PHASE AND FREQUENCY TRACKING CONSIDERATIONS FOR HETERODYNE OPTICAL COMMUNICATIONS

John E. Kaufmann*
Massachusetts Institute of Technology
Lincoln Laboratory
Lexington, Massachusetts 02173

ABSTRACT

In heterodyne optical communications, phase or frequency tracking is generally needed to avoid performance degradation when signaling in the presence of laser frequency jitter and Doppler shifts. This paper examines a phase-lock loop approach for BPSK and two forms of frequency tracking for MFSK. Using a statistical model for laser frequency instability, the performance of these schemes is calculated by a linear analysis of the tracking loop in the small-error regime.

I. INTRODUCTION

Recently there has been considerable interest in heterodyne optical communications systems because of the potential for substantial performance improvement over direct detection systems. The expected improvement is at least 10 dB around 1 μm (Ref. 1). In heterodyne systems, however, the laser's frequency instability can impact the choice of modulation schemes, the required laser transmitter power, and the receiver implementation. In general, frequency or phase tracking will be needed at the receiver to avoid significant degradations in communications performance and the need for increased transmitter power unless very stable lasers are available.

This paper examines receiver phase and frequency tracking schemes suitable for heterodyne PSK and MFSK systems in a space communications context, although this work is also applicable to fiberoptic systems. Tracking information is derived directly from the received modulated signal--no separate beacon signal is provided for this purpose. A phase-lock loop approach is pursued for PSK signaling and two methods of frequency tracking--a decision-directed and a suboptimal frequency tracker--are investigated for MFSK. Linearization of the tracking system under the assumption that the tracking errors

* This work was sponsored by the Department of the Air Force.
The U.S. Government assumes no responsibility for the information presented.

are small when the loop is locked permits the derivation of analytic results for tracking performance.

Knowledge of the power spectral density of the frequency noise process is essential for an optimized design. Measurements of typical frequency-jitter spectra are reported in Ref. 2 for a HeNe laser and in Ref. 3 and Ref. 4 for GaAlAs. The spectral shape takes the form shown in Figure 1 with a low-frequency rise which my be attributed in part to "environmental" effects such as temperature and current fluctuations or mechanical vibrations, and at higher frequencies, a low-level spectrally "white" component associated with quantum fluctuations in the laser. Generally, it is the low-frequency fluctuations which will be tracked, depending upon the receiver signal-to-noise ratio. The untracked high-frequency components will degrade communications performance, however.

Finally, considerations of tracking Doppler shifts, which can be very large in satellite communications, are discussed.

II. PHASE SYNCHRONIZATION FOR PSK SIGNALING

For binary signaling, antipodal PSK is the optimum signaling scheme and a phase-lock loop (PLL) is used for phase-tracking. The input to the PLL is the intermediate-frequency (IF) signal at the output of the receiver photodetector. The detector's output noise process n(t) can be modeled as white Gaussian with two-sided power spectral density

$\frac{N_o}{2} = \frac{h\nu}{2\eta}$ where h is Planck's constant, ν is the frequency of the optical field, and η is the quantum efficiency of the detector.

The received IF signal then takes the form

$$r(t) = \sqrt{2P}\, m(t) \cos [2\pi f_{IF} t + \phi(t)] + n(t) \qquad (1)$$

where P is the received signal power, m(t) is the binary modulation (±1), f_{IF} is the receiver's IF frequency and $\phi(t)$ is the phase jitter in the signal due to laser instability. Note that $\phi(t)$ is the difference between the phases of the PLL oscillator and the received signal. Figure 2 depicts the receiver with a squaring loop appropriate for BPSK. The bandwidth W_i of the IF bandpass filter is chosen to be just wide enough to pass the IF signal without appreciable distortion and to reject out-of-band noise. If the loop bandwidth, determined by the filter F(s), is much less than W_i (which is the case for sufficiently high signaling rates), the loop's internal noise can be modeled as being approximately white with a two-sided power spectral density $\frac{N_o'}{2}$,

$$\frac{N_o'}{2} = \frac{P\,h\nu}{\eta} + \frac{1}{2}\left(\frac{h\nu}{\eta}\right)^2 W_i \quad . \tag{2}$$

The effective signal power P' in the loop, after squaring, is $P' = (1/2)P^2$

In the locked condition where the tracking error is small (much less than one radian), the PLL may then be modeled as a linear system with input

$$r(t) = 2\,\phi(t) + n'(t) \tag{3}$$

where n'(t) is a white noise process with spectral height $N_o'/(2P')$ (Note that a double phase term arises in (3) because of the squaring operation). The approach taken here is to view the role of the linearized PLL as a filter which produces an estimate $\hat{\phi}(t)$ of $\phi(t)$. Using the mean-square error $E\{[\phi(t) - \hat{\phi}(t)]^2\}$ as the performance criterion, Wiener-filter theory may be applied to find the transfer function of the optimum PLL and its mean-square error performance. In this analysis, the frequency-noise power spectral density is modeled by

$$S_{\dot{\phi}}(\omega) = \frac{2\,\tau_c^{-1}\,d_f^2}{\omega^2 + \tau_c^{-2}} \tag{4}$$

where d_f is the rms frequency deviation (rad/sec) which accounts for the long-term spectral width of the heterodyne signal, and τ_c is the coherence time of the instantaneous frequency. The transfer function of the optimum PLL is found to be

$$H_o(s) = \frac{s(a-\tau_c^{-1}) + b}{s^2 + sa + b}$$

where

$$a = \left(\tau_c^{-2} + 4d_f\sqrt{\frac{2}{\gamma\tau_c}}\right)^{1/2} \tag{5}$$

$$b = 2d_f\sqrt{\frac{2}{\gamma\tau_c}}$$

$$\gamma = N_o'/2P' = \frac{2h\nu}{\eta P}\left(1 + \frac{1}{2}\,\frac{h\nu}{\eta P}\,W_i\right)$$

$H_o(s)$ specifies a second-order loop with natural frequency ω_n and damping factor ζ given by

$$\omega_n = \left(\frac{8d_f^2}{\gamma}\right)^{1/4}$$
$$\zeta = \frac{a}{2\omega_n} \tag{6}$$

The resultant minimum mean-square error ξ_ϕ^2 is

$$\xi_\phi^2 = \frac{\gamma}{4\tau_c}\left\{-1 + \left(1 + 4d_f\tau_c\sqrt{\frac{2\tau_c}{\gamma}}\right)^{1/2}\right\} \tag{7}$$

When an rms error requirement is imposed on the PLL for a given d_f and τ_c, equation (7) may be used to solve for the value of $P/h\nu$ needed. A simplified approximate form of the solution, valid for $P/h\nu >> 1$ is

$$\left(\frac{P}{h\nu}\right)_{req} = \eta^{-1}\left(\frac{d_f^2}{\xi_\phi^8\tau_c}\right)^{1/3} \tag{8}$$

Figure 3 plots the $P/h\nu$ requirenent as a function of d_f with τ_c as a parameter for $\eta = 1$ in the case where $\xi_\phi = .1$ radian. For a GaAlAs laser (in a stable temperature environment), typical parameter values might be $d_f/2\pi = 10$ Miz and $\tau_c = 10$ seconds.

III. FREQUENCY TRACKING FOR MFSK SIGNALING

A. Decision-Directed Tracker

Incoherent MFSK signaling may be an attractive alternative to PSK for heterodyne optical communications since an absolute phase reference is not required at the receiver and frequency modulation of some lasers, such as GaAlAs, is more easily accomplished than phase modulation. Furthernure, the use of a large MFSK symbol alphabet size, feasible with large available bandwidth, can yield performance superior to BPSK. (In fact, as $M \rightarrow \infty$, MFSK is asymptotically the optimum signaling scheme). However, some form of frequency tracking is necessary.

The tracking of frequency is based on the observation of a sequence of pulse waveforms

$$r_i(t) = \sqrt{2P}\cos[2\pi(f_{IF} + f^{(i)})t + \int_{-\infty}^{t}\dot{\phi}(s)ds + \theta_i] + n(t) \qquad (9)$$

$$(i-1)T_s \leq t \leq iT_s$$

where $f^{(i)}$ is one of M tone frequencies f_j, j = 1,2,...,M, n(t) is white Gaussian noise with two-sided spectral density $\frac{h\nu}{2}$, T_s is the duration of one symbol, and θ_i is a random phase uniformly distributed over $(0, 2\pi)$. The role of the frequency-tracker is to estimate the frequency jitter $\dot{\phi}(t)$.

The tracking problem is complicated by the a priori uncertainty in $f^{(i)}$. Therefore, a decision-directed approach, whereby knowledge of $f^{(i)}$ is obtained by waiting until demodulation of the i^{th} symbol, is pursued. For a receiver symbol-error probability P(E), the $f^{(i)}$ information supplied to the tracker is correct with probability 1-P(E), which is very nearly equal to one for small P(E). Because of the delay in obtaining the $f^{(i)}$ data to make the frequency estimate, the change in $\dot{\phi}(t)$ must be small from one symbol to the next for this scheme, to work.

The form of the frequency tracker is shown in Figure 4. Every T_s seconds an estimate is made of the tracking error and the frequency reference for the demodulator is updated. Perfect symbol tine-synchronization is assumed. For purposes of analysis, the frequency error can be assumed to remain essentially constant over the duration of a symbol. The problem of estimating the tracking error then becomes me of estimating a constant frequency offset. A bank of square-law detectors is employed–one pair for each tone frequency. If it was decided that f_k was the transmitted tone, data from only the detector pair corresponding to fk is used to form the error estimate.

If ρ (Hz) is the difference in frequency between the local reference and the received signal, the objective of the loop frequency-error estimator, shown in Figure 5, is to estimate ρ. If f_k is sent, the estimate $\hat{\rho}$ takes the form

$$\hat{\rho} = \frac{\pi^2}{16E_sT_s^2}(r_1 - r_2) \qquad (10)$$

where r_1 and r_2 are the squared magnitudes of the signal spectral components λ/T_s Hz ($\lambda \leq 1$) above and below f_k respectively,

$$r_1 = \left| \int_0^{T_s} r(t)\, e^{-j2\pi(f_k + \frac{\lambda}{T_s})t}\, dt \right|^2$$

$$r_2 = \left| \int_0^{T_s} r(t)\, e^{-j2\pi(f_k - \frac{\lambda}{T_s})t}\, dt \right|^2 \qquad (11)$$

and $E_s = PT_s$ is the symbol energy. The performance of this estimator is discussed in detail in Ref. 5. The multiplicative factor in front of (r_1-r_2) in (10) provides an unbiased estimate of ρ as $\rho \to 0$, and requires estimation of the symbol energy separately. $\lambda = 0.5$ is selected here on the basis of minimizing the bias and variance of $\hat{\rho}$. The bias and variance depend upon the actual value of ρ. Around $\rho = 0$, it is shown in Ref. 5 that $\hat{\rho}$ is very nearly unbiased with variance

$$\mathrm{Var}(\hat{\rho} \mid \rho = 0) = \left(\frac{\pi}{8}\right)^2 \left[\frac{h\nu}{\eta E_s T_s^2} \quad 1 + \frac{1}{2} \left(\frac{\pi}{2}\right)^2 \quad \frac{h\nu}{\eta E_s} \right] \triangleq \sigma_n^2 \qquad (12)$$

For $\eta E_s / h\nu >> 1$, (12) is very nearly equal to the Cramér-Rao bound for frequency estimation for a constant-amplitude T_s -second pulse of energy E_s (Ref. 6),

$$\sigma_{CR}^2 = \frac{6}{(2\pi)^2 T_s^2} \quad \frac{h\nu}{\eta E_s} \quad . \qquad (13)$$

The sequence of frequency estimates is then filtered by a discrete-time filter $F(z^{-1})$ whose output drives a VCO to correct the tracking error. The closed-loop impulse response h_k is chosen to take the form

$$h_k = 2b\, e^{-kb} \cos(kb) \qquad (14)$$

which is the discrete-time counterpart of a second-order analog loop with damping factor $1/\sqrt{2}$. The impulse response is specified completely by the parameter b so that finding the optimum loop is accomplished by optimizing with respect to b. The frequency noise power spectral density is assumed to be the same as (4). When the frequency drift in T_s seconds is much less than T_s^{-1}, the b which minimizes the mean-square loop error

$\xi_{\dot{\phi}}^2 = E\{[\dot{\phi}(t) - \hat{\dot{\phi}}(t)]^2\}$ is found to be

$$b = \frac{d_f}{2\pi\, \sigma_n} \sqrt{\frac{T_s}{3\tau_c}} \qquad (15)$$

and a simplified, approximate expression for the resultant minimum mean-square error, useful for first-cut design purposes, is

$$\xi_{\dot{\phi}}^2 = \frac{d_f \sigma_n}{2\pi} \sqrt{\frac{3T_s}{\tau_c}} \qquad (Hz^2) \tag{16}$$

Fixing an rms requirement $\xi_{\dot{\phi}}$ on loop performance allows determination of the required P/hν as a function of the various parameters,

$$\left(\frac{P}{h\nu}\right)_{req} = -22.3 + 10 \log\{Z[1 + \sqrt{1 + 421.7\, Z^{-1} T_s^{-1}}]\} \text{ dB-Hz} \tag{17}$$

$$\text{where } Z = \frac{d_f^2 T_s^{-2}}{\tau_c \xi_{\dot{\phi}}^4}$$

The result in (17) is valid for $b << 1$, which is the case for slow frequency drifts and corresponds to the optimum performance when (15) is such that $b >> T_s / \tau_c$.

$\left(\frac{P}{h\nu}\right)_{req}$ is plotted in Figure 6 for two values of T_s as a function of d_f with τ_c as a parameter.

B. A Sub-Optimum Frequency Tracker

A possible drawback to the decision-directed tracking approach is its complexity which increases with the symbol alphabet size M. A suboptinum tracking scheme which does not require symbol demodulation or time synchronization is presented here.

Figure 7 depicts the form of this tracking loop. The idea here is to derive the tracking information from only the MFSK tones f_1 and f_M instead of all the tones as in the decision-directed scheme. The system uses upper and lower band-edge bandpass filters whose outputs are passed through square-law detectors and a threshold decision device to provide an error signal for tracking. The tracking error is estimated by forming the difference between the detector outputs. The loop attempts to reduce the error by centering the overall spectrum between the band-edge filters.

Ideally, the band-edge filters produce a signal output only when the lowest or highest frequency tones are present and noise output the remainder of the time. When M is large, noise is observed most of the time. To prevent noise-induced tracking degradations when no signal is present in the filters, a threshold detector is provided. The detector allows its input to pass through to the loop only when it decides a signal is present on the basis of a threshold test. Otherwise, no signal is passed on. A small change in frequency error over the time it takes to accumulate data to form a frequency-error estimate is assumed. (Since an estimate can be formed only after reception of f_1 and f_M, the tracking error can be corrected, on average, only once every M symbols, or every MT_s seconds).

The closed-loop impulse response is assumed to be the same as that specified by (14). Optimization of the loop then proceeds in a manner very similar to that for the decision-directed loop. The minimum attainable mean-square tracking error is found to be

$$\xi_{\phi}^2 = \frac{d_f \sigma_n'}{2\pi} \sqrt{\frac{3MT_s}{\tau_c}} \quad (\text{Hz}^2) \tag{18}$$

σ_n' is given by

$$\sigma_n'^2 = \left(\frac{\pi}{8}\right)^2 \frac{h\nu}{\eta E_s P_D} \left\{1 + \frac{1}{2}\left(\frac{\pi}{2}\right)^2 \frac{h\nu}{\eta E_s}\left[1 + \frac{P_F}{P_D}(M-1)\right]\right\} \tag{19}$$

where P_D is the signal-detection probability for the threshold detector for $\rho = 0$ and P_F is the false-alann probability. For $P_D = 1$ and $P_F = 0$, the result in (19) becomes the same as (12) for the decision-directed scheme, that is, $\sigma_n' = \sigma_n$. A comparison of (18) with (16) points out the basic difference between the two frequency trackers. All parameters being equal for both systems, the mean-square error is larger by the factor $\sqrt{M}$ with the band-edge method. To provide the same performance as the decision-directed loop, the band-edge tracker requires M times more received signal power when $\eta E_s/h\nu >> 1$ and $\sqrt{M}$ times more power if $\eta E_s/h\nu << 1$ for the same frequency jitter. From another point of view, for a given received-signal power, the band-edge tracker sacrifices the ability to track larger rates-of-change in frequency: to maintain the same rms error, τ_c must be increased by the factor M or d_f reduced by $\sqrt{M}$.

IV. DOPPLER-SHIFT CONSIDERATIONS

When there is relative motion between platform, optical space communications systems experience much larger Doppler shifts than microwave systems. Tracking the Doppler shifts as well as the laser jitter is essential.

Consider an optical link between a synchronous and a low-altitude satellite. The largest shifts occur in the case where the synchronous satellite and the orbit of the low-altitude satellite are coplanar, for which it is straightforward to show that the Doppler frequency is

$$f_d(t) = \frac{f_o \omega R_s R_L \cos \omega t}{c(R_s^2 + R_L^2 - 2R_L R_s \sin \omega t)^{1/2}} \quad (\text{Hz}) \tag{20}$$

where ω is the angular frequency of the low-altitude satellite orbit with respect to the earth, R_S and R_L are the synchronous and low-orbit radii, respectively, f_o is the optical carrier frequency (Hz), and c is the speed of light. The time origin t = 0 is chosen such that $t = \frac{\pi}{2\omega}$ gives the time of closest approach of the two satellites. If one assures a 90-minute period for the low-orbit satellite and an optical wavelength of .83 μm (typical of

GaAlAs lasers), the maximum Doppler shift is in excess of 9 GHz with a maximum rate-of-change of nearly 13 MHz/second in magnitude. The tunability of the frequency control system of the heterodyne receiver mist be sufficient to accommodate these shifts.

While laser frequency/phase jitter my be tracked by a loop which is closed through a VCO at the IF, the range of Doppler shifts to be tracked will generally require a Doppler tracking loop, closed through the local oscillator as shown in Figure 8. Tracking Doppler is much simpler than tracking the laser frequency jitter, however, since the Doppler shifts change smoothly and, in fact, predictably over time whereas laser jitter does not. Therefore, a Doppler-tracking design is not given here, but rather it is noted that one must be incorporated into the final overall system design.

V. CONCLUSION

The performance of a PLL for PSK communications and two different frequency-tracking loops for MFSK have been analyzed in the regime of small tracking errors. Knowledge of the frequency-noise power spectral density is needed to optimize loop performance. Further analysis of phase and frequency tracking to determine performance for different frequency-noise spectra, to quantify other aspects of loop behavior such as pull-in performance, and loss of lock or cycle-slipping is warranted. Examination of typical laser parameters indicates that phase and frequency tracking are important considerations in a sound communications system design.

REFERENCES

1. Chan, V., “Coherent Optical Space Communication System Architecture and Technology Issues,” Proceedings on Control and Communication Technology in Laser Systems, SPIE Meeting, San Diego, California, August 25-26,-1981.
2. Siegman, A. E., et al., “Preliminary Measurements of Laser Short-Term Frequency Fluctuations,” IEEE Journal of Quantum Electronics, Vol. QE-3, No. 5, May 1967, pp. 180-189.
3. Tsuchida, H., et al., “Frequency Stability Measurement of Feedback Stabilized AlGaAs DH Laser,” Japanese Journal of Applied Physics, Vol. 19, No. 12, December 1980, pp. L721-L724. -
4. Walther, F. G., M.I.T. Lincoln Laboratory, work to be published.
5. Mui, S. Y., “A Weighted-Average Frequency Tracker,” National Telecommunications Conference, Houston, Texas, Nov. 30-Dec. 4, 1980.
6. Skolnik, M. I., Radar Handbook, McGraw-Hill Book Co., New York, 1970, pp. 4-4 - 4-7.

FIGURE 1

POWER SPECTRAL DENSITY OF LASER FREQUENCY NOISE

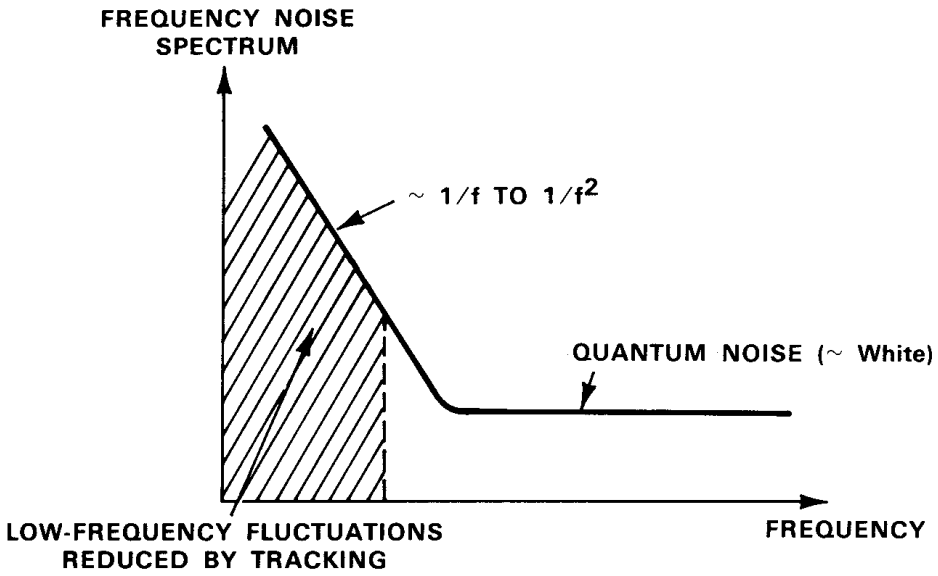

FIGURE 2

OPTICAL HETERODYNE BPSK RECEIVER WITH PHASE-LOCK LOOP

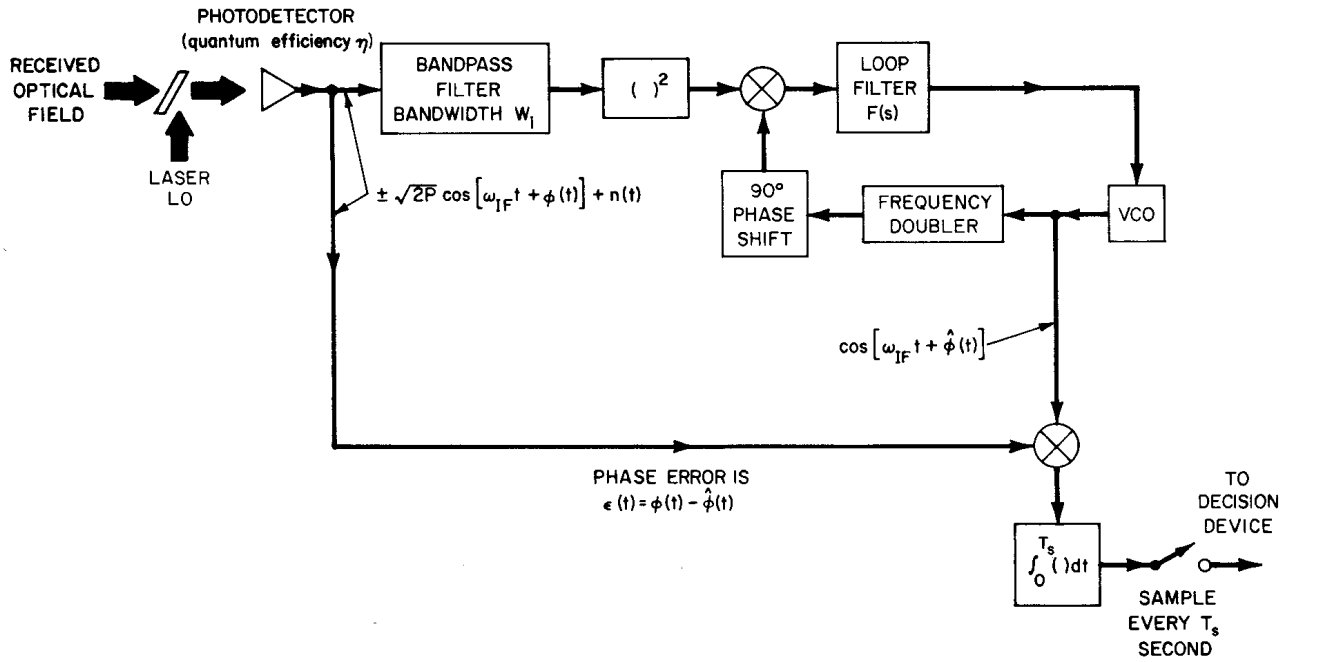

FIGURE 3

RECEIVED POWER REQUIREMENT FOR PHASE-TRACKING
(0.1-rad RMS Error, η = 1)

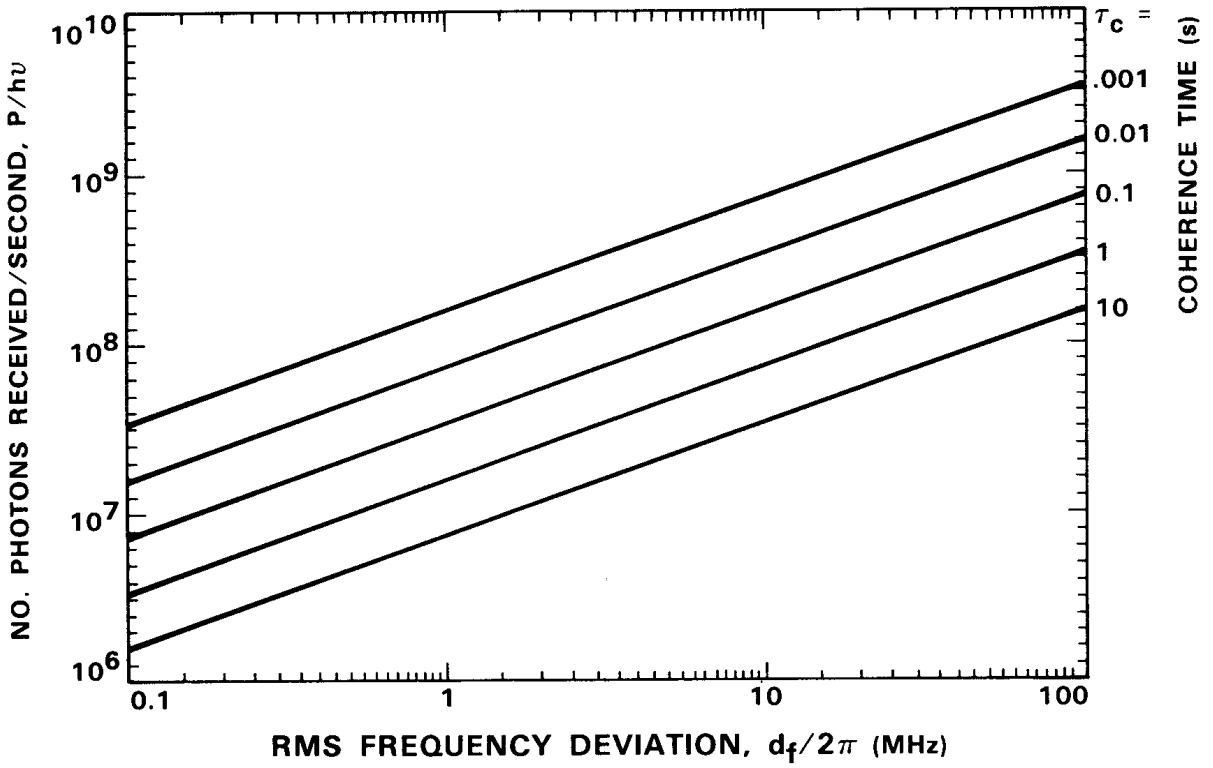

FIGURE 4

DECISION-DIRECTED FREQUENCY-TRACKING LOOP FOR MFSK

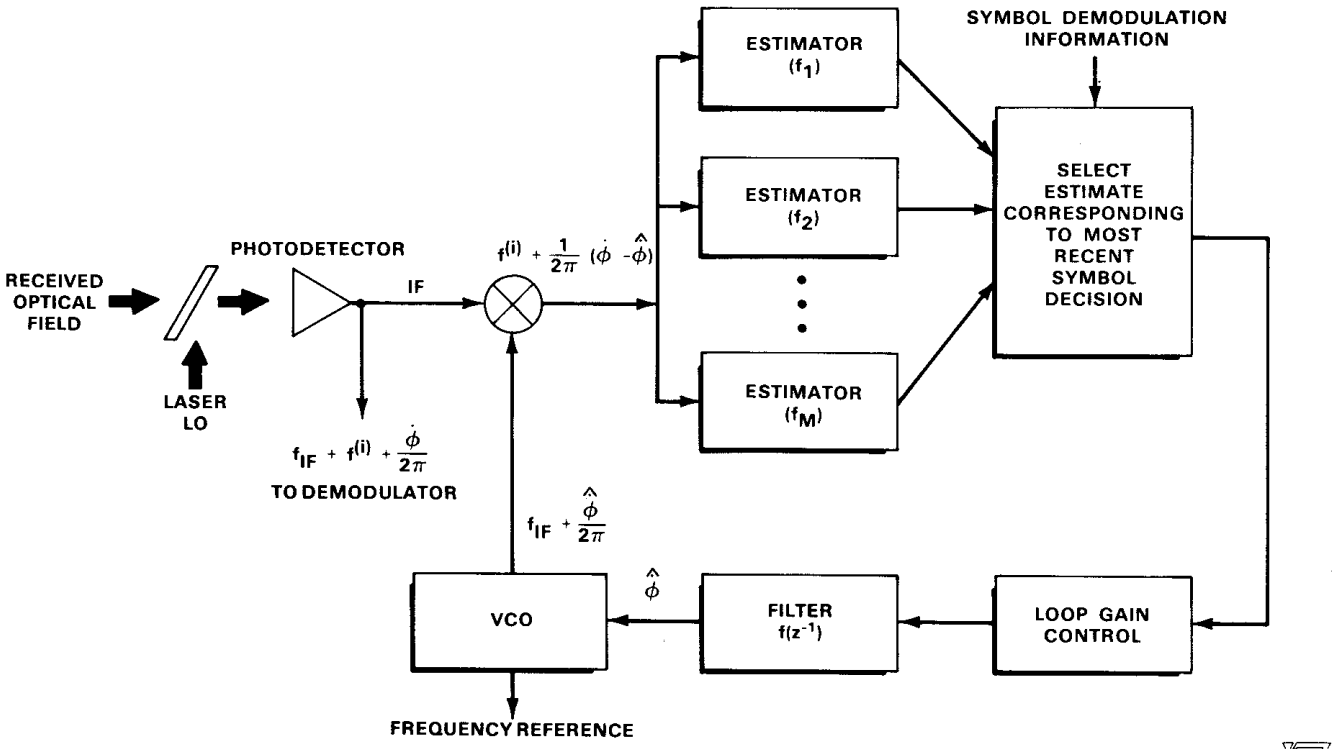

FIGURE 5

FREQUENCY-ERROR ESTIMATOR FOR TRACKING LOOP

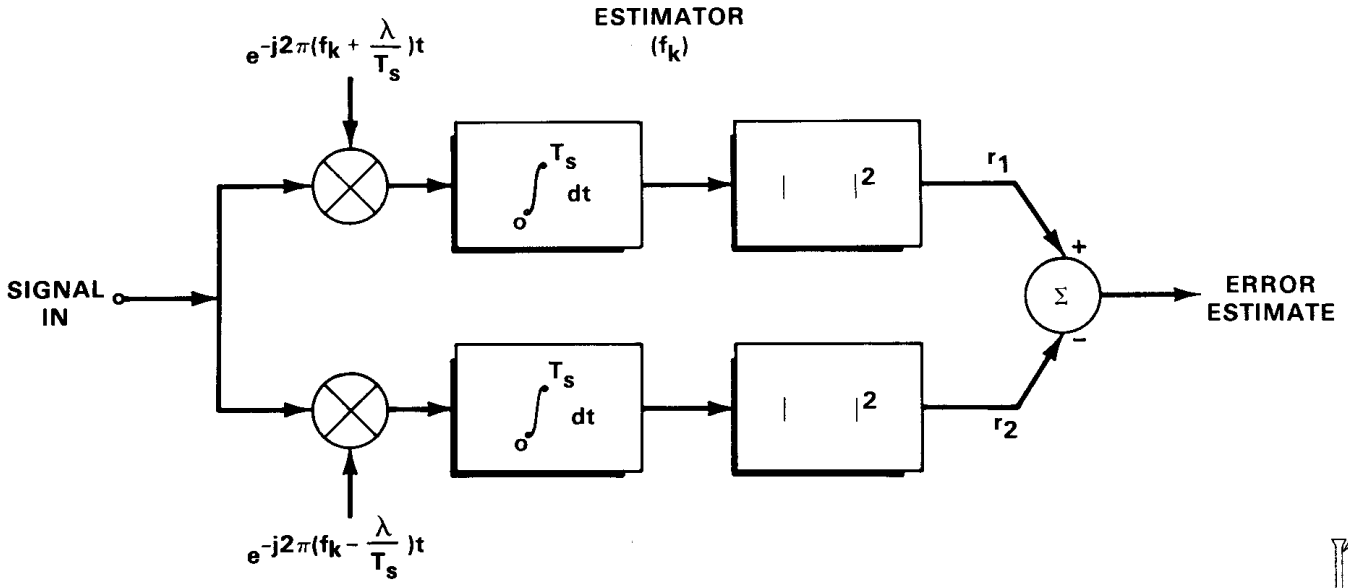

FIGURE 6

RECEIVED POWER REQUIREMENT FOR FREQUENCY TRACKING
(0.1 T_s^{-1} RMS Error, η = 1)

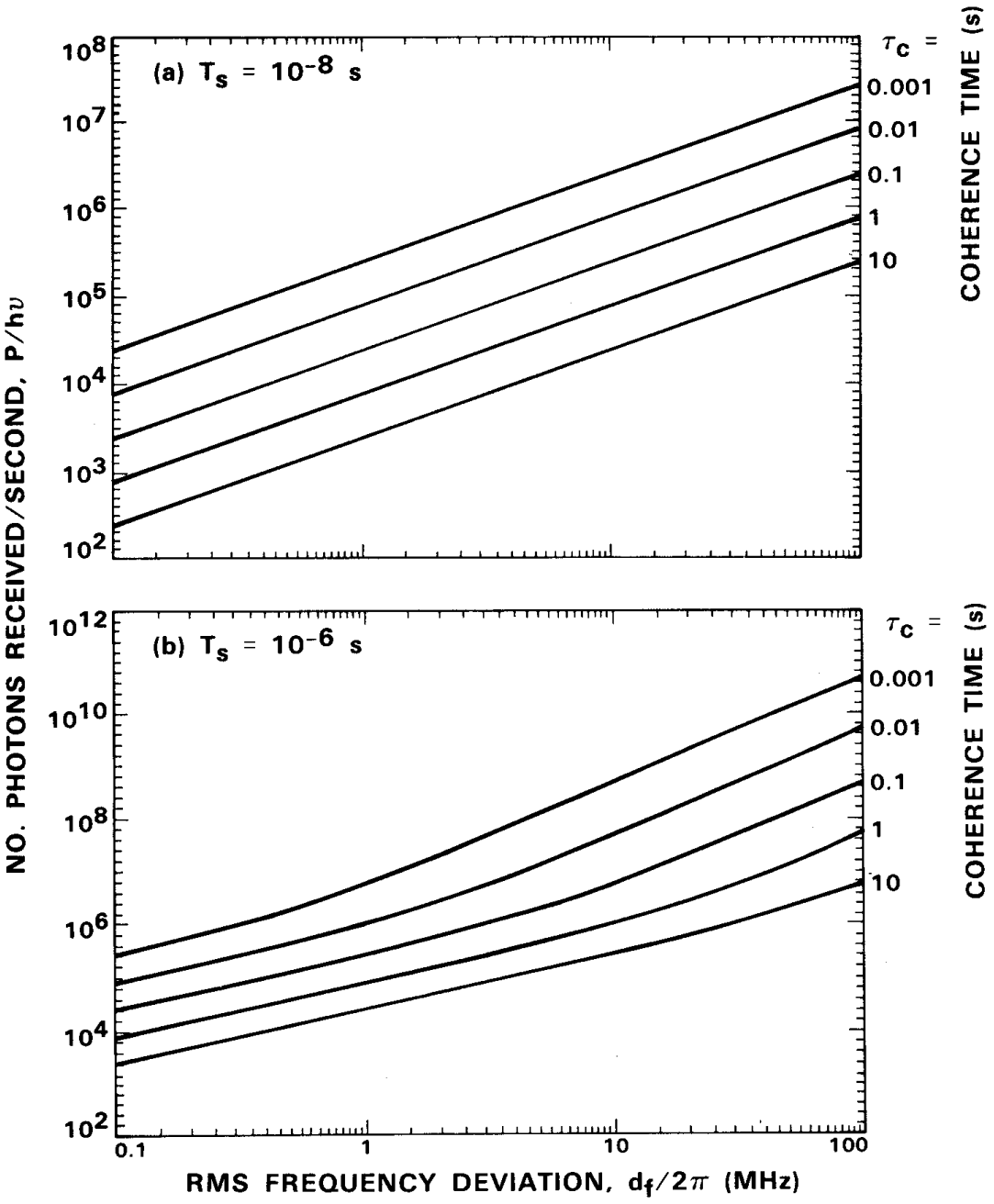

FIGURE 7

BAND-EDGE FREQUENCY TRACKER FOR MFSK

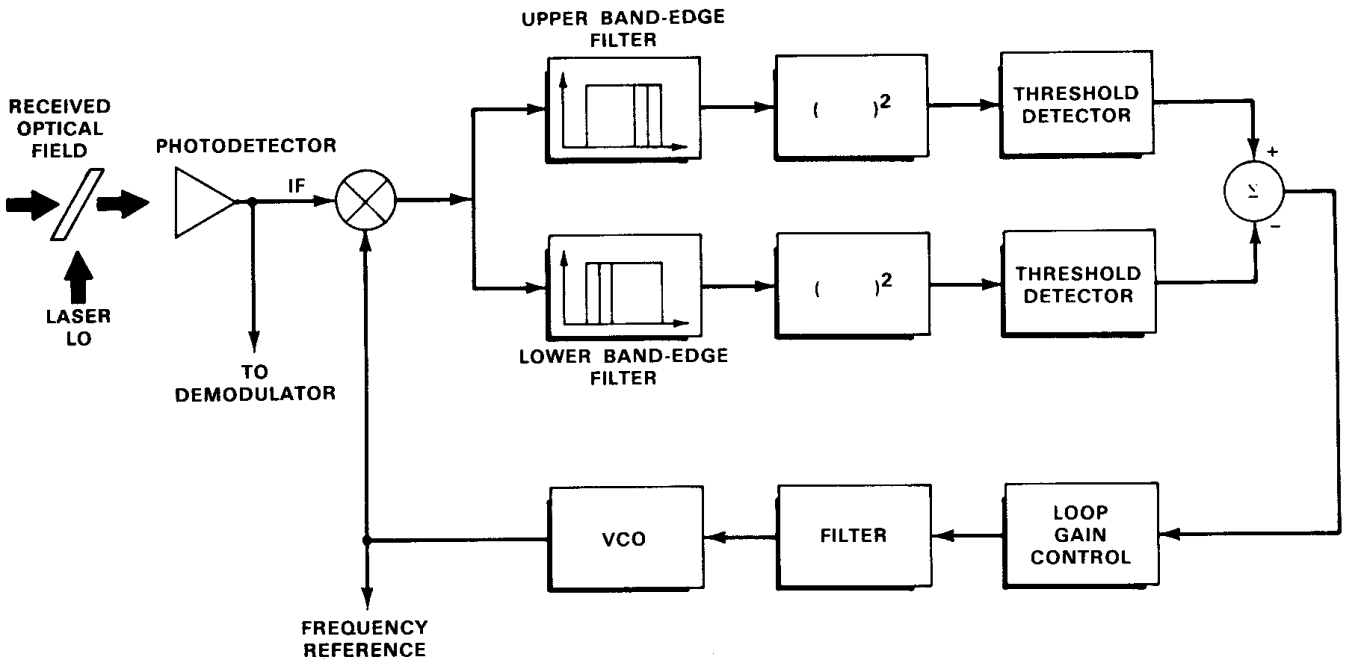

FIGURE 8

FREQUENCY TRACKING WITH DOPPLER SHIFT