channels has attracted considerable attention recently. TBMA has been proposed by the authors [1] and by Liu and Sayeed [2], independently. Works prior to TBMA (*e.g.,*[3, 4]) assumed that each sensor is allocated an orthogonal channel to transmit its observation as in TDMA, FDMA or CDMA.

Several asymptotic optimality properties of TBMA have been proved under the assumption of conditionally i.i.d. data and identical channel gains [1, 2, 5]. Similarly, it has been shown in [2, 5] that TBMA achieves the best *error exponent* in target detection. The intuition behind these optimality results is that the effect of noise w on z (eqn. (3)) becomes negligible as $n \to \infty$. As a result, the asymptotic performance of TBMA is as if the fusion center has direct access to histogram, which is sufficient to get optimal performance. Thus, the main conclusion that can be drawn from [1, 2, 5] is that *the asymptotic performance of TBMA is as if the fusion center has direct access to $X_1, \cdots, X_n$ in case of conditionally i.i.d. data and identical channel gains*.

In other orthogonal approaches such as TDMA, the bandwidth requirement grows linearly with n. In TBMA, however, the bandwidth requirement is independent of n—only k orthogonal dimensions are needed. This implies that TBMA is significantly more bandwidth efficient than other orthogonal allocation methods, when the number of sensors, n, is large compared to k. This is likely to be the case in a target detection scenario with binary sensor observations ($k = 2$).

In this paper, we first propose a detector that is asymptotically optimal in terms of providing the best error exponent in Bayesian hypothesis testing. Next, we provide an error exponent analysis of the TBMA scheme with *i.i.d. random channel gains* and conditionally i.i.d. data. In particular, we show that for the case that the channel gains $h_1, \cdots, h_n$ have *non-zero mean*, the detection error probabilities decay exponentially with n. Numerical simulations are provided to compare the performance of TBMA with other orthogonal allocation methods. We also provide a general characterization of the detection error exponents for non-i.i.d. channel gains and data using large deviation theory.

Organization of the paper is as follows. In Section 2, we review some results of large deviation theory and propose the *minimum-rate* detector. In Section 3, error exponents for i.i.d. channel gains and conditionally i.i.d. data are provided and analyzed. Next, the error exponents for the general (non-i.i.d.) case are provided. In Section 4, some examples and simulation results are presented. Section 5 concludes the paper.

2. LARGE DEVIATIONS AND THE MINIMUM-RATE DETECTOR

In this section, we are interested in the hypothesis testing problem

$$\mathcal{H}_0 : \theta = \theta_0 \text{ vs. } \mathcal{H}_1 : \theta = \theta_1. \tag{4}$$

Consider the TBMA scheme with received signal z. Upon reception of z, the fusion center decides on whether $\mathcal{H}_0$ or $\mathcal{H}_1$ is true. For a given decision rule at the fusion center, let $\alpha = \Pr\{\mathcal{H}_0 \to \mathcal{H}_1\}$ denote the probability that $\mathcal{H}_1$ is decided although $\mathcal{H}_0$ was true. Notation $\beta = \Pr\{\mathcal{H}_1 \to \mathcal{H}_0\}$ is defined analogously. The α and β are generally called *Type-I* and *Type-II* error probabilities in literature. We will use the notations α_n, β_n, when the dependence on n needs to be made explicit. In this paper, we will be primarily interested in characterizing *the error exponents*[2] (*i.e.,* the rate of decay)

$$-\lim_{n\to\infty} \frac{1}{n} \log \alpha_n, \quad -\lim_{n\to\infty} \frac{1}{n} \log \beta_n \tag{5}$$

of the Type-I and Type-II error probabilites in various situations of interest.

In the TBMA scheme, the statistic used for hypothesis testing is the *inner product* (notation $\langle \cdot, \cdot \rangle$) between z and

[2]Throughout the paper, the notation log refers to the natural logarithm.

the waveforms $s_1, \cdots, s_k$. Let

$$y := \frac{1}{\sqrt{E}n}[\langle z, s_1\rangle \ \cdots \ \langle z, s_k\rangle]^T \quad (6)$$

$$= \frac{1}{n}\sum_{i=1}^{n} h_i e_{X_i} + \tilde{w}, \quad (7)$$

where $e_1, \cdots, e_k$ are the standard basis vectors, and $\tilde{w} \sim \mathcal{N}(0, \frac{\sigma^2}{En^2}I)$. In order to compute asymptotic error probabilities, we need to understand the asymptotics of the random vector y. The theory of large deviations characterizes the probability of large excursions of y from its "mean" behavior by quantifying by the so-called rate function, which is defined below.

Definition For a set $B \subset \mathbb{R}^k$, let int(B) denote the interior of B and cl(B) denote the closure of B. The sequence of random variables y for $n = 1, 2, \cdots$ is said to satisfy the *large deviations principle* with *rate function* I if for any measurable set B

$$\begin{aligned} -\inf_{x\in\text{int}(B)} I(x) &\leq \liminf_{n\to\infty}\frac{1}{n}\log\Pr(y\in B) \\ &\leq \limsup_{n\to\infty}\frac{1}{n}\log\Pr(y\in B) \quad (8) \\ &\leq -\inf_{x\in\text{cl}(B)} I(x), \end{aligned}$$

where $I : \mathbb{R}^k \to \mathbb{R}_+ \cup \{\infty\}$. The effective domain of the function I is defined as $\mathcal{D}_I = \{x : I(x) < \infty\}$.

Remark 1: The sets of interest B in hypothesis testing mostly satisfy the so-called *I-continuity* property:

$$\inf_{x\in\text{int}(B)} I(x) = \inf_{x\in\text{cl}(B)} I(x), \quad (9)$$

which implies

$$\lim_{n\to\infty}\frac{1}{n}\log\Pr(y\in B) = -\inf_{x\in B} I(x).$$

The rate function I admits the following interpretation. For an integer k and $r \in \mathbb{R}^k$, let $B_\epsilon(r)$ be the open ball in $\mathbb{R}^k$ centered at r with radius $\epsilon > 0$. If I is continuous in the interior of its domain $\mathcal{D}_I$, then it satisfies the condition in Remark 1 for all balls $B = B_\epsilon(r)$ inside $\mathcal{D}_I$. The essence of the large deviations principle is that[3]

$$\Pr\{y \in B_\epsilon(x)\} \doteq e^{-n(I(x)+O(\epsilon))}, \quad x \in \text{int}(\mathcal{D}_I), \quad (10)$$

where $O(\epsilon)$ is a function that goes to zero as $\epsilon \to 0$. In other words, the probability that the y turns out to be in the close vicinity of x behaves as $e^{-nI(x)}$.

[3]The "$\doteq$" notation in (10) means $\lim_{n\to\infty}\frac{1}{n}\log\Pr\{y \in B_\epsilon(x)\} = -n(I(x)+O(\epsilon))$. This notation should be understood similarly in the rest of the paper.

We will consider the detection error exponents with the ML detector. The exact computation of the likelihood function of y, however, is generally intractable in our setup. To alleviate the problem, we propose a variant of the ML detector as follows. Suppose y satisfies LDP with rate function I_i under hypothesis $\theta = \theta_i$. We define the *minimum-rate detector* as the decision rule with decision regions

$$\Gamma_0 = \{x \in \mathbb{R}^k : I_0(x) \leq I_1(x)\}, \qquad \Gamma_1 = \mathbb{R}^k \setminus \Gamma_0. \quad (11)$$

Here, the detector decides that $\mathcal{H}_0$ is true if $e^{-nI_0(y)} \geq e^{-nI_1(y)}$ holds (*i.e.,* the asymptotic likelihood of y under $\mathcal{H}_0$ is higher). The decision region Γ_1 can be interpreted similarly. We expect the error exponents of this detector to be same as that of the exact ML detector.

Suppose we restrict ourselves to the class of detectors $\mathcal{C}$ based on y alone (and not on n directly). Define a detector in $\mathcal{C}$ to be *max-min optimal* if it maximizes the minimum of the Type-I and Type-II error exponents amongst all detectors in $\mathcal{C}$. Note that if a detector is max-min optimal, then it also has the best exponent for probability of error in a bayesian setting since the probability of error decays exponentially with the lower of the Type-I and Type-II exponents (the so called *worst exponent wins rule* [6]). Next, we provide some results on the asymptotic performance of the minimum-rate detector.

Theorem 1 *Suppose that y satisfies the LDP principle with rate functions I_0 and I_1 under hypthesis $\theta = \theta_0$ and $\theta = \theta_1$ respectively. Let Γ_0 (defined in (11)) be I_1-continuous and Γ_1 be I_0-continuous (eqn. (9)). Then,*
i) *The minimum-rate detector is max-min optimal in $\mathcal{C}$.*
ii) *The error exponents of the minimum-rate detector are given by*

$$-\lim_{n\to\infty}\frac{1}{n}\log\alpha_n = \inf_{x\in\Gamma_1} I_0(x), \quad (12)$$

$$-\lim_{n\to\infty}\frac{1}{n}\log\beta_n = \inf_{x\in\Gamma_0} I_1(x). \quad (13)$$

If the infimums in (12) and (13) are attained at the boundary $\partial\Gamma_1 = \{x : I_0(x) = I_1(x)\}$, then the exponents of α_n and β_n are the same and equal to

$$\eta := \inf_{x\in\partial\Gamma_1} I_0(x).$$

Proof Refer to [7].

Next, we will characterize the rate functions of y under varying assumptions on the data and channel statistics.

3. ERROR EXPONENTS IN FADING CHANNELS

3.1. Conditionally i.i.d. data and non-zero mean i.i.d. channels

Let $X_1, \cdots, X_n$ be conditionally i.i.d. with pmf p_θ. Suppose that the channel gains $h_1, \cdots, h_n$ are i.i.d. and independent of $X_1, \cdots, X_n$. In this section, we will show that the error probabilities $\alpha_n,\ \beta_n$ *decay exponentially* with the network size n under the condition that the channel gains have *non-zero* mean $h := \mathbb{E}(h_i)$ with the minimum-rate detector.

From the law of large numbers and Slutsky's Theorem [8], it follows that in this case

$$y \to \mathbb{E}(h_i e_{X_i}) = h p_\theta, \tag{14}$$

in probability as $n\to\infty$, where p_θ is viewed as a vector $[p_\theta(1)\ \cdots\ p_\theta(k)]^T$. Hence, for large n one would expect to have $y \approx hp_{\theta_i}$ under hypothesis $\mathcal{H}_i$. Detection errors typically happen when y is close neither to hp_{θ_0} nor to hp_{θ_1}.

The next theorem characterizes the rate function of y. We adopt the following notations from [9]: $D(Q||P)$ denotes the *relative entropy* between the probability density functions (pdfs) Q and P. For random variables X and Y, $D(Y||X)$ denotes the relative entropy between the pdfs of Y and X.

Theorem 2 *Suppose that the moment generating function of $h_1, h_2, \cdots$ satisfies*

$$\varphi(t) = \mathbb{E}e^{th_i} < \infty, \quad \forall t \in \mathbb{R}. \tag{15}$$

Let $I_h : \mathbb{R}\to\mathbb{R}_+ \cup \{\infty\}$ be the function defined as

$$I_h(r) \quad = \inf_{\tilde{h}:\ \mathbb{E}(\tilde{h})=r} D(\tilde{h}||h_i), \quad r \in \mathbb{R}, \tag{16}$$

where the minimization is over real valued random variables $\tilde{h}$. Similarly, define $I : \mathbb{R}^k\to\mathbb{R}_+ \cup \{\infty\}$

$$I(x) \quad = \quad \inf_{\tilde{p}} \{D(\tilde{p}||p_\theta) + \sum_{j=1}^{k} \tilde{p}_j I_h(\frac{x_j}{\tilde{p}_j})\}, \tag{17}$$

for each $x \in \mathbb{R}^k$, where the minimization is over all probability vectors $\tilde{p} \in \mathbb{R}^k$. Then, y satisfies the large deviations principle with the rate function I.

Proof The proof uses the Gärtner-Ellis Theorem presented in Section 3.2. For details refer to [7].

3.2. Non-I.I.D. Data and Channels

In general as mentioned before, sensor observations and channel gains may be dependent. In this section, we provide a generalization of the results of Section 3.1 for possibly dependent and not identically distributed data and channels.

Let $\{(X_i, h_i)\}_{i=1}^{\infty} = \{(X_1, h_1), (X_2, h_2), \cdots\}$ be a sequence of $\{1, \cdots, k\} \times \mathbb{R}$ valued random vectors. The distribution of $\{(X_i, h_i)\}_{i=1}^{\infty}$ is determined by whether $\mathcal{H}_0$ or $\mathcal{H}_1$ is the correct hypothesis. Let

$$\varphi(t) = \mathbb{E}e^{\langle t,y\rangle}, \quad t \in \mathbb{R}^k, n \in \mathbb{N}$$

be the moment generating function of y. The following theorem generalizes Theorem 2.

Theorem 3 (Gärtner-Ellis Theorem [6]) *Suppose that*

$$\lim_{n\to\infty} \frac{1}{n} \log \varphi(nt) = \Lambda(t), \quad \forall t \in \mathbb{R}^k \tag{18}$$

exists as an extended real number. If Λ is an essentially smooth, lower-semicontinuous function,[4] then, y satisfies the Large Deviations Principle with rate function

$$\Lambda^*(x) = \sup_{t\in\mathbb{R}^k} [\langle x, t\rangle - \Lambda(t)], \quad x \in \mathbb{R}^k.$$

The function Λ^* is called the Legendre Transform *of* Λ.

Remark 3: The above theorem does not require any independence, etc. assumption on the sequence $\{(X_i, h_i)\}_{i=1}^{\infty}$. It only needs the existence of the *asymptotic log moment generating function (mgf)* (18) and its continuity, smoothness. Theorem 3 also characterizes the error exponents of the proposed minimum-rate detector.

4. EXAMPLES AND SIMULATION RESULTS

4.1. ON/OFF Channels

Consider the ON/OFF channel, *i.e.,* h_i is Bernoulli $\{0, 1\}$ distributed with mean h. For this scenario, $I_h(\cdot)$ is the relative entropy function between two Bernoulli variables. Using the Lagrange multipliers method, it is easy to get

$$I(x) = XD(\tilde{x}||p_\theta) + X \log \frac{X}{h} + (1 - X) \log \frac{1 - X}{1 - h},$$

where $X = h \sum_{j=1}^{k} x_j$ and $\tilde{x} = x/(\sum_{j=1}^{k} x_j)$. Using Theorem 1, the error exponent of α_n is obtained as

$$\eta = X^*C + X^* \log \frac{X^*}{h} + (1 - X^*) \log \frac{1 - X^*}{1 - h}, \tag{19}$$

where

$$C = -\min_{0\le\lambda\le1} \log(\sum_{j=1}^{k} p_{\theta_0}^{\lambda}(j)p_{\theta_1}^{1-\lambda}(j))$$

is the so-called *Chernoff information* [9], and $X^* = he^C/(1-h + he^C)$.

Simulation results for the hypotheses

$$\mathcal{H}_0 : X_i \overset{\text{i.i.d.}}{\sim} p_{\theta_0} = [0.8\ 0.2], \ \ \mathcal{H}_1 : X_i \overset{\text{i.i.d.}}{\sim} p_{\theta_1} = [0.2\ 0.8] \tag{20}$$

are given in Fig. 1. The LD (Large Deviations) estimate refers to $e^{-n\eta}$. SNR $= E/\sigma^2 =$ 3dB. For the other orthogonal method (TDMA), the antipodal constellation and the ML detector based the received signal are used (not based on decoded symbols). Some remarks are in order:

[4]See [6] for precise meanings of these terms.

i) In the On/OFF channel, unlike the minimum-rate detector, the exact ML detector is computationally intractable.

ii) TBMA outperforms the TDMA scheme in channels with non-zero mean.

iv) The Chernoff information C is the optimal exponent obtained when the fusion center has direct access to $X_1, \cdots, X_n$ [9]. From (19), it is seen that $\eta \to C$ as $h \to 1$. In other words, the asymptotic performance of the TBMA scheme approaches the optimal one as the channel fading disappears.

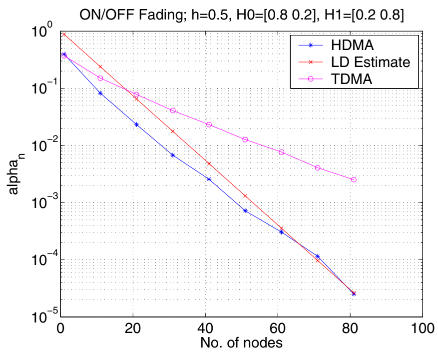

Fig. 1. ON/OFF channel with binary observations.

4.2. Gausssian Channels

Suppose that $h_1, \cdots, h_n$ are i.i.d. Gaussian with certain mean and variance. Depending on the mean, variance and p_{θ_i}'s the Gaussian channels present a rich set of behaviors. In the Gaussian channel $h_i \sim \mathcal{N}(\mu, \sigma_h^2)$ with binary observations, the vector y has a mixed Gaussian distribution, and the likelihood function of y can be computed numerically, unlike the situation, for example, in ON/OFF channels.

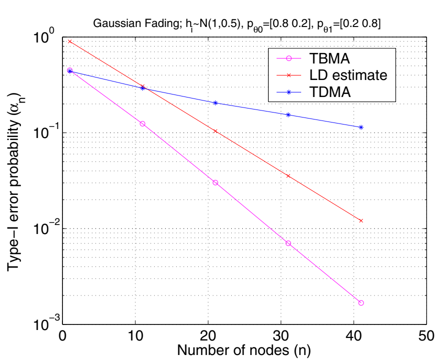

Fig. 2. Non-zero mean Gaussian channel.

Figure 2 shows the error probabilities of TBMA and TDMA schemes when the channel has non-zero mean ($\mu = 1, \sigma_h^2 = 0.5$, SNR $= -20$dB). The large deviations estimate of α_n is rather coarse in this channel; the slope of the curve $e^{-\eta n}$ appears to be same as the slope of α_n, however there is a non-vanishing gap between $e^{-\eta n}$ and α_n. The TBMA error performance again outperforms TDMA.

5. CONCLUSIONS

In this paper, we have considered the transmission of sensor observations over a multiaccess fading channel for the purpose of detection. We analyzed the performance of the Type-Based Multiple Access using large deviations theory. An asymptotic version of the Maximum Likelihood detector is proposed, and its error exponents are characterized. Simulation results are presented to validate the theoretical findings and check the accuracy of the large deviations approximations. For the detection scenarios considered, TBMA is significantly more bandwidth efficient than the conventional approaches such as TDMA, FDMA, CDMA.

6. REFERENCES

[1] G. Mergen and L. Tong, "Estimation over deterministic multiaccess channels," in *42nd Annual Allerton Conf. on Commun., Control and Comp.*, Oct. 2004.

[2] Ke Liu and A. M. Sayed, "Optimal distributed detection strategies for wireless sensor networks,," in *42nd Annual Allerton Conf. on Commun., Control and Comp.*, Oct. 2004.

[3] J.-F. Chamberland and V. V. Veeravalli, "Asymptotic results for decentralized detection in power constrained wireless sensor networks," *IEEE JSAC Special Issue on Sensor Networks*, 2004.

[4] B. Chen, R. Jiang, T. Kasetkasem, and P.K. Varshney, "Fusion of decisions transmitted over fading channels in wireless sensor networks," in *the 36th Asilomar Conference*, 2002.

[5] G. Mergen and L. Tong, "Sensor-fusion center communication over multiaccess fading channels," submitted to *ICASSP'05*, Sep. 2004.

[6] F. Den Hollander, *Large Deviations (Fields Institute Monographs, 14)*, American Mathematical Society, 2000.

[7] G. Mergen, V. Naware, and L. Tong, "Asymptotic detection performance of type-based multiple access over multiaccess fading channels," *submitted to IEEE Trans. Sig. Proc.*, May 2005.

[8] P. Billingsley, *Probability and Measure 3rd. Ed.*, Wiley Inter-Science, New York, NY, 1995.

[9] T. Cover and J. Thomas, *Elements of Information Theory*, John Wiley & Sons, Inc., 1991.