AN OPTIMUM PATHWISE FILTER FOR OPTICAL HETERODYNE DPSK RECEIVER

T.E. Dabbous* and K.M. Hassan*

ABSTRACT
In this paper we propose a linear time invariant filter for optical heterodyne DPSK receiver to combat laser phase noise. Using optimal control techniques, we convert the filtering problem into an equivalent identification problem in which the filter parameters are the unknowns. Using variations arguments, we derive the necessary conditions of optimality on the basis of which the filter parameters can be determined. Simulation results show that the proposed filter is capable of minimizing the effects of phase and shot noise, that are superimposed on the DPSK signal, regardless the transmitted sequence.

KEY WORDS: Linear Filtering, Optimal Control Theory, Optical Communication.

1. INTRODUCTION

Coherent optical communication systems offer two basic advantages over direct detection optical systems: (1) improving in receiver sensitivity by about 10-20 dB, and (2) the possibility of using frequency division multiplexing to utilize more of the enormous bandwidth of a single mode fiber [1]. These systems suffer, however, from phase noise which is a major engineering problem that has to be solved before the use of coherent optical transmission. There are two ways that can be used for solving this problem. The first is the use of advanced fabrication techniques of high performance lasers with narrow linewidth. In fact, distributed feedback (DFB) and distributed Bragg reflector (DBR) lasers are the basic candidates for the narrow linewidth sources. Further, the projected linewidth of frequency less than 100 Khz requires more technological progress on large area wafers [2]. The second way is to design and implement an optical receiver that is capable of accommodating the inevitable phase noise, with the present technology, to achieve high performance reception. This is the main interest of the paper.

For Amplitude Shift Keying (ASK) and Frequency Shift Keying (FSK), the signal space constellations are asymmetric. Consequently, the effect of phase noise is not very severe. On the other hand, phase modulation possesses a great challenge due to the fact that the information and noise are both embedded into the phase. The simplest phase modulation format is binary shift keying (PSK) which gives the best quantum-limited receiver sensitivity (about 20 photons/bit for $1 \times 10^{-9}$ BER using heterodyne detection [3]). However, due to phase noise, PSK can be implemented only with phase locked loops due to the time varying nature of the phase noise process [4]. The alternative approach is to convey the information in the change of phase instead of its value. This is the differential phase shift keying (DPSK). In this case, the effective phase distortion is limited only to the current and previous bit intervals. Note that DPSK is less sensitive to phase noise than PSK.

In this paper we propose a novel linear time invariant filter model to detect the DPSK sequences in the presence of phase and shot noises. The filter model is assumed to be governed by a set of linear time invariant differential equations with unknown parameters. Using any of the optimization techniques available in the literature, one is able to obtain the (optimum) filter parameters with the help of which the effect of phase and shot noises can be reduced. It should be noted that the filter is driven by the DPSK sequences with phase and shot noise being superimposed on it and hence it is clear that the filter parameters depend on the input (or received) signal. In other words, the parameters of the proposed filter will be adapted to the received signal. This is known as pathwise filtering. In fact, this approach is different from that proposed in [5] where the authors used a fixed parameter for all received signals. Although the filter proposed in [5] with fixed parameter is easier for implementation, one may expect that the filter may work well for some signals while it may not give satisfactory results for other signals. That is why we have introduced the concept of pathwise filtering here to make use of the available information to improve the filter performance. It is worth mentioning that the filter proposed in this paper can be implemented by means of operational amplifiers and it can be also modified to suit for the detection of PSK signal which is proven to possess a better performance over the other modulation formats. The paper is organized as follows. In section 2 we formulate the (pathwise) filtering problem and present the filter equations. In section 3 we utilize optimal control theory to develop the necessary conditions for optimal identification of filter parameters. In section 4 we propose a numerical scheme, which is based on the necessary conditions, for computing the (optimum) filter parameters along with some numerical simulations to show the effectiveness of the proposed filter.
2. Problem Formulation and Filter Equations.

Let $x(t); -T \leq t \leq T$, denote the detected signal from the photo detector as shown in Fig. 1. Taking into consideration phase and shot noise, we assume that $x(t)$ is given by [5]:

$$x(t) = s(t) e^{j\theta(t)} + n(t), \quad -T < t < T,$$

(2.1)

where $s(t)$ represents the modulated data, $\theta(t)$ is the combined phase noise of transmitter and local oscillator lasers, and $n(t)$ is the shot noise introduced by the detection process. For each period $[0, T]$, the transmitted data $s(t)$ may take the values $+S$ or $-S$ for any $t \in [0, T]$, where $T$ is the system bit period. The phase noise $\theta$ is modeled as a Brownian motion with zero mean and variance $2\beta T; 0 \leq t \leq T$, where $\beta$ is the sum of linewidths of transmitter and local oscillator [5]. The shot noise $n(t)$ is modeled as a zero mean complex white Gaussian noise with variance $\gamma t$. The problem is to design a linear time invariant filter that reduces the effects of phase and shot noise and hence one is able to obtain a good estimate for the transmitted data $s(t); -T \leq t \leq T$. This filtering problem was considered in [5] where the authors used delay and multiply type heterodyne binary DPSK receiver, which includes two bandpass filters, to estimate the transmitted signal at the end of each bit period $T$. For the design of these filters the authors used the following criterion to determine (optimally) the filters parameters

$$J = E \left\{ \left( \frac{S^2}{2} \right)^2 \right\},$$

(2.2)

where

$$v(t) = \frac{1}{2} R \{ y(t)y^*(t - T) \}; 0 \leq t \leq T.$$

(2.3)

Here $R\{X\}$ denotes the real part of the complex variable $X$. $E\{v\}$ is the mathematical expectation of the random variable $v$, $y$ is the filter output, and $\frac{S^2}{2}$ is the filter output in the absence of noise. Minimizing $J$, the authors obtained the filter parameters with the help of which one is able to estimate $s(t)$ at the end of each period. Simulated results showed that the probability of error ($Pr\{v(T) < 0\}$) is increased as $\beta T$ increases. For the range $0.001 \leq \beta T \leq 0.1$, the probability of error ranges from $10^{-6}$ to $10^{-1}$ for SNR = 16 dB. Further, for each $\beta T$ and SNR, the authors obtained the filter parameter that minimizes the probability of error. Simulated results showed also that the optimal parameter increases with the increase of SNR and $\beta T$. 
In this paper we propose a linear time invariant filter that is governed by a set of ordinary differential equations, driven by the received signal, with unknown parameters. The filtering problem is then formulated as an identification problem in which the filter parameters are the unknowns. Using control theory, we develop the necessary conditions of optimality on the basis of which the filter parameters can be determined. Based on these necessary conditions, we propose a numerical scheme for computing the filter (optimal) parameters. In this section we present the filter model and show how the filtering problem can be treated as a pathwise identification problem in which the (optimal) parameters depend on the received signal \( x(t) \), which is known to the filter.

**Filter Equations**

Let \( x(t); -T < t < T \), denote the received signal, and let \( x_R(t) \) and \( x_I(t); -T < t < T \), denote the real and imaginary parts of the signal \( x(t) \), respectively. Let \( y(t) \) denote the filter output due to the input \( x(t); -T < t < T \), and let \( y_R \) and \( y_I \) be the corresponding real and imaginary parts of \( y \). Define

\[
\begin{align*}
x_1(t) &= x_R(t - T); \\
x_2(t) &= x_I(t - T); \\
x_3(t) &= x_R(t); \\
x_4(t) &= x_I(t); \\
y_1(t) &= y_R(t - T); \\
y_2(t) &= y_I(t - T); \\
y_3(t) &= y_R(t); \\
y_4(t) &= y_I(t); \quad 0 \leq t \leq T.
\end{align*}
\]

Using the above definitions, we assume that the proposed filter is governed by the following linear (stochastic) differential equations in which the real and imaginary parts of the received signal are the inputs

\[
\begin{align*}
\frac{dy_1}{dt} &= a_1 y_1(t) + a_2 x_1(t), \\
\frac{dy_2}{dt} &= a_1 y_2(t) + a_2 x_2(t), \\
\frac{dy_3}{dt} &= a_1 y_3(t) + a_2 x_3(t), \\
\frac{dy_4}{dt} &= a_1 y_4(t) + a_2 x_4(t), \quad 0 \leq t \leq T.
\end{align*}
\]
The filter parameters $a_1$ and $a_2$ are unknown and can be determined through the minimization process of the following performance index

$$J_1(a_1, a_2) = \int_0^T \left[ \lambda_1 \left( y_1(t)y_3(t) - S^2 \right)^2 + \lambda_2 \left( y_1^2(t) + y_4^2(t) \right) \right] dt,$$

where $\lambda_i (i > 0), 1 \leq i \leq 2$, are some arbitrary constant weights. The main reason for choosing the above criterion is to reduce the effect of both phase and shot noises on the received signal. This is done by taking into consideration the following requirements:

(a) The multiplication of the real parts of the filter outputs at $(t-T)$ and $t$ (i.e., $y_1$ and $y_3$) is close to $S^2$ (multiplication of the filter outputs at $(t-T)$ and $t$ in the absence of noise). This is included in the first term of the performance index $J_1$.

(b) The imaginary parts of the filter outputs at $(t-T)$ and $t$ are minimum. This is also included in the second term of $J_1$.

It should be noted that the criterion (2.6) assumes that, in the absence of noise, the transmitted data over the period $[-T, T]$ is "1" which corresponds to $(+S, +S)$ or $(-S, -S)$. In the case where the transmitted data is "0", which corresponds to $(-S, +S)$ or $(+S, -S)$, the above criterion is modified to

$$J_2(a_1, a_2) = \int_0^T \left[ \lambda_1 \left( y_1(t)y_3(t) + S^2 \right)^2 + \lambda_2 \left( y_1^2(t) + y_4^2(t) \right) \right] dt.$$

**Remark 2.1**

Since the transmitted data ("0"'s or "1"'s), during the period $[-T, T]$, is not known, one should use both of the above criteria to determine the filter optimum parameters. The minimum value of $J_1$ should be determined first assuming that the symbol "1" has been transmitted. Then assuming that symbol "0" has been transmitted, one determine the minimum value of $J_2$. If $J_1 < J_2$, then we choose the filter parameters that correspond to $J_1$, otherwise we choose the filter parameters that correspond to $J_2$.

In the remaining part of this section we shall reformulate the above filtering problem in a way that is convenient for the derivation of the necessary conditions of optimality on the basis of which the filter parameters can be determined.

Let $X = (x_1, x_2, x_3, x_4)'$ and $Y = (y_1, y_2, y_3, y_4)'$ and define

$$A = \begin{pmatrix} a_1 & 0 & 0 & 0 \\ 0 & a_1 & 0 & 0 \\ 0 & 0 & a_1 & 0 \\ 0 & 0 & 0 & a_1 \end{pmatrix}, \quad B = \begin{pmatrix} a_2 & 0 & 0 & 0 \\ 0 & a_2 & 0 & 0 \\ 0 & 0 & a_2 & 0 \\ 0 & 0 & 0 & a_2 \end{pmatrix}.$$
Then the filter equations (2.5) and the performance index $J_1$ can be written as follows

$$
\frac{dY}{dt} = A(\alpha)Y(t) + B(\alpha)X(t); \quad 0 \leq t \leq T.
$$

(2.8)

$$
Y(0) = Y_0,
$$

$$
J_1(\alpha) = \int_0^T \left\{ \lambda_1 \left( Y'(t)Q_1Y(t) - S^2 \right)^2 + \lambda_2 Y'(t)Q_2Y(t) \right\} dt,
$$

(2.9)

where $\alpha \equiv \{a_1, a_2 \}$, $Y_0$ is the initial state of the filter and $A'$ denotes the transpose of the matrix $A$. With this preparation, we can now state the filtering problem as follows

**Problem (P) (Filtering Problem)**

Given the observed path (data) $X(t); 0 \leq t \leq T$, find the parameter $\alpha \equiv \{a_1, a_2 \}$ so that $J_1(\alpha)$ is minimum subject to the dynamic constraint (2.8).

In the following section we make use of variational arguments to derive the necessary conditions of optimality for the problem (P) on the basis of which the filter parameters can be determined.

3. Necessary Conditions of Optimality

Consider the filtering problem (P) and let $\alpha^0$ be its solution (i.e., $J_1(\alpha^0) \leq J_1(\alpha)$ for all $\alpha \in \delta$, where $\delta$ denotes the parameter set which is assumed to be compact and convex). Let $\alpha^\varepsilon \equiv \alpha^0 + \varepsilon (\alpha - \alpha^0); \varepsilon \in [0,1]$, and $\alpha, \alpha^0 \in \delta$. Let $Y^\varepsilon(t) \equiv Y(t, \alpha^\varepsilon)$ and $Y^0(t) \equiv Y(t, \alpha^0); 0 \leq t \leq T$, be the solutions of (2.8) with $\alpha$ being replaced by $\alpha^0$ and $\alpha^\varepsilon$, respectively. Let

$$
\widetilde{Y}(t) \equiv \bar{Y}(t, \alpha^0, \alpha - \alpha^0) \equiv \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \left( Y^\varepsilon(t) - Y^0(t) \right),
$$

(3.1)

denote the (Gateaux) differential of $Y$ at $\alpha^0$ in the direction $(\alpha - \alpha^0)$. The following result shows that the differential $\widetilde{Y}$ exists and satisfies a related differential equation.
Lemma 3.1

Consider the problem (P) and suppose the parameter set \( \mathcal{A} \) is compact and convex. Then for each pair \( \alpha, \alpha^0 \in \mathcal{A} \), the (Gateaux) differential \( \widetilde{Y} \) of \( Y \) exists and satisfies the following differential equation

\[
\frac{d\widetilde{Y}}{dt} = \widetilde{A}Y^0(t) + A^0\widetilde{Y}(t) + \widetilde{B}X(t); \quad 0 \leq t \leq T,
\]

\[
\widetilde{Y}(0) = 0,
\]

where \( \widetilde{A} = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} (A(\alpha^\varepsilon) - A(\alpha^0)) \), denotes the (Gateaux) differential of \( A \).

Proof

The proof follows from standard computations (see for example [6,7,8]).

With the help of the above lemma, we can now present the necessary conditions of optimality for the filtering problem (P) which is the main result of this paper.

Theorem 3.2 (Necessary Conditions of Optimality)

Consider the problem (P) and suppose Lemma 3.1 hold. Then the optimal parameter \( \alpha^0 \) can be determined by the simultaneous solution of the differential equation

\[
\frac{dY^0}{dt} = A(\alpha^0)Y^0(t) + B(\alpha^0)X(t); \quad 0 \leq t \leq T,
\]

\[
Y^0(0) = Y_0.
\]

The adjoint equation

\[
- \frac{d\psi}{dt} = A'(\alpha^0)\psi(t) + \lambda_1 ((Y^0(t))'Q_1Y^0(t) - 2S^2)Q_1 + Q'_1 Y^0(t) + \lambda_2 [Q_2 + Q'_2] Y^0(t); \quad 0 \leq t \leq T,
\]

\[
\psi(T) = 0,
\]

and the inequality

\[
\int_0^T (\psi'(t)[\widetilde{A}Y^0(t) + \widetilde{B}X(t)]) dt \geq 0.
\]
Proof

Define

\[ J(\alpha^e) = \int_0^T \left\{ \lambda_1 ((Y^e(t))^\prime Q_1 Y^e(t) - S^2) + \lambda_2 (Y^e(t))^\prime Q_2 Y^e(t) \right\} dt, \quad (3.6) \]

\[ J(\alpha^o) = \int_0^T \left\{ \lambda_1 ((Y^o(t))^\prime Q_1 Y^o(t) - S^2) + \lambda_2 (Y^o(t))^\prime Q_2 Y^o(t) \right\} dt, \quad (3.7) \]

and

\[ \mathcal{J}_0(\alpha - \alpha^0) = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon}(J(\alpha^e) - J(\alpha^o)), \quad (3.8) \]

where \( \mathcal{J}_0(\alpha - \alpha^0) \) denotes the (Gateaux) differential of \( J \) at \( \alpha^0 \) in the direction \( \alpha - \alpha^0 \). In order that \( J \) attains its minimum at \( \alpha^0 \), it is necessary that

\[ \mathcal{J}_0(\alpha - \alpha^0) = \int_0^T \left\{ \lambda_1 \left[ (Y^o(t))^\prime Q_1 Y^o(t) - 2S^2 \right] (Q_1 + Q_1^\prime) Y^o(t) 
+ \lambda_2 (Q_2 + Q_2^\prime) Y^o(t) \right\} dt \geq 0. \quad (3.9) \]

where \( \tilde{Y} \) is the solution of (3.2). The inequality (3.9) can be further simplified by introducing the adjoint variable \( \psi(t) \), which satisfies the (backward) equation (3.4). Using equations (3.2), (3.4) and (3.9) and noting that

\[ \int_0^T \frac{d}{dt} (\tilde{Y}^\prime(t)\psi(t)) dt = 0, \]

one can easily verify that

\[ \mathcal{J}_0(\alpha - \alpha^0) = \int_0^T \psi^\prime(t) \left( \tilde{A} Y^o(t) + \tilde{B} X(t) \right) dt. \quad (3.10) \]

The inequality (3.5) now follows from (3.9) and (3.10). This completes the proof.

4. Algorithm and Numerical Simulations

In this section we make use of the necessary conditions developed in section 3 to propose a numerical scheme for computing the filter optimum parameters \( a_1 \) and \( a_2 \). The proposed filter and the corresponding algorithm are then tested by a numerical examples to illustrate the effectiveness of the filter.
Algorithm
The proposed algorithm may be summarized in the following steps:

1. Set $n=1$, and guess the filter parameters $\left\{ a_1^n, a_2^n \right\}$.
2. Using Gaussian random generator, generate the phase noise $\theta(t)$, the shot noise $n(t)$ and the process $x(t)$; $-T \leq t \leq T$.
3. Given the data $x(t)$; $-T \leq t \leq T$, find $x_i(t)$; $1 \leq i \leq 4$, (see equation (2.4)).
4. Solve the differential equation (3,3) and obtain $Y(n)^{(n)}(t); 0 \leq t \leq T$.
5. Given $Y(n)^{(n)}(t); 0 \leq t \leq T$, obtain $\psi(n)^{(n)}(t); 0 \leq t \leq T$, by solving (3.4).
6. Using the inequality (3.5), obtain the gradient vector $g(n) = \left( g_1(n), g_2(n) \right)$.
7. Update the filter parameters using the following relations:
   
   \[ a_1^{(n+1)} = a_1^{(n)} + \varepsilon g_1^{(n)} \]
   \[ a_2^{(n+1)} = a_2^{(n)} + \varepsilon g_2^{(n)} \]

   where $\varepsilon (> 0)$ is chosen so that $J_1(a_1^{(n+1)}, a_2^{(n+1)}) < J_1(a_1^{(n)}, a_2^{(n)})$.

8. If $\left| J_1(a_1^{(n+1)}, a_1^{(n+1)}) - J_1(a_1^{(n)}, a_2^{(n)}) \right| \leq \delta$, where $\delta (> 0)$ is sufficiently small, then stop; otherwise set $n=n+1$ and $a_i^{(n)} = a_i^{(n+1)}$, $1 \leq i \leq 2$, and go to step 4.

Numerical Simulations
Let $T=2$, and $S=2$, as normalized values, and let $\Delta t (= 0.01)$ denote the integration step size. Using random number generator, we have generated the received data $x(t); -2 \leq t \leq 2$, for different variances $\beta \Delta t$ and $\gamma \Delta t$. (see figs. 2, 5, 8, 11, 14, 17, 20). In the figs. 2, 5, 8, 11, we have taken $\gamma \Delta t=5$, whereas $\beta \Delta t = 5, 10, 15, 20$, receptively, and transmitted data is assumed $s(t) = 2; -T \leq t \leq 0$, and $s(t) = -2; 0 \leq t \leq T$, (i.e., symbol “0” being transmitted). The real parts of the filter output corresponding to the values of $\beta \Delta t$ given above are depicted in figs. 6, 9, and 12. Note that the real and imaginary parts of the filter output, denoted respectively by $y_R(t)$ and $y_I(t)$, are given by

\[ y_R(t) = y_1(t) + y_3(t), \]
\[ y_I(t) = y_2(t) + y_4(t); -2 \leq t \leq 2. \]
From these figures, it is clear that the filter output $y_R$ has a behaviour which is close to that of the DPSK transmitted signal and hence one can easily tell that symbol "0" has been transmitted. Another way for detecting the transmitted data is shown in figs. 4, 7, 10, and 13. In these figures we have used the variable $v(t)$ which is given by

$$v(t) = y_R(t)y_R(t - T) - y_1(t)y_1(t - T); \quad 0 \leq t \leq 2,$$

as a criterion for detecting the transmitted signal. If $v(T) < 0$, then symbol "0" has been transmitted otherwise symbol "1" has been transmitted. As shown in figs. 4, 7, 10, and 13, the value of $v(2) \cong -4 = -S^2$ which indicate that symbol "0" has been transmitted. In figs. 14-16, we have taken $\beta \Delta t = 5$ and $\gamma \Delta t = 20$ to study the effect of shot noise. Again, it is clear that symbol "0" has been transmitted. From figs. 17-22 we observe, once again, that symbol "1" has been transmitted. From these results, it is clear that our proposed filter is capable of detecting the transmitted data and the effect of phase and shot noise has been remarkably reduced.

5. Conclusions

In this paper we have considered the filtering problem for DPSK optical signal. A linear time invariant filter has been proposed to reduce the effect of phase and shot noise that are superimposed on the DPSK signal. The filter consists of four differential equations in which the received signal being its input. In other words, the filter is driven by the input signal and hence the filter parameters are adapted to the received signal. Using control theory approach, we have formulated the filtering problem as an equivalent identification problem in which the filter parameters are the unknowns. Using a suitable performance index (i.e., $J_1$ or $J_2$), we have developed the necessary conditions for optimal identification. Based on these necessary conditions, we have proposed an iterative scheme for determining the filter optimum parameters along with some numerical simulations to illustrate the effectiveness of the proposed filter. From these simulations, we have seen that the filter is capable of reducing the effect of phase and shot noises. In fact, we have seen that the filter output $y_R(t)$ has a behavior which is very close to that of the transmitted signal $s(t)$, and hence one can easily decide for which symbol has been transmitted (see figs. 3, 6, 9, 12, 15, 18 and 21). Further, we have introduced the variable $v(t)$ which can also be used to decide for the symbol being transmitted. In figs. 4, 7, 10, 13, and 16, $v(T)$ is close to -4.0 indicating that symbol "0" has been transmitted. On the other hand, in figs. 19 and 22, $v(T)$ is close to 4.0 indicating that symbol "1" has been transmitted. Finally, it is worth mentioning here that the proposed filter can be implemented using operational amplifiers and also can be easily modified to suit for the detection of PSK signal which is proven to have a better performance over the other modulation formats.
6. References


Fig. 1 Optical Heterodyne Binary DPSK Receiver
\[ \beta \Delta t = 5 \& \gamma \Delta t = 5 \]

\[ \beta \Delta t = 10 \& \gamma \Delta t = 5 \]
$\beta \Delta t = 15 \& \gamma \Delta t = 5$

![Graphs showing signal variations over time for different values of $\beta \Delta t$ and $\gamma \Delta t$.]

$\beta \Delta t = 20 \& \gamma \Delta t = 5$

![Graphs showing signal variations over time for different values of $\beta \Delta t$ and $\gamma \Delta t$.]
$\beta \Delta t = 5 \& \gamma \Delta t = 20$

$\beta \Delta t = 15 \& \gamma \Delta t = 5$
\[ \beta \Delta t = 10 \& \gamma \Delta t = 20 \]

Fig. 20

Fig. 21

Fig. 22