

An Equivalent Linear Model for Marked and Filtered Doubly Stochastic Poisson Processes with Application to MMSE Linear Estimation for Synchronous m -ary Optical Data Signals

W. A. GARDNER, MEMBER, IEEE

Abstract—An equivalent linear model for minimum mean-squared error (MMSE) linear estimation of marked and filtered doubly stochastic Poisson processes is presented. The equivalence is employed to determine the structure of the MMSE noncausal steady-state linear receiver for synchronous m -ary optical data signals.

I. EQUIVALENT LINEAR MODEL

Consider the process $Y(t)$ composed of the sum of a marked and filtered doubly stochastic Poisson process [1] and a statistically independent zero-mean process $V(t)$,

$$Y(t) = \sum_{i=1}^{\infty} g(t, \tau_i, u_i) + V(t), \quad \forall t \geq t_0, \quad (1)$$

where $\{\tau_i\}_{1}^{\infty}$ are the random occurrence times of doubly stochastic Poisson counting process $N(t)$ which starts at $t = t_0$ and has stochastic rate process $\lambda(t)$, and the random marks $\{u_i\}_{1}^{\infty}$ are statistically independent of each other and of $\{\tau_i\}_{1}^{\infty}$, and are identically distributed sets of random variables, and where $g(\cdot, \cdot, \cdot)$ is a deterministic function.¹ Denote the minimum mean-squared error (MMSE) linear causal estimate of the stochastic rate process $\lambda(t)$, given observations $\{Y(\tau); t_0 \leq \tau \leq t\}$, by $\hat{\lambda}(t)/Y$ for $t \geq t_0$.

Consider also the (hypothetical) process $Z(t)$ defined by

$$Z(t) \triangleq \int_{t_0}^{\infty} E\{g(t, \tau, u)\} \lambda(\tau) d\tau + V(t) + W(t), \quad \forall t \geq t_0, \quad (2)$$

where u has the same distribution as u_i , and $W(t)$ is a zero-mean nonstationary process that is statistically independent of the remainder of $Z(t)$ and has autocorrelation function

$$k_W(t, \tau) = \int_{t_0}^{\infty} E\{\lambda(\sigma)\} E\{g(t, \sigma, u)g(\tau, \sigma, u)\} d\sigma, \quad (3)$$

and $\lambda(t)$, $V(t)$, and g are the same as in (1). Denote the MMSE linear causal estimate of $\lambda(t)$, given observations $\{Z(\tau); t_0 \leq \tau \leq t\}$ by $\hat{\lambda}(t)/Z$ for $t \geq t_0$.

Theorem: The MMSE linear causal filters for $\hat{\lambda}(t)/Z$ and $\hat{\lambda}(t)/Y$ are identical, and their impulse response function h is given by the solution to the linear integral equation

$$\int_{t_0}^t h(t, \sigma) k_Z(\sigma, \tau) d\sigma = k_{\lambda Z}(t, \tau), \quad \forall t, \tau \ni t_0 \leq \tau \leq t. \quad (4)$$

Paper approved by the Editor for Communication Theory of the IEEE Communications Society for publication without oral presentation. Manuscript received October 28, 1975; revised March 4, 1976.

The author is with the Department of Electrical Engineering, University of California, Davis, CA 95616.

¹We denote random quantities with boldface letters and samples of those random quantities with corresponding nonboldface letters.

Proof: It is well known that the optimum filter h_Z for $\hat{\lambda}(t)/Z$ is the solution to (4) (cf., [4, sect. 9.6]). Similarly, it is well known that the optimum filter h_Y for $\hat{\lambda}(t)/Y$ is the solution to the linear integral equation

$$\int_{t_0}^t h(t, \sigma) k_Y(\sigma, \tau) d\sigma = k_{\lambda Y}(t, \tau), \quad \forall t, \tau \ni t_0 \leq \tau \leq t. \quad (5)$$

Hence, we need only prove that $k_Y = k_Z$ and $k_{\lambda Y} = k_{\lambda Z}$. Straightforward evaluation from (2) yields

$$k_Z(t, \tau) = \int_{t_0}^{\infty} \int_{t_0}^{\infty} E\{g(t, \sigma, u)\} E\{g(\tau, \gamma, u)\} k_{\lambda}(\sigma, \gamma) d\sigma d\gamma + k_V(t, \tau) + k_W(t, \tau) \quad \forall t, \tau \geq t_0 \quad (6)$$

$$k_{\lambda Z}(t, \tau) = \int_{t_0}^{\infty} E\{g(\tau, \sigma, u)\} k_{\lambda}(\sigma, t) d\sigma \quad \forall t, \tau \geq t_0. \quad (7)$$

Also,

$$K_Y(t, \tau) \triangleq E\{Y(t)Y(\tau)\} = E\{[E\{Y(t)Y(\tau)/\{\lambda\}_{t_0}^{\infty}\}]\}. \quad (8)$$

From [1, eqs. (4.11) and (4.16)] we obtain

$$k_Y(t, \tau) = E\left\{\int_{t_0}^{\infty} \lambda(\sigma) E\{g(t, \sigma, u)g(\tau, \sigma, u)\} d\sigma + \int_{t_0}^{\infty} \int_{t_0}^{\infty} \lambda(\sigma)\lambda(\gamma) E\{g(t, \sigma, u)\} E\{g(\tau, \gamma, u)\} d\sigma d\gamma + k_V(t, \tau)\right\}. \quad (9)$$

Performing the expectation yields $k_Y \equiv k_Z$. Continuing,

$$k_{\lambda Y}(t, \tau) \triangleq E\{\lambda(t)Y(\tau)\} = E\{\lambda(t)[E\{Y(\tau)/\{\lambda\}_{t_0}^{\infty}\}]\}. \quad (10)$$

Employing [1, eq. (4.11)] yields

$$K_{\lambda Y}(t, \tau) = E\left\{\int_{t_0}^{\infty} \lambda(t)\lambda(\sigma) E\{g(t, \sigma, u)\} d\sigma\right\}. \quad (11)$$

Performing the expectation yields $k_{\lambda Y} \equiv k_{\lambda Z}$.

Remark 1: If we let $V \equiv 0$ and we let g be defined by

$$g(t, \tau, u) = \delta(t - \tau), \quad (12)$$

the results of this theorem reduce to the well-known result for unmarked and unfiltered processes [1, sect. 6.5.1]. See also [2] for a similar result for the special case where $\lambda(t)$ is a synchronous PAM signal.

Remark 2: The equivalence of this theorem is valid for prediction and noncausal smoothing as well as causal filtering. For a general estimation problem with arbitrary forward and reverse memory, (4) becomes

$$\int_{T_t} h(t, \sigma) k_Z(\sigma, \tau) d\sigma = k_{\lambda Z}(t, \tau) \quad \forall \tau \in T_t, \forall t \geq t_0, \quad (13)$$

where the interval T_t is determined by the desired memory of the filter h .

Remark 3: The equivalence of this theorem is valid for zero error-mean minimum error-variance [i.e., minimum variance, linear, unbiased (MVLU)] estimation as well as MMSE estimation, by replacing (13) with²

$$\int_{T_t} h'(t, \sigma) [k_Z(\sigma, \tau) - m_Z(\sigma)m_Z(\tau)] d\sigma = [k_{\lambda Z}(t, \tau) - m_{\lambda}(t)m_Z(\tau)] \quad \forall \tau \in T_t, \forall t \geq t_0; \quad (14)$$

the MVLU estimate is then given by

$$\hat{\lambda}'(t)/Y = \int_{T_t} h'(t, \tau) [Y(\tau) - m_Y(\tau)] d\tau + m_{\lambda}(t) \quad \forall t \geq t_0. \quad (15)$$

II. APPLICATION TO OPTICAL DIGITAL DATA COMMUNICATION

A. Equivalent Model

If we let g in (1) be defined by (with abuse of notation)

$$g(t, \tau_i, u_i) = u_i g(t - \tau_i), \quad (16)$$

then (1) is an appropriate model for the output of an optical detector with random photon arrival times $\{\tau_i\}$ and corresponding random avalanche gains $\{u_i\}$, and with dispersion function $g(\cdot)$, and additive thermal noise $V(t)$ (see [5] and refs. [3] and [4] therein). $\lambda(t)$ is the average rate of arrival of photons, which is assumed modulated by a stochastic information bearing signal. Specifically, let $\lambda(t)$ be a synchronous m -ary digital data signal

$$\lambda(t) = \sum_{n=1}^{\infty} p(t - t_0 - nT, a_n) + \lambda_0 \quad \forall t \geq t_0, \quad (17)$$

where the constant λ_0 is the rate of background optical radiation from the channel, and is not part of the information bearing signal. We assume that the data sequence $\{a_n\}_{1}^{\infty}$ that modulates the translates of the pulse $p(t, \cdot)$, is a stationary-of-order-2 sequence of discrete random variables with m allowable values $\{\alpha_q\}_{1}^m$. Denote the probability that $a_n = \alpha_q$ by P_q and the joint probability that $a_n = \alpha_q$ and $a_m = \alpha_r$ by $P_{qr}(n - m)$. Denote the mean and mean-squared values of u_i by \bar{u} and \bar{u}^2 . Then the equivalent linear "baseband" model for the output of the optical detector is, from (2),

$$Z(t) = \bar{u} \int_{t_0}^{\infty} g(t - \tau) \lambda(\tau) d\tau + X(t) \quad \forall t \geq t_0, \quad (18)$$

where X is the sum of V and W , and therefore has autocorrelation function

$$k_X(t, \tau) = k_V(t - \tau) + k_W(t, \tau) \quad \forall t, \tau \geq t_0, \quad (19)$$

² In (14) and (15) m_Z, m_{λ}, m_Y denote mean functions for Z, λ, Y , respectively.

$$k_W(t, \tau) = \bar{u}^2 \int_{t_0}^{\infty} m_\lambda(\sigma) g(t - \sigma) g(\tau - \sigma) d\sigma \quad \forall t, \tau \geq t_0. \quad (20)$$

We have assumed that the thermal noise is wide-sense stationary (WSS) for $t \geq t_0$. Since $m_\lambda(t)$ is periodic with period T , for $t \geq t_0$, then $k_W(t, \tau)$ is jointly periodic in t and τ , for $t \geq t_0$ and $\tau \geq t_0$. Hence, the effective additive noise $X(t)$ is wide-sense cyclostationary (WSCS) with period T for $t \geq t_0$. Furthermore, the cross correlation (7) is

$$k_{\lambda Z}(t, \tau) = \bar{u} \int_{t_0}^{\infty} g(t - \sigma) k_\lambda(\sigma, \tau) d\sigma \quad \forall t, \tau \geq t_0. \quad (21)$$

We are interested only in steady-state estimation (i.e., $t \geq t_0$), and the steady-state receiver is affected by the behavior of quantities in (17)–(21) only for $t \geq t_0$ and $\tau \geq t_0$. As a result, the only difference between the steady-state problem here and the steady-state estimation problem for the additive noise channel, which was solved in [3], is that here the effective additive noise is WSCS rather than WSS, and here the signal has a constant additive term λ_0 . The constant does not affect the method of solution of the integral equation (13) for the receiver; however, the nonstationarity of the noise requires a slight modification of the solution technique presented in [3, appendix]. The results are presented in the following subsection.

B. Receiver Structure

1) *Estimation of Posterior Probabilities of Data Digits:* Let the probability that $a_n = \alpha_q$, given observations $Y \triangleq \{Y(\tau); t_0 \leq \tau < \infty\}$ be denoted by $\Pr[a_n = \alpha_q/Y]$. This posterior probability can be interpreted as the image of Y under the functional $\Pr[a_n = \alpha_q/(\cdot)]$. This functional maps the sample waveform Y into the number $\Pr[a_n = \alpha_q/Y]$. Now, the ensemble of all such images together with the probability distribution for this ensemble, which derives from the probability distribution for Y , is a random variable; i.e., $\Pr[a_n = \alpha_q/Y]$ is a random posterior probability. It can be shown (cf., [3], [6]) that the MMSE linear noncausal steady-state estimate of this random posterior probability, given observations Y , is

$$\hat{\Pr}[a_n = \alpha_q/Y] = \int_{-\infty}^{\infty} h_q(nT - \tau) Y(\tau) d\tau \quad \forall n \geq 1 \quad (22)$$

$$h_q(t) \triangleq \sum_{q'=1}^m \sum_{i=-\infty}^{\infty} C_{q'q}(i) \mu_{q'}(t - iT), \quad (23)$$

where μ_q is the solution to the linear integral equation

$$\int_{-\infty}^{\infty} k_X(t, \tau) \mu_q(-\tau) d\tau = \tilde{p}_0(t, \alpha_q), \quad \forall t \in (-\infty, \infty). \quad (24)$$

The kernel $k_X(t, \tau)$ is defined by (19), with $t_0 = -\infty$ and

$$k_W(t, \tau) = \int_{-\infty}^{\infty} m_0(\sigma) g(t - \sigma) g(\tau - \sigma) d\sigma, \quad \forall t, \tau \in (-\infty, \infty) \quad (25)$$

$$m_0(\sigma) \triangleq \bar{u}^2 \left[\sum_{i=-\infty}^{\infty} \sum_{q=1}^m P_q p(\sigma - iT, \alpha_q) + \lambda_0 \right], \quad \forall \sigma \in (-\infty, \infty). \quad (26)$$

The function $\tilde{p}_0(t, \alpha_q)$ is defined by

$$\tilde{p}_0(t, \alpha) \triangleq \int_{-\infty}^{\infty} g(t - \tau) p_0(\tau, \alpha) d\tau \quad (27)$$

$$p_0(t, \alpha) \triangleq \bar{u} [p(t, \alpha) + \lambda_0(t)] \quad (28)$$

$$\lambda_0(t) \triangleq \begin{cases} \lambda_0, & \forall |t| \leq T/2, \\ 0, & \forall |t| > T/2. \end{cases} \quad (29)$$

$\mu_q(t)$ is the impulse response function of a generalized matched filter. It is matched to the modified form $\tilde{p}_0(t, \alpha_q)$ of the q th signal $p(t, \alpha_q)$ in additive WSCS noise $X(t)$. The coefficients $C_{qq'}(n)$ are given by the (qq') th element of the $m \times m$ matrix

$$C(n) = T \int_{-1/2T}^{1/2T} T(f) \exp(j2\pi n T f) df, \quad (30)$$

where the matrix $T(f)$ is given by

$$T(f) \triangleq (1/\bar{u}) [Q(f)L(f) + I]^{-1} Q(f). \quad (31)$$

The (qq') th elements of the matrices $Q(f)$ and $L(f)$ are defined by

$$Q_{qq'}(f) \triangleq \sum_{n=-\infty}^{\infty} P_{qq'}(n) \exp(-j2\pi n T f) \quad (32)$$

$$L_{qq'}(f) \triangleq \frac{1}{T} \sum_{i=-\infty}^{\infty} G(f - i/T) P_0(f - i/T, \alpha_q) M_{q'}(f - i/T), \quad (33)$$

where G , P_0 , and M_q are the Fourier transforms of g , p_0 , and μ_q , respectively. In (31), I is the $m \times m$ identity matrix.

The transfer function of the filter for $\hat{\Pr}[a_n = \alpha_q/Y]$ [Fourier transform of $h_q(\cdot)$] is from (23) and (30)

$$H_q(f) = \sum_{q'=1}^m T_{q'q}(f) M_{q'}(f). \quad (34)$$

The sequence of coefficients $\{C_{qq'}(n); -\infty < n < \infty\}$ have the interpretation of being the tap weights of a tapped-delay line with unit delay T (or the discrete pulse response function of a nonrecursive sampled data filter).³ Thus, the m posterior estimates $\{\hat{\Pr}[a_n = \alpha_q/Y]\}$ are the outputs, at time nT , of a parallel bank of m generalized matched filters, with transfer functions $\{M_q(f)\}$, followed by an $m \times m$ multiport tapped-delay line with matrix of transfer functions $T(f)$. This result directly parallels that obtained in [3]; see [3, fig. 2]. The only difference between h_q here in (23) and h_q in [3, eqs. (18) and (19)] is that the additive noise $X(t)$ here is WSCS rather than WSS (when $m_\lambda(t)$ is constant, $X(t)$ is WSS), and that the modified signal $p_0(t, \alpha_q)$, defined by (28), enters here in place of the unmodified signal $p(t, \alpha_q)$. The modification is simply the addition of a rectangular pulse $\lambda_0(t)$, defined by (29). $\lambda_0(t)$ is caused by the background optical radiation from the channel.

³The coefficients $C_{pq}(n)$ in this paper play the same role as the coefficients $c_{pq}(-n)$ in [3].

There are two special cases worthy of mention for which an explicit solution to (24) for the generalized matched filter can be obtained:

a) If the photoelectron dispersion function g is sufficiently narrow compared with the inverse bandwidth of the data pulses $\{p(t, \alpha_q)\}$ so that, under the integrals in (18)–(21), it can be replaced with a Dirac delta function with area $A =$ the area of g (this is equivalent to assuming perfect observation of the occurrence times $\{\tau_i\}$ when the additive thermal noise is negligible), and if the additive thermal noise $V(t)$ is white with intensity N_0 , then

$$\mu_q(t) = \frac{p(-t, \alpha_q) + \lambda_0(-t)}{N_0 + u^2 A^2 m_\lambda(-t)} \quad (35)$$

b) If the mean-rate function $m_\lambda(t)$ is a constant m_λ , then

$$M_q(f) = \frac{\bar{u}G^*(f)[P^*(f, \alpha_q) + \Lambda_0^*(f)]}{K_V(f) + u^2 A^2 m_\lambda |G(f)|^2}, \quad (36)$$

where $K_V(f)$ is the power spectral density for the additive thermal noise $V(t)$, and $\Lambda_0(f)$ is the Fourier transform of the rectangular pulse $\lambda_0(t)$.

2) *Estimation of Data Digits and Signal Waveform:* As discussed in [6] and [3] the MMSE linear noncausal steady-state estimates of the data digits $\{a_n\}$ and the signal waveform $\lambda(t)$ are easily obtained from the posterior probability estimates:

$$\hat{a}_n/Y = \sum_{q=1}^m \hat{\text{Pr}}[a_n = \alpha_q/Y] \alpha_q, \quad \forall n \gg 1, \quad (37)$$

$$\hat{\lambda}(t)/Y = \sum_{q=1}^m \sum_{n=n_0}^{\infty} \hat{\text{Pr}}[a_n = \alpha_q/Y] p_0(t - nT, \alpha_q), \quad (38)$$

$$\forall t \geq t_0 + n_0 T, n_0 \gg 1.$$

The estimate of the signal component of $\lambda(t)$, $s(t) \triangleq \lambda(t) - \lambda_0$, is simply

$$\hat{s}(t)/Y = \hat{\lambda}(t)/Y - \lambda_0. \quad (39)$$

Again, this result directly parallels that obtained in [3]. The posterior probability estimates are employed by an averaging device that computes an average alphabet letter for \hat{a}_n/Y or an average train of m -ary pulses for $\hat{s}(t)/Y$.

If one employs the criterion of MVLU rather than MMSE, then the preceding results in this section are all valid if λ_0 is replaced by 0 everywhere except in (39), and $Y(\tau)$ is replaced by $Y(\tau) - m_Y(\tau)$ in (22) and $\hat{\lambda}(t)/Y$ is replaced by $\hat{\lambda}(t)/Y + m_\lambda(t)$ in (39), and $P_{qq'}(i - j)$ is replaced by $P_{qq'}(i - j) - P_q P_{q'}$ in (32). Then, the rectangular pulse $\lambda_0(t)$ that occurs in additive combination with the signal pulses $\{p(t, \alpha_q)\}$ in the receiver structure will vanish. In this case, the receiver more closely parallels the receiver derived in [3] for additive noise channels.

It can be shown that the steady-state MMSE resulting from linear noncausal estimation of the data digits $\{a_n\}$ is, for $n \gg 1$,

$$\text{MMSE} = T \int_{-1/2T}^{1/2T} \alpha' [Q(f)L(f) + I]^{-1} Q(f) \alpha df. \quad (40)$$

It can also be shown that the steady-state minimum-error

variance resulting from MVLU noncausal estimation of $\{a_n\}$ is also given by (40), provided that $P_{qq'}(i - j)$ is replaced by $P_{qq'}(i - j) - P_q P_{q'}$ in the defining equation for Q . Formula (40) is also valid for additive noise channels, if the definitions of Q and L given in [3] are employed.

3) *Detection of Data Digits:* If it is desired to employ the linear receiver to make a hard decision as to which of the m letters $\{\alpha_q\}$ the n th digit a_n took on, then two alternatives arise:

a) *Linearly Constrained Estimation Theorist's Decision:* Decide $a_n = \alpha_q$ iff $|\hat{a}_n/Y - \alpha_q| \leq |\hat{a}_n/Y - \alpha_p|$ for $p = 1, 2, \dots, m$; i.e., pick the closest allowable value to the estimated value.

b) *Linearly Constrained Hypothesis Tester's Decision:* Decide $a_n = \alpha_q$ iff $\hat{\text{Pr}}[a_n = \alpha_q/Y] \geq \hat{\text{Pr}}[a_n = \alpha_p/Y]$ for $p = 1, 2, \dots, m$; i.e., pick the most pseudoprobable value, given observations Y .

It has been shown for one-shot detection for an additive noise channel [6] that a) yields lower probability of error than b) for PAM, where $p(t, \alpha_q) = \alpha_q p(t)$. In fact, b) is useless for PAM with $m > 2$. However, b) yields lower probability of error than a) for PSK. In fact, a) is useless for PSK with $m > 2$. A more general analysis of these two decision rules will be presented in a future paper.

III. CONCLUSIONS

The theorem on equivalent linear models presented in Section I enables engineers who are familiar with MMSE linear filtering for additive noise channels to transfer their knowledge and intuition directly to problems of MMSE linear filtering for optical channels with outputs that are modeled as marked and filtered doubly stochastic Poisson processes with additive noise. As an example of this, the theorem is employed in Section II to derive the structure and MSE of the MMSE linear noncausal steady-state receiver for synchronous m -ary optical signals. The parallels between this receiver and that derived in [3] for additive noise channels are pointed out.

The practical value of the solution for the optimum receiver presented in Section II is twofold.

1) Knowledge of the matched-filter-tapped-delay-line structure of the optimum linear receiver provides a basis for the design of practical (including adaptive) optical receivers.

2) Knowledge of the minimum attainable (with a linear receiver) mean-squared error enables comparison with the mean-squared error of previously considered linear optical receivers, and enables analysis of the effects on mean-squared error of photon pulse dispersion, random avalanche gain, and thermal noise in the optical detector, and of intersymbol interference due to optical signal pulse dispersion in the channel. In some cases, useful bounds on probability of error can be obtained from mean-squared error.

REFERENCES

- [1] D. L. Snyder, *Random Point Processes*. New York: Wiley, 1975.
- [2] D. G. Messerschmitt, "Optimum mean-square equalization for digital fiber optic systems," in *Proc. Int. Conf. Communications*, 1975, pp. 43.2–43.5.
- [3] W. A. Gardner, "The structure of least-mean-square linear estimators for synchronous m -ary signals," *IEEE Trans. Inform. Theory*, vol. IT-19, pp. 240–243, Mar. 1973. (Errata: The following errors in [3] should be noted. 1) Q_n^{ja} in (11) and $p_n(j, q)$ in (5) are identical; 2) $\{e_i\}$ should be replaced with $\{c_i\}$ in the first sentence of the last paragraph of Section III. 3) The definition following (A-6) should read $\tilde{c}_p(z) \triangleq \sum_{n=-\infty}^{\infty} c_n^p(z^{-n})$.)

- [4] L. E. Franks, *Signal Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1969.
- [5] E. V. Hoversten, D. L. Snyder, R. O. Harger, and K. Kurimoto, "Direct-detection optical communication receivers," *IEEE Trans. Commun.*, vol. COM-22, pp. 17-27, Jan. 1974.
- [6] W. A. Gardner, "Structurally constrained receivers for signal detection and estimation," *IEEE Trans. Commun.*, vol. COM-24, pp. 578-592, June 1976.