2011 50th IEEE Conference on Decision and Control and
European Control Conference (CDC-ECC)
Orlando, FL, USA, December 12-15, 2011

Channel Estimation for Free-Space Optical Communication

Arash Komaee

***Abstract*— Optical fade caused by atmospheric turbulence impairs the performance of optical communication through atmosphere. Adaptive techniques of signal detection, power control, and channel coding can be employed to reduce the degrading effect of this phenomenon. For implementation of these techniques, the knowledge of channel state is required. This paper develops a channel estimator to extract the strength of optical fade from the observations of channel output. Further, it is shown how to incorporate this estimation in an adaptive threshold test for the purpose of optimal binary signal detection.**

I. INTRODUCTION

Atmospheric turbulence caused by differential heating of air results in random variations in the refractive index of this optical medium. This, in turn, introduces random fluctuations in the intensity and phase of an optical beam propagating through atmosphere [1]–[3]. This phenomenon is a major limitation in optical communication over atmospheric channels, since the turbulence induced optical fade can significantly reduce the signal-to-noise ratio (SNR) at the receiver.

The performance degrading effect of optical fade can be partially compensated using adaptive techniques of signal detection, power control, and channel coding. These techniques mainly rely on the availability of channel state information [4]–[6]. The channel state (strength of optical fade) is estimated at the receiver and the resulting estimate is either used locally for optimal signal detection or is sent to the transmitter (via a feedback channel) for adaptive power control or channel coding.

The major contribution of this paper is the development of a low-complexity channel estimator for atmospheric optical channels with on-off keying (OOK) modulation. In addition, an optimal scheme to incorporate the channel estimation in binary signal detection is established. Compared to our prior results on this topic [7], the channel estimator developed in this paper requires a much slower update rate which allows for a more efficient implementation. This improvement is achieved by using two different sampling rates for detection and channel estimation.

II. MODEL AND PROBLEM STATEMENT

We consider an atmospheric optical channel which transmits a sequence of binary messages by OOK modulation of an optical beam. As the modulated optical beam propagates through atmosphere, the turbulence induced optical fade randomly changes its intensity. The optical fade is described by a slowly varying lognormal process [2], [3], [8]–[10]. At the receiving terminal, the optical beam strikes a photodetector which converts optical power into an electrical signal. This signal is regarded as the channel output and is further processed to reconstruct the transmitted message and to estimate the channel state. In this section, the channel model is presented and its associated channel estimation and signal detection problems are stated.

The channel estimator and signal detector in this paper operate at different sampling periods T_s and T_b, respectively. Here, T_b is the bit transmission time and is much smaller than the sampling period T_s of the channel estimator. For simplicity of analysis, it will be assumed that $T_s/T_b = N$ is an integer. A discrete-time signal with the sampling period T_s will be shown by s_k while $s[n]$ stands for a signal with the sampling period T_b.

A. Transmitter and Receiver

Let $\{m[n]\}_{n=1}^{\infty}$ be a sequence of binary messages with values in $\{0, 1\}$ and construct the continuous-time signal

$$m(t) = \sum_{n=1}^{\infty} m[n] \, \Pi(t - nT_b + T_b), \tag{1}$$

where $\Pi(\cdot)$ is a rectangular pulse defined as

$$\Pi(t) = \begin{cases} 1, & 0 \leqslant t \leqslant T_b \\ 0, & \text{otherwise.} \end{cases}$$

To transmit this signal, an optical source with the maximum power $P_M > 0$ is employed at the transmitter and its instantaneous power $P_T(t)$ is modulated by the message $m(t)$ as

$$P_T(t) = P_M m(t).$$

The generated optical beam propagates through the turbulent atmosphere and at a distance from the transmitter strikes the optical receiver. Under the assumptions which will be discussed later in this section, the power fluctuations caused by atmospheric turbulence can be described by a multiplicative nonnegative stochastic process $\{\alpha(t)\}$. The antenna gain and path attenuation can be also characterized by a multiplicative constant $G > 0$. Thus, the instantaneous optical power at the receiver is given by

$$P_R(t) = G\alpha(t) P_T(t) = GP_M \alpha(t) m(t). \tag{2}$$

The receiver is equipped with a direct-detection photodetector to convert the absorbed optical power to an electrical signal. The output $s(t)$ of this device is a stream of small

This work was supported in part by the Army Research Office under ODDR&E MURI01 Program Grant No. DAAD19-01-1-0465 to the Center for Communicating Networked Control Systems (through Boston University).

The author is with the Department of Aerospace Engineering, University of Maryland, College Park, MD 20742, USA, akomaee@umd.edu.

electrical pulses each one representing the absorption of a photon from the incident optical field. This signal can be ideally modeled by a Poisson impulse process, defined as the derivative of a counting Poisson process [3], [8], [11]–[13]. The rate of this process is the sum of two terms: a signal term proportional to the instantaneous optical power and a constant term representing dark current noise (caused by thermal effects) and background radiations. Mathematically, the rate of the Poisson process is given by

$$\lambda(t) = \eta P_R(t) + \lambda_b, \tag{3}$$

where $\eta > 0$ is the sensor sensitivity and $\lambda_b > 0$ stands for the dark current and background noise. In the present model where the rate is a stochastic process rather than a deterministic function, the Poisson process must be replaced by a doubly stochastic (conditionally) Poisson process [14].

B. Statistical Characterization of Optical Fade

Atmospheric turbulence caused by differential heating of air results in random variations in the refractive index of this optical medium. The refractive index at every point r can be modeled by $n_a(r) = \bar{n}_a + \delta n_a(r)$, where the constant $\bar{n}_a$ is the mean refractive index and the stochastic field $\delta n_a(r)$ characterizes the fluctuations caused by turbulence. The statistical properties of $\delta n_a(r)$ are predicted by Kolmogorov's turbulence theory [1], [8], [9].

Rytov's method on approximate solution of Maxwell's equations is frequently used to study propagation of optical fields in turbulent atmosphere [8], [9]. This method expresses the complex amplitude of the optical field as

$$U(r) = T(r)\, U_0(r), \tag{4}$$

where $U_0(r)$ denotes the complex amplitude in the absence of turbulence and $T(r)$ is a stochastic field representing the fluctuation term $\delta n_a(r)$. Using Rytov's method, the stochastic field $\chi(r) \triangleq \ln|T(r)|$ is expressed as a weighted integral of $\delta n_a(r)$. Applying the central limit theorem [15] to this linear form, $\chi(r)$ is approximated by a Gaussian random variable which leads to a lognormal distribution [8], [9] for the fade coefficient

$$\exp(2\chi(r)) = \frac{|U(r)|^2}{|U_0(r)|^2}.$$

The variance σ_χ^2 of $\chi(r)$ is obtained from Kolmogorov's turbulence theory and depends on the wavelength of light, propagation distance, refractive index structure constant, and the shape of optical field [9]. For a 1 km optical link, the numerical value of σ_χ varies from 10^{-2} to 1, depending on the strength of turbulence [16]. The fact that turbulence does not absorb energy implies [9] that $\mathrm{E}[\exp(2\chi(r))] = 1$ which requires that $\mathrm{E}[\chi(r)] = -\sigma_\chi^2$.

For an isotropic turbulence, the spatial autocorrelation

$$b_\chi(\|r\|) = \frac{\mathrm{Cov}(\chi(r_1 + r), \chi(r_1))}{\sigma_\chi^2}$$

is obtained in [1], [9] as

$$b_\chi(\|r\|) = \exp\left(-3.44\,(\|r\|/r_0)^{5/3}\right).$$

Here, r_0 is the Fried parameter or turbulence coherence length which can be calculated from the explicit expressions provided in [1], [9]. This parameter plays an important role in our analysis because when the diameter of the receiving aperture is small enough compared to r_0, the stochastic field $\chi(r)$ is highly uniform over the aperture such that it can be approximated by a random variable, not depending on spatial coordinate. This condition is normally valid for short range applications on the order of 1 km length and under the weak to moderate turbulence regime [9]. Noting that the total received optical power is the integral of $|U(r)|^2$ over the aperture, and since $|T(r)|^2$ does not depend on r under this condition, the received optical power can be simply obtained by integrating $|U_0(r)|^2$ over the aperture and then multiplying the result by a random variable. This random variable is regarded as optical fade and is given by $\exp(2\chi(\bar{r}))$, where $\bar{r}$ is the center of the receiving aperture.

Since optical fade is a time-dependent phenomenon, we need to mathematically characterize its temporal variations in order to complete our model. Although a full description for the temporal evolution of optical fade has not been proposed yet, at least the autocorrelation of its logarithm $\chi(r,t)$ can be approximated using Taylor's frozen-flow hypothesis [1]. According to this hypothesis, the refractive index of air as a function of position r and time t can be represented as

$$n_a(r,t) = \bar{n}_a + \delta n_a(r - vt),$$

where v is the component of local wind perpendicular to the light propagation direction. From this representation, it is easy to determine the temporal autocorrelation $\rho(\tau)$ in terms of the spatial autocorrelation $b_\chi(\cdot)$ as

$$\rho(\tau) = \frac{\mathrm{Cov}(\chi(r, t+\tau), \chi(r,t))}{\sigma_\chi^2} = b_\chi(\|v\| \cdot |\tau|).$$

This function can be explicitly written as

$$\rho(\tau) = \exp\left(-(|\tau|/\tau_0)^{5/3}\right),$$

where $\tau_0 = 0.477 r_0/\|v\|$ is the fade coherence time.

In the rest of this subsection, we obtain a discrete-time Markov process to represent the optical fade. Assume that the sampling period T_s is much smaller than the fade coherence time τ_0 (e.g., 1 ms versus 10 ms) such that $\alpha(t)$ is almost constant over $[kT_s - T_s, kT_s)$. Thus, for every t in this interval, $\alpha(t)$ can be approximated by a random variable $\alpha_k = \alpha(kT_s)$. To determine a Markov model for the discrete-time process $\{\alpha_k\}$, let $\{x_k\}$ be a discrete-time stochastic process satisfying two conditions: first, x_k is a zero mean unit variance Gaussian random variable, and second, $\mathrm{E}[x_{k'+k} x_{k'}] = \rho(kT_s)$ holds for every integers k and k'. Noting that $\chi(\bar{r})$ can be expressed as $\chi(\bar{r}) = \sigma_\chi x_k - \sigma_\chi^2$, the lognormal fade $\alpha_k = \exp(2\chi(\bar{r}))$ is represented by

$$\alpha_k = a(x_k), \tag{5}$$

where the mapping $a(\cdot): \mathbb{R} \to \mathbb{R}^+$ is defined as

$$a(x) = \exp\left(2\sigma_\chi x - 2\sigma_\chi^2\right).$$

The first condition on $\{x_k\}$ is easy to maintain if this stochastic process is obtained from a linear dynamical system driven by white Gaussian noise and with a Gaussian initial state. With a finite-dimensional linear system, the second condition can be only approximated. In this study, we follow [10] in using a first order system to approximate $\{x_k\}$.

Let $\{x_k\}_{k=1}^{\infty}$ be the solution of the state-space equation

$$x_{k+1} = x_k \cos\epsilon + w_k \sin\epsilon, \tag{6}$$

where $\{w_k\}$ is a zero mean unit variance white Gaussian noise and ϵ is a constant. The initial state x_1 is a zero mean unit variance Gaussian random variable independent of $\{w_k\}$. Then, $\{x_k\}$ is a sequence of zero mean unit variance Gaussian random variables with the autocorrelation function $\mathrm{E}[x_{k'+k}x_{k'}] = (\cos\epsilon)^k$ and the transition probability density function

$$q_x(\xi|\zeta) = p_{x_{k+1}}(\xi|x_k = \zeta) = \Phi\left(\xi; \zeta\cos\epsilon, \sin^2\epsilon\right).$$

Here, $\Phi\left(x; \bar{x}, \sigma_x^2\right)$ denotes a Gaussian density function with mean $\bar{x}$ and variance σ_x^2. We solve a least squares problem to obtain the value of ϵ which minimizes the distance between the autocorrelation function $(\cos\epsilon)^k$ and $\rho(kT_s)$. The solution to this optimization problem is given by

$$\cos\epsilon = \exp(-1.07 T_s/\tau_0).$$

Fig. 1 compares the exact autocorrelation function $\rho(kT_s)$ with its approximation $(\cos\epsilon)^k$.

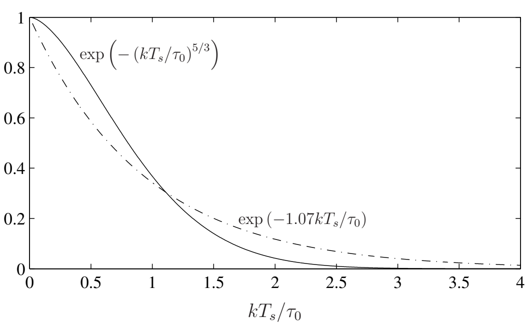

Fig. 1. Comparison between the exact (solid line) and the approximate (dashed line) autocorrelation functions.

C. Sampling the Channel Output

The channel estimator is a discrete-time system with the input signal $\{s_k\}$ defined for $k = 1, 2, 3, \ldots$ as

$$s_k = \int_{kT_s - T_s}^{kT_s} s(t)\,dt.$$

Since $\{s(t)\}$ is a doubly stochastic Poisson process, its integral over any arbitrary interval is a conditionally Poisson random variable. Thus, s_k is distributed according to

$$\Pr\{s_k = \ell|\lambda_k\} = \frac{e^{-\lambda_k}\lambda_k^{\ell}}{\ell!}, \quad \ell = 0, 1, 2, \ldots,$$

where the Poisson rate λ_k is defined as

$$\lambda_k = \int_{kT_s - T_s}^{kT_s} \lambda(t)\,dt.$$

Substituting (1), (2), and (3) into this integral and approximating $\alpha(t)$ with α_k over the integration interval, λ_k can be approximated as

$$\lambda_k = \eta G P_M T_b \alpha_k \left(\sum_{n=1}^{N} m[(k-1)N + n]\right) + \lambda_b N T_b.$$

Replacing α_k with (5) in this expression, it can be written in terms of the state variable x_k as

$$\lambda_k = \mu a(x_k) m_k + Nb,$$

where $\mu = \eta G P_M T_b$, $b = \lambda_b T_b$, and $\{m_k\}$ is a discrete-time signal defined as

$$m_k = \sum_{n=1}^{N} m[(k-1)N + n].$$

In a similar manner, $s[n]$ and $\lambda[n]$ are defined as the integrals of $s(t)$ and $\lambda(t)$ over the interval $[nT_b - T_b, nT_b)$. Then, conditioned on $\lambda[n]$, the random variable $s[n]$ is Poisson distributed according to

$$\Pr\{s[n] = \ell|\lambda[n]\} = \frac{e^{-\lambda[n]}\lambda^{\ell}[n]}{\ell!}, \quad \ell = 0, 1, 2, \ \ldots.$$

Here, the rate $\lambda[n]$ can be explicitly written as

$$\lambda[n] = \mu a(x_{\nu+1}) m[n] + b,$$

where $\nu = \lfloor (n-1)/N \rfloor$ and $\lfloor\cdot\rfloor$ denotes the floor function.

D. Problem Statement

Suppose the sequence of binary messages $\{m[n]\}$ are independent[†] random variables with the probability distribution

$$\Pr\{m[n] = 1\} = 1 - \Pr\{m[n] = 0\} = p,$$

where $0 < p < 1$ is a known constant. Assume that the discrete-time signal $\{s_i\}$ is observed during $1 \leqslant i \leqslant k$ and define the observation set $I_k = (s_1, s_2, \cdots, s_k)$ for $k \geqslant 1$ and $I_0 = \emptyset$. In this framework, the following problems are considered:

P-1) Given the observation set I_k, for any arbitrary $k \geqslant 0$ determine the posterior density

$$\pi_k(\xi) = p_{x_{k+1}}(\xi|I_k)$$

and then compute the minimum mean squared error estimation $\hat{\alpha}_k = \mathrm{E}[\alpha_{k+1}|I_k]$ from the integral[‡]

$$\hat{\alpha}_k = \int_{-\infty}^{+\infty} \pi_k(\xi)\, a(\xi)\, d\xi. \tag{7}$$

P-2) Given $s[n]$ and the observation set I_ν, for any integer n determine the detection rule

$$\hat{m}[n] = \mathcal{D}_\nu(s[n], I_\nu) \in \{0, 1\}$$

which minimizes the probability of error

$$P_e = \Pr\{\hat{m}[n] \neq m[n]\}. \tag{8}$$

[†]The case that the binary messages are statistically dependent due to channel coding can be treated with some more efforts.

[‡]The notation $\hat{\alpha}_{k+1|k}$ might better describe $\mathrm{E}[\alpha_{k+1}|I_k]$; however, for the sake of simplicity, we use $\hat{\alpha}_k$ here.

III. Optimal Estimation and Detection

Theorems 1 and 2 below present the solutions to the channel estimation (P-1) and signal detection (P-2) problems.

Theorem 1: Define the mapping $\theta(\cdot): \mathbb{R} \times \mathbb{Z}^+ \to \mathbb{R}^+$ as

$$\theta(x, s) = \sum_{\ell=0}^{N} B(\ell; N, p)\, e^{-\ell\mu a(x)} \left(1 + (\ell/N)\,\beta a(x)\right)^s,$$

where $\beta = \mu/b$ and $B(\ell; N, p)$ is the binomial distribution

$$B(\ell; N, p) = \frac{N!}{\ell!\,(N-\ell)!}\, p^{\ell} (1-p)^{N-\ell}.$$

Consider the model of Section II and assume that for every $k \geqslant 1$, the observation set I_k is provided to estimate x_{k+1}. Then, the posterior density $\pi_k(\xi) = p_{x_{k+1}}(\xi|I_k)$ is determined from the recursive equation

$$\pi_k(\xi) = \frac{\int_{-\infty}^{+\infty} q_x(\xi|\zeta)\, \theta(\zeta, s_k)\, \pi_{k-1}(\zeta)\, d\zeta}{\int_{-\infty}^{+\infty} \theta(\zeta, s_k)\, \pi_{k-1}(\zeta)\, d\zeta} \tag{9}$$

with the initial density $\pi_0(\xi) = \Phi(\xi; 0, 1)$.

Proof: Assuming that $p_{x_k}(\xi|I_k)$ is known, $p_{x_{k+1}}(\xi|I_k)$ is obtained from

$$p_{x_{k+1}}(\xi|I_k) = \int_{-\infty}^{+\infty} p_{x_{k+1}}(\xi|x_k = \zeta, I_k)\, p_{x_k}(\zeta|I_k)\, d\zeta.$$

Noting that

$$p_{x_{k+1}}(\xi|x_k = \zeta, I_k) = p_{x_{k+1}}(\xi|x_k = \zeta) = q_x(\xi|\zeta),$$

this equation can be rewritten as the state update equation

$$p_{x_{k+1}}(\xi|I_k) = \int_{-\infty}^{+\infty} q_x(\xi|\zeta)\, p_{x_k}(\zeta|I_k)\, d\zeta. \tag{10}$$

To obtain the measurement update equation, i.e., the equation which determines $p_{x_k}(\xi|I_k)$ in terms of $p_{x_k}(\xi|I_{k-1})$ and s_k, Bayes' rule is employed to write

$$\begin{aligned} p_{x_k}(\xi|I_k) &= p_{x_k}(\xi|s_k, I_{k-1}) \\ &= \frac{\Pr\{s_k|x_k = \xi, I_{k-1}\}\, p_{x_k}(\xi|I_{k-1})}{\Pr\{s_k|I_{k-1}\}} \\ &= \frac{\Pr\{s_k|x_k = \xi, I_{k-1}\}\, p_{x_k}(\xi|I_{k-1})}{\int_{-\infty}^{+\infty} \Pr\{s_k|x_k = \zeta, I_{k-1}\}\, p_{x_k}(\zeta|I_{k-1})\, d\zeta}. \end{aligned} \tag{11}$$

The conditional probability distribution in the last equality can be expressed as

$$\begin{aligned} \Pr\{s_k|x_k = \xi, I_{k-1}\} &= \Pr\{s_k|x_k = \xi\} \\ &= \sum_{\ell=0}^{N} \Pr\{s_k|m_k = \ell, x_k = \xi\} \Pr\{m_k = \ell|x_k = \xi\} \\ &= \frac{e^{-Nb}(Nb)^{s_k}}{s_k!}\, \theta(\xi, s_k). \end{aligned}$$

This result is concluded from the fact that m_k is independent of x_k and has a binomial distribution $B(\ell; N, p)$, and from the fact that conditioned on $m_k = \ell$ and $x_k = \xi$, s_k is a Poisson random variable with the rate $\mu a(\xi)\,\ell + Nb$. Substituting this result and $\pi_{k-1}(\xi) = p_{x_k}(\xi|I_{k-1})$ into (11), the measurement update equation is obtained as

$$p_{x_k}(\xi|I_k) = \frac{\theta(\xi, s_k)\, \pi_{k-1}(\xi)}{\int_{-\infty}^{+\infty} \theta(\zeta, s_k)\, \pi_{k-1}(\zeta)\, d\zeta}.$$

Replacing $p_{x_k}(\xi|I_k)$ from this equation into the state update equation (10) results in the recursive equation (9). ■

Theorem 2: Consider the model of Section II and assume that $s[n]$ and the observation set I_ν are provided for detection of $m[n]$. Then, the detection rule which minimizes the probability of error (8) is the threshold test

$$\hat{m}[n] = \begin{cases} 1, & L[n] \geqslant \gamma \\ 0, & L[n] < \gamma \end{cases} \tag{12}$$

with the threshold $\gamma = (1-p)/p$ and the likelihood function

$$L[n] = \int_{-\infty}^{+\infty} \pi_\nu(\xi)\, e^{-\mu a(\xi)} \left(1 + \beta a(\xi)\right)^{s[n]} d\xi, \tag{13}$$

where $\beta = \mu/b$.

Proof: It is shown in [17] that the detection rule which minimizes the probability of error (8) is the threshold test

$$\hat{m}[n] = \begin{cases} 1, & \Lambda[n] \geqslant 1 \\ 0, & \Lambda[n] < 1, \end{cases} \tag{14}$$

where

$$\Lambda[n] = \frac{\Pr\{m[n] = 1|s[n], I_\nu\}}{\Pr\{m[n] = 0|s[n], I_\nu\}}. \tag{15}$$

Application of Bayes' rule leads to

$$\Pr\{m[n]\,|s[n], I_\nu\} = \frac{\Pr\{s[n]\,|I_\nu, m[n]\} \Pr\{m[n]\,|I_\nu\}}{\Pr\{s[n]\,|I_\nu\}}.$$

Substituting this result into (15), $\Lambda[n]$ can be expressed as

$$\Lambda[n] = L[n] \cdot \frac{\Pr\{m[n] = 1|I_\nu\}}{\Pr\{m[n] = 0|I_\nu\}},$$

where the likelihood function $L[n]$ is given by

$$L[n] = \frac{\Pr\{s[n]\,|I_\nu, m[n] = 1\}}{\Pr\{s[n]\,|I_\nu, m[n] = 0\}}.$$

Noting that $m[n]$ is independent of I_ν, one can write

$$\Pr\{m[n] = 1|I_\nu\} = \Pr\{m[n] = 1\} = p.$$

Thus, the detection rule (14) can be written as (12).

The likelihood function $L[n]$ can be determined in terms of $s[n]$ and the posterior density $\pi_\nu(\cdot)$. For $m[n] = 0$, the channel output does not depend on $x_{\nu+1}$ which leads to

$$\Pr\{s[n] = \ell|I_\nu, m[n] = 0\} = \frac{e^{-b} b^{\ell}}{\ell!}. \tag{16}$$

For $m[n]=1$, one can write

$$\begin{aligned}\Pr\{s[n]=\ell|I_\nu,m[n]=1\}\\&=\int_{-\infty}^{+\infty}p_{x_{\nu+1}}(\xi|I_\nu,m[n]=1)\\&\quad\times\Pr\{s[n]=\ell|I_\nu,m[n]=1,x_{\nu+1}=\xi\}\,d\xi\\&=\int_{-\infty}^{+\infty}\pi_\nu(\xi)\frac{e^{-(\mu a(\xi)+b)}(\mu a(\xi)+b)^\ell}{\ell!}d\xi,\end{aligned}\quad(17)$$

where the second equality is concluded from

$$\Pr\{s[n]|I_\nu,m[n],x_{\nu+1}\}=\Pr\{s[n]|m[n],x_{\nu+1}\}$$

and

$$p_{x_{\nu+1}}(\xi|I_\nu,m[n])=p_{x_{\nu+1}}(\xi|I_\nu)=\pi_\nu(\xi).$$

Finally, the likelihood function (13) can be obtained by dividing (17) by (16). ■

Remark 1: Representation (13) of the likelihood function explains the structure of optimal detector as a combination of a channel estimator and a single bit detector. The information provided by the past bits for detection of the present bit is accumulated in the posterior density $\pi_\nu(\cdot)$. If the perfect knowledge of the channel state is available, the likelihood function will be given by

$$L[n]=e^{-\mu\alpha_{\nu+1}}(1+\beta\alpha_{\nu+1})^{s[n]}.$$

In the absence of this knowledge, the likelihood function must be averaged over the posterior density of optical fade.

IV. Filtering Problem

Determining the posterior density of the log-amplitude fade from the recursive equation (9) requires an excessive computational load. To reduce this computational complexity, the posterior density can be approximated by means of a finite-dimensional nonlinear filter. Our proposed procedure for developing a nonlinear filter of this type is as follows. At $k=0$, the posterior density is Gaussian with mean $\hat{x}_0=0$ and variance $v_0=1$. Substitute $\pi_0(\xi)=\Phi(\xi;0,1)$ into the recursive equation (9) to compute $\pi_1(\xi)$, which is not necessarily Gaussian. Computing $\pi_2(\xi)$ from this non-Gaussian function is difficult. Alternatively, approximate $\pi_1(\xi)$ with a Gaussian function with mean $\hat{x}_1$ and variance v_1 and then compute $\pi_2(\xi)$ from this Gaussian approximation. Repeat these steps for every k to obtain recursive equations for evolution of the mean $\hat{x}_k$ and the variance v_k. Nonlinear filtering based on successive Gaussian approximations has been studied by several researchers [18]–[21].

The Gaussian approximation step in this filtering algorithm can be formulated in terms of determining a Gaussian density function $\Phi(\xi;\hat{x}_{k+1},v_{k+1})$ with the minimum distance from the probability density function

$$\tilde{\pi}_{k+1}(\xi)\triangleq\frac{\int_{-\infty}^{+\infty}q_x(\xi|\zeta)\,\theta(\zeta,s_{k+1})\,\Phi(\zeta;\hat{x}_k,v_k)\,d\zeta}{\int_{-\infty}^{+\infty}\theta(\zeta,s_{k+1})\,\Phi(\zeta;\hat{x}_k,v_k)\,d\zeta}.\quad(18)$$

Adopting relative entropy [22] as the measure of distance between $\tilde{\pi}_{k+1}(\xi)$ and its Gaussian approximation, this distance is minimized when the mean $\hat{x}_{k+1}$ and the variance v_{k+1} of the Gaussian approximation are chosen as

$$\begin{aligned}\hat{x}_{k+1}&=\int_{-\infty}^{+\infty}\tilde{\pi}_{k+1}(\xi)\,\xi d\xi\\v_{k+1}&=\int_{-\infty}^{+\infty}\tilde{\pi}_{k+1}(\xi)(\xi-\hat{x}_{k+1})^2\,d\xi.\end{aligned}\quad(19)$$

Lemma 1: Define the mappings $\phi_i(\cdot):\mathbb{R}\times\mathbb{R}^+\times\mathbb{Z}^+\to\mathbb{R}$ for $i=0,1,2$ as

$$\phi_i(x,v,s)=\int_{-\infty}^{+\infty}(\zeta-x)^i\,\theta(\zeta,s)\,\Phi(\zeta;x,v)\,d\zeta,\quad(20)$$

and based on that, define

$$\psi_1(x,v,s)=\frac{\phi_1(x,v,s)}{\phi_0(x,v,s)}$$

and

$$\psi_2(x,v,s)=\frac{\phi_2(x,v,s)}{\phi_0(x,v,s)}-\frac{\phi_1^2(x,v,s)}{\phi_0^2(x,v,s)}-v.$$

Then, for $\tilde{\pi}_{k+1}(\cdot)$ given by (18), the integrals in (19) can be expressed in terms of $\hat{x}_k$, v_k, and s_{k+1} as

$$\begin{aligned}\hat{x}_{k+1}&=\hat{x}_k\cos\epsilon+\psi_1(\hat{x}_k,v_k,s_{k+1})\cos\epsilon\\v_{k+1}&=v_k\cos^2\epsilon+\sin^2\epsilon+\psi_2(\hat{x}_k,v_k,s_{k+1})\cos^2\epsilon.\end{aligned}\quad(21)$$

Proof: The proof follows by substituting (18) into the integrals in (19), reversing the order of double integration, and noting that

$$\int_{-\infty}^{+\infty}q_x(\xi|\zeta)\,\xi d\xi=\zeta\cos\epsilon$$

$$\int_{-\infty}^{+\infty}q_x(\xi|\zeta)(\xi-\hat{x}_{k+1})^2\,d\xi=(\zeta\cos\epsilon-\hat{x}_{k+1})^2+\sin^2\epsilon.$$

■

The recursive equations (21) with the initial state $\hat{x}_0=0$ and $v_0=1$ represent a nonlinear filter which approximates the mean and variance of the posterior density (9). Based on these approximate values, an approximation of the posterior density is given by $\pi_k(\xi)\simeq\Phi(\xi;\hat{x}_k,v_k)$. This approximation can be replaced into (7) to obtain the channel estimate

$$\hat{\alpha}_k=\exp\left(2\sigma_\chi\hat{x}_k+2\sigma_\chi^2(v_k-1)\right).\quad(22)$$

In a similar manner, $\pi_\nu(\xi)$ in (13) can be approximated by the Gaussian density function $\Phi(\xi;\hat{x}_\nu,v_\nu)$ to obtain the likelihood function $L[n]$. When the variance v_ν is small, this integral can be further approximated as

$$L[n]=e^{-\mu\hat{\alpha}_\nu}(1+\beta\hat{\alpha}_\nu)^{s[n]}.$$

This expression for the likelihood function simplifies the detection rule (12) into the threshold test

$$\hat{m}[n]=\begin{cases}1,& s[n]\geqslant u[n]\\0,& s[n]<u[n],\end{cases}$$

with the adaptive threshold

$$u[n]=\frac{\ln\gamma+\mu\hat{\alpha}_\nu}{\ln(1+\beta\hat{\alpha}_\nu)}.$$

V. Simulation Results

We examined the performance of the proposed nonlinear filtering scheme via simulations. For this purpose, the stochastic process $\{x_k\}$ (log-amplitude fade) is generated from (6) and the estimators $\{\hat{x}_k\}$ and $\{v_k\}$ (mean and variance) are determined recursively from (21) with the initial state $\hat{x}_1 = 0$ and $v_1 = 0$. The fade estimation $\{\alpha_k\}$ is obtained from (22). The parameters used in the simulations are $\epsilon = 0.046$, $\sigma_\chi = 0.2$, $\mu = 50$, $b = 5$, $N = 1000$, and $p = 0.5$. The results of this study are illustrated in Fig. 2. Parts (a) and (b) of this figure indicate that the estimators $\hat{x}_k$

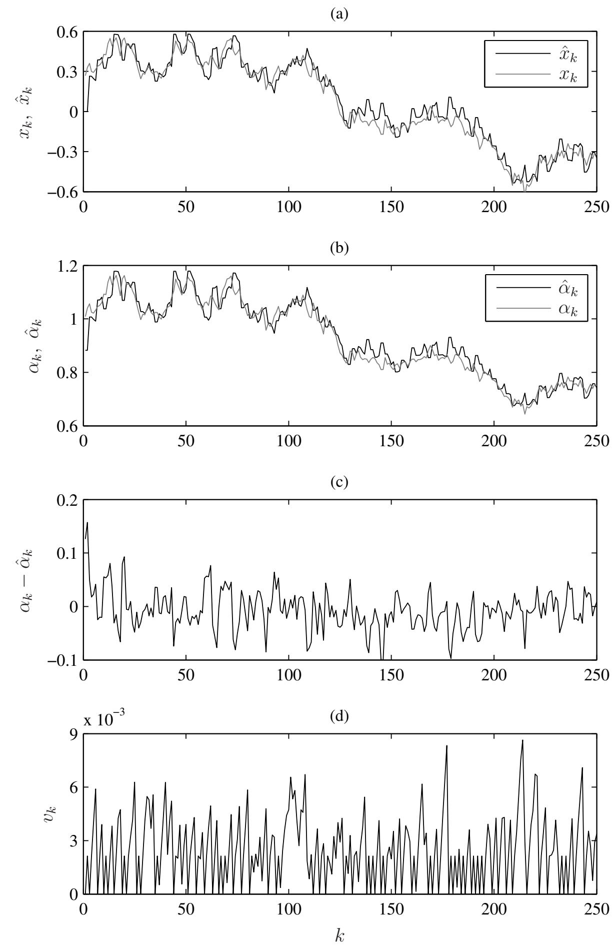

Fig. 2. Simulation results: (a) log-amplitude fade x_k (gray line) and its estimate $\hat{x}_k$ (black line); (b) optical fade α_k (gray line) and its estimate $\hat{\alpha}_k$ (black line); (c) estimation error $\alpha_k - \hat{\alpha}_k$; (d) estimation variance v_k.

and $\hat{\alpha}_k$ closely track their targets x_k and α_k. The estimation error $\alpha_k - \hat{\alpha}_k$ in part (c) is bounded between -0.1 and 0.1 (except for a short initial period), which shows an error of less than 10% in estimation of α_k.

VI. Conclusion

The problem of channel estimation for atmospheric optical channels with on-off keying modulation has been considered. The performance of these channels is degraded by the optical fade caused by atmospheric turbulence. The knowledge of optical fade can improve the channel performance when incorporated in adaptive techniques of signal detection, power control, and channel coding. A channel estimator has been developed to estimate the strength of optical fade from the observations of channel output. The approximate implementation of this estimator using a finite-dimensional nonlinear filter has been studied. The low update rate of this filter is its chief advantage over the prior results on this problem. An optimal scheme to incorporate the channel estimate in binary signal detection has been presented.

References

[1] M. C. Roggemann, B. M. Welsh, and R. Q. Fugate, "Improving the resolution of ground-based telescopes," *Rev. Mod. Phys.*, vol. 69, pp. 437 – 506, Apr. 1997.

[2] J. H. Shapiro, "Imaging and optical communication through atmospheric turbulence," in *Laser Beam Propagation in the Atmosphere* (J. W. Strohbehn, ed.), ch. 6, pp. 171–222, Berlin Germany: Springer-Verlag, 1978.

[3] E. V. Hoversten, R. O. Harger, and S. J. Halme, "Communication theory for the turbulent atmosphere," *Proc. IEEE*, vol. 58, pp. 1626 – 1650, Oct. 1970.

[4] A. Goldsmith, *Wireless Communications*. New York: Cambridge University Press, 2005.

[5] C. C. Davis, I. I. Smolyaninov, and S. D. Milner, "Flexible optical wireless links and networks," *IEEE Commun. Mag.*, vol. 41, pp. 51 – 57, Mar. 2003.

[6] J. Li and M. Uysal, "Achievable information rate for outdoor free space optical communication with intensity modulation and direct detection," in *Proc. of IEEE Global Telecommunications Conference (GLOBECOM)*, pp. 2654 – 2658, Dec. 2003.

[7] A. Komaee, "Detection and channel estimation for optical communication over atmospheric turbulent channels," in *Proc. of the Conference on Information Sciences and Systems (CISS)*, Mar. 2011.

[8] J. W. Goodman, *Statistical Optics*. New York: Wiley, 1985.

[9] S. Karp, R. M. Gagliardi, S. E. Moran, and L. B. Stotts, *Optical Channels: Fibers, Clouds, Water, and the Atmosphere*. New York: Plenum Press, 1988.

[10] X. Zhu and J. M. Kahn, "Markov chain model in maximum-likelihood sequence detection for free-space optical communication through atmospheric turbulence channels," *IEEE Trans. Commun.*, vol. 51, pp. 509 – 516, Mar. 2003.

[11] R. M. Gagliardi and S. Karp, *Optical Communication*. New York: Wiley, 1995.

[12] S. Karp, E. L. O'neill, and R. M. Gagliardi, "Communication theory for the free-space optical channel," *Proc. IEEE*, vol. 58, pp. 1611 – 1626, Oct. 1970.

[13] O. Macchi and B. C. Picinbono, "Estimation and detection of weak optical signals," *IEEE Trans. Inform. Theory*, vol. 18, pp. 562 – 573, Sep. 72.

[14] D. L. Snyder, *Random Point Processes*. New York: Wiley, 1975.

[15] A. N. Shiryaev, *Probability*. New York: Springer, 2nd ed., 1996.

[16] X. Zhu and J. M. Kahn, "Free-space optical communication through atmospheric turbulence channels," *IEEE Trans. Commun.*, vol. 50, pp. 1293 – 1300, Aug. 2002.

[17] H. V. Poor, *An Introduction to Signal Detection and Estimation*. New York: Springer-Verlag, 2nd ed., 1994.

[18] H. J. Kushner and A. S. Budhiraja, "A nonlinear filtering algorithm based on an approximation of the conditional distribution," *IEEE Trans. Automat. Contr.*, vol. 45, pp. 580 – 585, Mar. 2000.

[19] K. Ito and K. Xiong, "Gaussian filters for nonlinear filtering problems," *IEEE Trans. Automat. Contr.*, vol. 45, pp. 910 – 927, May 2000.

[20] D. Brigo, "On the nice behaviour of the Gaussian projection filter with small observation noise," *Systems & Control Letters*, vol. 26, pp. 363 – 370, 1995.

[21] A. Komaee, "State estimation from space-time point process observations with an application in optical beam tracking," in *Proc. of the Conference on Information Sciences and Systems (CISS)*, pp. 1 – 6, Mar. 2010.

[22] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.