

AFFTC-PA-11011



On the Ungerboeck and Forney Observation Models for Offset QPSK

Michael Rice, Mohammad Saquib

AIR FORCE FLIGHT TEST CENTER
EDWARDS AFB, CA

12 MARCH 2012

A
F
F
T
C

Approved for public release; distribution is unlimited.

AIR FORCE FLIGHT TEST CENTER
EDWARDS AIR FORCE BASE, CALIFORNIA
AIR FORCE MATERIEL COMMAND
UNITED STATES AIR FORCE

On the Ungerboeck and Forney Observation Models for Offset QPSK

Michael Rice
Brigham Young University
Provo, UT 84602
Email: mdr@byu.edu

Mohammad Saquib
The University of Texas at Dallas
Richardson, TX 75080
Email: saquib@utdallas.edu

Abstract—The Ungerboeck and Forney observation models for offset QPSK are derived for three different representations for OQPSK. The representations and their corresponding Ungerboeck observation models are different, but the trellises used by the maximum likelihood sequence estimators possess the same complexity. The Forney observation models are obtained in straight-forward way for the bit-based representations for OQPSK, but there is no Forney observation model in the traditional sense for the symbol-based OQPSK representation. In the case of the symbol-based OQPSK representation, a method for whitening the resulting non-linear system is described.

I. INTRODUCTION

Frequency selective channels generate intersymbol interference (ISI) that can cause severe degradations in bit error rate performance. The most common ISI mitigation technique is some form of equalization using a linear filter (designed to meet the zero-forcing or minimum mean-squared error criteria), a non-linear decision-feedback filter, or, when the ISI span is finite, a maximum likelihood (ML) sequence estimator. These approaches have been well-studied for non-offset linear modulations. In 1972, Forney [1] published an early, rigorous analysis of the application of the Viterbi algorithm to perform ML sequence estimation in an ISI channel. Forney showed that the ML sequence estimator operates on symbol-spaced samples of the output of a filter matched to the channel-distorted pulse shape. The channel-distorted pulse shape does not generally satisfy the Nyquist No-ISI condition. Consequently, the noise samples at the output of the matched filter are correlated. Because the Viterbi algorithm requires uncorrelated noise samples, Forney describes a discrete-time noise whitening procedure preceding the ML sequence detector. The resulting equivalent discrete-time channel seen at the Viterbi algorithm input is called the *Forney observation model*. In 1974, Ungerboeck [2] described an approach that does not require the noise whitening process. Ungerboeck's approach operates directly on the symbol-spaced samples of the matched filter output and employs a modified version of the Viterbi algorithm. The resulting equivalent discrete-time channel seen at the Viterbi algorithm input is called the *Ungerboeck observation model*.

In general, ML sequence estimation performs equally well with the two observation models. However, this is not the case when reduced-state sequence estimation is employed. Hafeez and Stark [3], Schober, Chen, and Gerstacker [4], and Lončar

and Rusek [5] showed that the performance of a reduced state sequence estimator is better over the Forney observation model than over the Ungerboeck model.

ML sequence estimation for offset QPSK (OQPSK) has received little attention in the open literature. This is almost certainly due to the similarities in the structure and analysis techniques for the offset and non-offset cases. In the context of equalization, the decision feedback equalizer for OQPSK was developed by Bello and Pahlavan in 1984 [6] and revisited by Tu [7] in 1993.

In this paper, we derive the Ungerboeck and Forney observation models for OQPSK and identify the of the ML sequence estimator. Because there are different representations for OQPSK, we derive the Ungerboeck and Forney observation models for some representative cases. In particular we show the following:

- [Approach 1, Section III-A] The representation that treats OQPSK as a binary modulation with overlapping pulse shapes yields an Ungerboeck observation model that is a discrete-time LTI system — the Forney observation model is derived in a straight-forward way.
- [Approach 1', Section III-B] The representation that extends Approach 1 to explicitly incorporate the real/imaginary alternation of the binary symbols produces a linear (but not time invariant) system for the Ungerboeck observation model. In this case, the Forney observation model is obtained by whitening a linear operator, not by performing a power spectral factorization of an LTI system.
- [Approach 2, Section III-C] The representation that preserves the quaternary symbol-based structure of OQPSK produces a non-linear system for the Ungerboeck observation model. A noise-whitening procedure is possible, but it is not clear in what sense the resulting system represents a Forney observation model.

In each case, the Ungerboeck observation model is derived from basic principles. Next, the equivalent discrete-time system defined by the Ungerboeck observation model is identified and the equivalent discrete-time system is whitened to produce the Forney observation model.

II. BACKGROUND AND NOTATION

Let $s(t)$ denote the OQPSK signal. The OQPSK signal passes through a channel with impulse response $c(t)$ and is immersed in additive white Gaussian noise. Consequently, the received signal is

$$r(t) = s(t) * c(t) + w(t) \quad (1)$$

where $*$ is the convolution operator and $w(t)$ is a complex-valued, zero-mean, white Gaussian random process with power spectral density N_0 W/Hz. The log-likelihood function is

$$\tilde{\Lambda} = -\frac{1}{N_0} \int |r(t) - s(t) * c(t)|^2 dt. \quad (2)$$

The maximum likelihood sequence is symbol sequence that maximizes Λ or, after the elimination of data-independent terms, maximizes

$$\Lambda = 2\text{Re} \left[\int r(t) [s(t) * c(t)]^* dt \right] - \int |s(t) * c(t)|^2 dt. \quad (3)$$

There are different signal representations for OQPSK that may be used.¹ The chosen representations depends on the point of view one wishes to adopt for OQPSK. The choice has a profound impact on the resulting Ungerboeck and Forney observation models.

The representations define the transmission of $2N$ binary symbols $a(0), a(1), \dots, a(2N-1)$ where $a(n) \in \{-A, +A\}$. Motivated by the non-offset arrangement, the binary symbols may be thought of as quaternary symbols where $a(2k) + ja(2k+1)$ is the k th quaternary symbol for $k = 0, 1, \dots, N-1$. The first representation considered in this paper is

$$s_1(t) = \sum_{n=0}^{2N-1} x(n)p(t - nT_b) + w(t) \quad (4)$$

where $p(t)$ is a real-valued unit-energy square-root Nyquist pulse shape,² T_b is the bit time (sec/bit), and the binary symbols $x(n)$ are defined by

$$x(n) = \begin{cases} a(n) & n \text{ even} \\ ja(n) & n \text{ odd} \end{cases}. \quad (5)$$

Here, OQPSK is thought of as a binary modulation with overlapping pulse shapes. This formulation is preferred when the desire is to hide the real/imaginary alternation of the binary symbols from the intermediate signal processing. Keeping track of the real/imaginary alternation is pushed to the equalization and detection phase of the process. The second representation is the more traditional model:

$$s_2(t) = \sum_{k=0}^{N-1} a(2k)p(t - kT_s) + ja(2k+1)p(t - kT_s - T_s/2) \quad (6)$$

¹The two most common signal representations are used here: see Equations (4) and (6). Others include the cross-correlated trellis-coded quadrature modulation (XTCQM) [8, Chapter 3] (see also [9]) and an equivalent CPM representation [8, Chapter 3] (see also [10]).

²Because both T_b and $T_s = 2T_b$ spaced pulse trains are considered, there is potential for ambiguity in what is meant by a Nyquist pulse shape. Here, Nyquist pulse shape refers to the zero-ISI condition for T_s -spaced samples of $p(t) * p(-t)$, not to T_b -spaced samples.

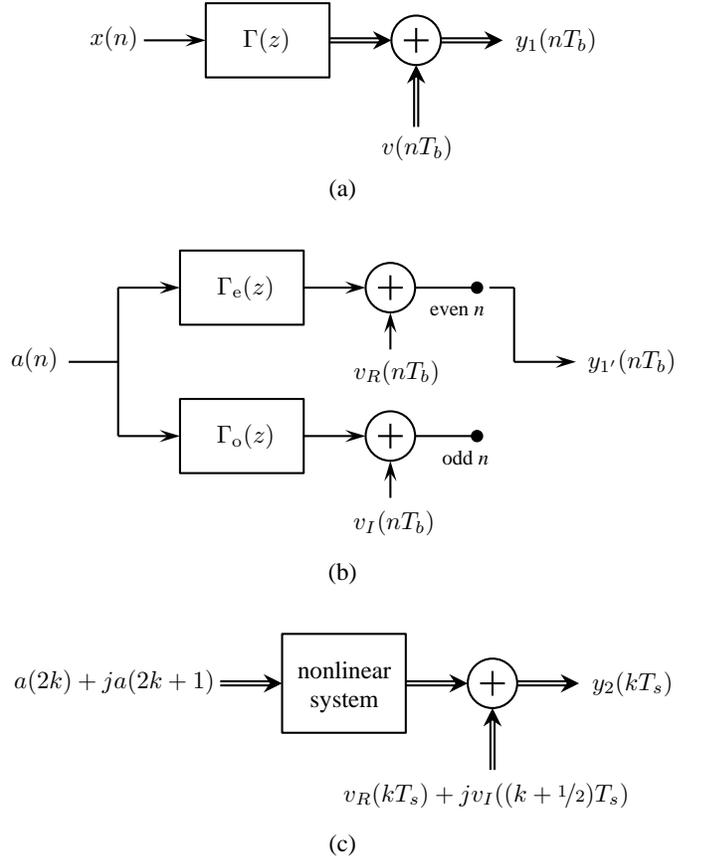


Fig. 1. The three equivalent discrete-time systems representing the Ungerboeck observations models for the approaches discussed in this paper: (a) The discrete-time system for Approach 1; (b) The discrete-time system for Approach 1'; (c) The discrete-time system for Approach 2. In (a), $\Gamma(z)$ is the z -transform of $\gamma(nT_b)$ [see (9)]. In (b), $\Gamma_e(z)$ is the z -transform of $h_e(\ell)$ and $\Gamma_o(z)$ is the z -transform of $h_o(\ell)$ [see (17)].

where $a(n)$ and $p(t)$ are the same as before and $T_s = 2T_b$ is the symbol time (sec/symbol). Here, the notion of a quaternary symbol is captured by using T_s as the time reference for the pulse train. This point of view is the one usually taken for OQPSK and preserves the explicit relationship between the non-offset and offset versions of QPSK.

III. THREE APPROACHES

A. Approach 1

Using $g(t) = p(t) * c(t)$ to represent the composite pulse shape, the modified log-likelihood function Λ for $s(t) = s_1(t)$ may be expressed as

$$\text{Re} \left(\sum_{n=0}^{2N-1} x^*(n) \left[2y(nT_b) - \sum_{\ell=-L_b}^{L_b} x(n-\ell)\gamma(\ell T_b) \right] \right) \quad (7)$$

where

$$y(nT_b) = \int r(t)g^*(t - nT_b)dt \quad (8)$$

$$\gamma(\tau) = \int g(t)g^*(t - \tau)dt. \quad (9)$$

The variable $y(kT_b)$ is interpreted as the output, at the instant $t = kT_b$, of a filter whose input is $r(t)$ and whose impulse response is matched to the composite pulse shape $g(t)$. The variables $\gamma(mT_b)$ are T_b -spaced samples of the autocorrelation of the composite pulse shape $g(t)$. The autocorrelation function $\gamma(\tau)$ is assumed to be zero outside the interval

$$-L_b T_b \leq \tau \leq L_b T_b \quad \text{or} \quad -L_s T_s \leq \tau \leq L_s T_s. \quad (10)$$

(Because $T_b = 1/2T_s$, L_s is approximately $1/2L_b$.)

The ML sequence estimator finds the sequence that maximizes (7). The argument of (7) may be re-written in a recursive way to produce a more suitable form for the Viterbi Algorithm. Following Ungerboeck, a useful recursion is obtained by exploiting the symmetry $\gamma(\tau) = \gamma^*(-\tau)$ and eliminating the term $|x(n)|^2 \gamma(0) = A^2 \gamma(0)$, which is the same for any sequence. Using the notation $\mathbf{a}_n = a(0), a(1), \dots, a(n)$, the result is

$$M_1(\mathbf{a}_n) = M_1(\mathbf{a}_{n-1}) + \text{Re} \left(x^*(n) \left[y(nT_b) - \sum_{\ell=1}^{L_b} x(n-\ell) \gamma(\ell T_b) \right] \right). \quad (11)$$

This formulation has the usual interpretation in the context of the Viterbi algorithm:

- The state at time step n may be defined by

$$x(n-1), x(n-2), \dots, x(n-L_b)$$

which are uniquely determined by

$$a(n-1), a(n-2), \dots, a(n-L_b).$$

The number of states (or nodes) at each time step is 2^{L_b} .

- The state transition from time step $n-1$ to time step n is defined by $a(n)$. Because $a(n) \in \{-A, +A\}$, there are 2 branches leaving each state and 2 branches entering each state.
- Each time step in the trellis corresponds to a bit time T_b .
- $M_1(\mathbf{a}_{n-1})$ is the partial path metric corresponding to the surviving path at time step $n-1$.
- The second term on the right-hand side of (11) defines the branch metric for the branch connecting the states corresponding to \mathbf{a}_{n-1} and \mathbf{a}_n .

The system model defined by (12) is illustrated in Figure 1 (a) where $\Gamma(z)$ is the z -transform of $\gamma(\ell T_b)$. The input is $x(n)$, the output is $y_1(n)$, and the relationship between the two is

$$y_1(nT_b) = \sum_{\ell=-L_b}^{L_b} x(n-\ell) \gamma(\ell T_b) + v(nT_b) \quad (12)$$

where $v(nT_b)$ is the sample, at $t = nT_b$ of $v(t) = w(t) * g^*(-t)$. The $v(nT_b)$ are a sequence of complex-valued zero-mean Gaussian random variables with autocorrelation function $R_1(k) = N_0 \gamma(kT_b)$. Equation (12) defines the Ungerboeck observation model for Approach 1. The Forney observation model is obtained from Figure 1 (a) by whitening the noise. The most common way to approach this is to perform a

power spectral factorization $\Gamma(z) = F(z)F^*(1/z^*)$ and use the system $1/F^*(1/z^*)$ as the whitening filter. See [1, pp. 366–368] or [11, Section 9.3-2] for more details.

The standard noise whitening approach produces an uncorrelated sequence of T_b -spaced noise samples. This is more than what is needed: in the next section we show that an uncorrelated sequence of $T_s = 2T_b$ -spaced samples is all that is required. Clearly, the more strict approach (uncorrelated T_b -spaced noise samples) satisfies the minimum requirement (uncorrelated T_s -spaced noise samples). But in the event the more strict requirement forces one to “work too hard,” we are motivated to ask if realizing the minimum requirement reduces the computational complexity.

For future reference, the Ungerboeck observation model (12) has the following vector/matrix formulation:

$$\mathbf{y}_1 = \Gamma_1 \mathbf{x} + \mathbf{v} \quad (13)$$

where \mathbf{y}_1 , \mathbf{x} , and \mathbf{v} are column vectors formed in the obvious way from $y_1(nT_b)$, $x(n)$, and $v(nT_b)$, respectively, and Γ_1 is the convolution matrix formed from $\gamma(\ell T_b)$. The autocorrelation matrix of noise vector is³

$$\mathbf{E}[\mathbf{v}\mathbf{v}^\dagger] = N_0 \Gamma_1. \quad (14)$$

B. Approach 1'

Approach 1 does not account for the fact that the binary data symbols $x(n)$ alternate between purely real and purely imaginary quantities; cf., (5). Using the substitutions (5) in (7) and using the notation $y(t) = y_R(t) + jy_I(t)$ and $\gamma(\tau) = \gamma_R(\tau) + j\gamma_I(\tau)$ produces a form of Λ similar to (12) except where the each term in the sum over n depends on the parity⁴ of n . For even n , the summand is

$$a(n) \left[2y_R(nT_b) - \sum_{\ell=-L_b}^{L_b} a(n-\ell) h_e(\ell) \right] \quad (15)$$

whereas for odd n , the summand is

$$a(n) \left[2y_I(nT_b) - \sum_{\ell=-L_b}^{L_b} a(n-\ell) h_o(\ell) \right] \quad (16)$$

where

$$h_e(\ell) = \begin{cases} \gamma_R(\ell T_b) & \ell \text{ is even} \\ -\gamma_I(\ell T_b) & \ell \text{ is odd} \end{cases} \quad (17)$$

$$h_o(\ell) = \begin{cases} \gamma_R(\ell T_b) & \ell \text{ is even} \\ \gamma_I(\ell T_b) & \ell \text{ is odd} \end{cases}$$

As before the resulting expression for Λ may be formulated recursively to create a metric suitable for use by the Viterbi algorithm. Again, exploiting the symmetry of $\gamma(\ell T_b)$ and eliminating the data-independent term yields

$$M_{1'}(\mathbf{a}_n) = M_{1'}(\mathbf{a}_{n-1}) + B(\mathbf{a}_n, \mathbf{a}_{n-1}) \quad (18)$$

³The transpose of the vector \mathbf{v} is \mathbf{v}^\top . The conjugate-transpose of \mathbf{v} is \mathbf{v}^\dagger . The complex conjugate of \mathbf{v} is \mathbf{v}^* .

⁴The *parity* of an integer is its property of being even or odd.

where the branch metric is connecting the state corresponding to \mathbf{a}_{n-1} to the state corresponding to \mathbf{a}_n is

$$B(\mathbf{a}_n, \mathbf{a}_{n-1}) = a(n) \left[y_R(nT_b) - \sum_{\ell=1}^{L_b} a(n-\ell)h_e(\ell) \right] \quad (19)$$

for even n , and

$$B(\mathbf{a}_n, \mathbf{a}_{n-1}) = a(n) \left[y_I(nT_b) - \sum_{\ell=1}^{L_b} a(n-\ell)h_o(\ell) \right] \quad (20)$$

for odd n . As before, $a(n-1), \dots, a(n-L_b)$ define the 2^{L_b} states; there are 2 transitions leaving and entering each state; and the state transitions occur at T_b -spaced intervals. The important differences include the following:

- For even n , only the real part of the matched filter output is required whereas for odd n , only the imaginary part of the matched filter output is required. Consequently, the real part of the matched filter output may be sampled at 1 sample/symbol. The imaginary part of the matched filter output may be sampled at 1 sample/symbol as well, but with sampling instants delayed (or offset) by T_b relative to the sampling instants applied to the real part. This is identical to the sampling regime for OQPSK in the AWGN environment.
- The branch metrics require only the real part of $\gamma(\ell T_b)$ for even ℓ or the imaginary part of $\gamma(\ell T_b)$ for odd ℓ .
- The branch metrics connecting the states are different depending on the parity of n . This difference is nothing more than the sign on $\gamma_I(\ell T_b)$ for odd ℓ [see (17)].

Whereas Approach 1 suggests that the ML sequence estimator for OQPSK requires more complexity than the ML sequence estimator for QPSK, Approach 1' shows that the ML sequence estimators for both OQPSK and QPSK have the same complexity.

The equivalent system defined by this version of Λ is illustrated in Figure 1 (b). The input is $a(n)$ [real-valued] and the output is $y_{1'}(n)$ [also real-valued]. The relationship between $a(n)$ and $y_{1'}(n)$ depends on the parity of n :

$$y_{1'}(n) = \begin{cases} \sum_{\ell=-L_b}^{L_b} a(n-\ell)h_e(\ell) + v_R(nT_b) & n \text{ even} \\ \sum_{\ell=-L_b}^{L_b} a(n-\ell)h_o(\ell) + v_I(nT_b) & n \text{ odd} \end{cases} \quad (21)$$

where $v_R(nT_b)$ and $v_I(nT_b)$ are the real and imaginary parts of $v(nT_b)$ defined in the previous section. The sequence of noise samples

$$v_R(0), v_I(T_b), v_R(2T_b), v_I(3T_b), \dots, v_I((2N-1)T_b)$$

is a sequence real-valued, zero-mean Gaussian random variables with autocorrelation function

$$R_{1'}(k, k') = \begin{cases} 0 & k - k' \text{ is odd} \\ 1/2 N_0 \gamma_R((k - k')T_b) & k \text{ and } k' \text{ are even} \\ 1/2 N_0 \gamma_I((k - k')T_b) & k \text{ and } k' \text{ are odd} \end{cases} \quad (22)$$

Equation (21) defines the Ungerboeck observation model for Approach 1'. Note that the equivalent discrete-time system of Figure 1 (b) is not an LTI system. One of the consequences is that the noise sequence is not stationary. For this reason, the standard approach (involving power spectral factorization) to produce the Forney observation model does not apply.

The equivalent discrete-time system of Figure 1 (b), however, a linear system. As such, the input output relationship may be expressed using vector/matrix notation and the Cholesky decomposition may be used to create a usable Forney observation model. The vector/matrix form of (21) is

$$\mathbf{y}_{1'} = \mathbf{\Gamma}_{1'} \mathbf{a} + \mathbf{v}_{1'} \quad (23)$$

where $\mathbf{a} = [a(0), a(1), \dots, a(2N-1)]^T$,

$$\mathbf{y}_{1'} = \begin{bmatrix} y_R(0) \\ y_I(T_b) \\ y_R(2T_b) \\ \dots \\ y_I((2N-1)T_b) \end{bmatrix} \quad \mathbf{v}_{1'} = \begin{bmatrix} v_R(0) \\ v_I(T_b) \\ v_R(2T_b) \\ \dots \\ v_I((2N-1)T_b) \end{bmatrix} \quad (24)$$

and

$$\mathbf{\Gamma}_{1'} = \begin{bmatrix} \gamma_R(0) & -\gamma_I(-T_b) & \gamma_R(-2T_b) & \dots & 0 \\ \gamma_I(T_b) & \gamma_R(0) & \gamma_I(-T_b) & \dots & 0 \\ \gamma_R(2T_b) & -\gamma_I(T_b) & \gamma_R(0) & \dots & 0 \\ \vdots & & & & \vdots \\ 0 & \dots & \gamma_R(2T_b) & \gamma_I(T_b) & \gamma_R(0) \end{bmatrix} \quad (25)$$

The autocorrelation matrix $\mathbf{v}_{1'}$ may be expressed in terms of the autocorrelation matrix of \mathbf{v} in the previous section using the identity

$$\mathbf{v}_{1'} = \mathbf{A}(\mathbf{v} + \mathbf{B}\mathbf{v}^*) \quad (26)$$

where

$$\mathbf{A} = \frac{1}{2} \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & -j & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ 0 & 0 & 0 & \dots & -j \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & -1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ 0 & 0 & 0 & \dots & -1 \end{bmatrix} \quad (27)$$

Based on the relationship (26), it is straight-forward to see that

$$\mathbf{E}[\mathbf{v}_{1'} \mathbf{v}_{1'}^T] = N_0 \mathbf{A} (\mathbf{\Gamma}_1 + \mathbf{B} \mathbf{\Gamma}_1^T \mathbf{B}) \mathbf{A}^T. \quad (28)$$

With some effort, it can be shown that

$$\mathbf{A} (\mathbf{\Gamma}_1 + \mathbf{B} \mathbf{\Gamma}_1^T \mathbf{B}) \mathbf{A}^T = \frac{1}{2} \mathbf{\Gamma}_{1'} \quad (29)$$

from which we have

$$\mathbf{E}[\mathbf{v}_{1'} \mathbf{v}_{1'}^T] = \frac{N_0}{2} \mathbf{\Gamma}_{1'}. \quad (30)$$

The relationship (30) forms the basis for generating the Forney observation model from the Ungerboeck observation model (23). Applying the Cholesky factorization $\mathbf{\Gamma}_a = \mathbf{\Upsilon} \mathbf{\Upsilon}^T$, the Forney observation model is produced by multiplying both sides of (23) by $\mathbf{\Upsilon}^{-1}$ to give

$$\tilde{\mathbf{y}}_{1'} = \mathbf{\Upsilon}^{-1} \mathbf{y}_{1'} = \mathbf{\Upsilon}^T \mathbf{a} + \tilde{\mathbf{v}}_{1'}. \quad (31)$$

$$\Lambda = \sum_{k=0}^{N-1} a(2k) \left(2y_R(kT_s) - \sum_{\ell=-L_s}^{L_s} \left[a(2(k-\ell))\gamma_R(\ell T_s) - a(2(k-\ell)+1)\gamma_I((\ell-1/2)T_s) \right] \right) \\ + \sum_{k=0}^{N-1} a(2k+1) \left(2y_I((k+1/2)T_s) - \sum_{\ell=-L_s}^{L_s} \left[a(2(k-\ell))\gamma_I((\ell+1/2)T_s) + a(2(k-\ell)+1)\gamma_R(\ell T_s) \right] \right) \quad (32)$$

$$M_2(\mathbf{s}_k) = M_2(\mathbf{s}_{k-1}) + a(2k) \left(y_R(kT_s) - \sum_{\ell=1}^{L_s} \left[a(2(k-\ell))\gamma_R(\ell T_s) - a(2(k-\ell)+1)\gamma_I((\ell-1/2)T_s) \right] \right) \\ + a(2k+1) \left(y_I((k+1/2)T_s) - \sum_{\ell=0}^{L_s} \left[a(2(k-\ell))\gamma_I((\ell+1/2)T_s) + a(2(k-\ell)+1)\gamma_R(\ell T_s) \right] \right) \quad (33)$$

$$y_2(kT_s) = \sum_{\ell=-L_s}^{L_s} \left[a(2(k-\ell))\gamma_R(\ell T_s) - a(2(k-\ell)+1)\gamma_I((\ell-1/2)T_s) \right] \\ + j \sum_{\ell=-L_s}^{L_s} \left[a(2(k-\ell))\gamma_I((\ell+1/2)T_s) + a(2(k-\ell)+1)\gamma_R(\ell T_s) \right] + v_R(kT_s) + jv_I((k+1/2)T_s) \quad (34)$$

It is easy to verify that the resulting noise vector $\tilde{\mathbf{v}}_{1'}$ is white. Equation (31) defines the Forney observation model for Approach 1'.

C. Approach 2

This approach uses $s_2(t)$ as the representation for the OQPSK signal. Substituting (6) into (3) and retaining the real part gives Equation (32) at the top of the page. The maximum likelihood sequence estimator finds the sequence that maximizes (32). The argument of (32) may be rewritten recursively by exploiting the symmetry of $\gamma(\tau)$ and eliminating the term $|a(2k) + ja(2k+1)|^2\gamma(0) = 2A^2\gamma(0)$ which is the same for any sequence. The result is given by (33) at the top of the page where $s(k) = a(2k) + ja(2k+1)$ is the k -th symbol and $\mathbf{s}_k = s(0), s(1), \dots, s(k)$. This formulation has the following interpretation in the context of the Viterbi algorithm:

- The state at time step k is defined by $a(2k-2) + ja(2k-1), \dots, a(2(k-L_s)) + ja(2(k-L_s)+1)$. The number of states at each time step is 4^{L_s} .
- The state transition from time step $k-1$ to time step k is defined by $a(2k) + ja(2k+1)$. Consequently, there are 4 branches leaving each state and 4 branches entering each state.
- Each time step in the trellis corresponds to a symbol time T_s .
- $M_2(\mathbf{s}_{k-1})$ is the partial path metric corresponding to the surviving path at time step $k-1$.
- The second term on the right-hand side of (33) defines the branch metric for the branch connecting the states corresponding to the sequences \mathbf{s}_{k-1} and \mathbf{s}_k .

Compared to the trellises of Approaches 1 and 1', the temporal spacing between the nodes is twice as long, but there are twice as many branches leaving each node. Because $L_s \approx L_b/2$, the

number of states here is approximately the same as the number of states in Approaches 1 and 1'. Consequently, the trellises have approximately the same complexity.

This system defined by this formulation is illustrated in Figure 1 (c). The inputs are the quaternary symbols $a(2k) + ja(2k+1)$ and the outputs are $y_2(kT_s) = y_R(kT_s) + jy_I((k+1/2)T_s)$ where $y_R(\cdot)$ and $y_I(\cdot)$ are the real and imaginary parts, respectively, of the matched filter output (8). The input/output relationship is given by (34) at the top of the page and defines the Ungerboeck observation model for Approach 2. This relationship does not represent an LTI system. Not only is it the case that the relationship (34) cannot be written as a convolution involving $a(2k) + ja(2k+1)$ and a complex-valued impulse response, but it is also the case that $y_2(kT_s)$ cannot be written as a linear operation on $a(2k) + ja(2k+1)$. In light of these observations, it does not appear that Approach 2 has a Forney observation model in the traditional sense.

However, all is not lost. Using vector/matrix notation it can be shown, with some effort, that the vector of outputs may be expressed as

$$\mathbf{y}_2 = \underbrace{\left[\mathbf{G}_1 + \frac{j}{2}\mathbf{G}_2(\mathbf{I} + \mathbf{F}) \right]}_{\mathbf{H}_1} \mathbf{s} + \underbrace{\left[\frac{j}{2}\mathbf{G}_2(\mathbf{I} - \mathbf{F}) \right]}_{\mathbf{H}_2} \mathbf{s}^* \\ + \underbrace{\mathbf{P}\mathbf{v} + \mathbf{P}\mathbf{B}\mathbf{v}^*}_{\mathbf{v}_2} \quad (35)$$

where

$$\mathbf{s} = \begin{bmatrix} a(0) + ja(1) \\ a(2) + ja(3) \\ \dots \\ a(2N-2) + ja(2N-1) \end{bmatrix},$$

$$\mathbf{G}_1 = \begin{bmatrix} \gamma_R(0) & \gamma_R(-T_s) & \gamma_R(-2T_s) & \dots & 0 \\ \gamma_R(T_s) & \gamma_R(0) & \gamma_R(T_s) & \dots & 0 \\ \gamma_R(2T_s) & \gamma_R(T_s) & \gamma_R(0) & \dots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & \dots & \gamma_R(0) \end{bmatrix},$$

$$\mathbf{G}_2 = \begin{bmatrix} \gamma_I(0.5T_s) & \gamma_I(-0.5T_s) & \dots & 0 \\ \gamma_I(1.5T_s) & \gamma_I(0.5T_s) & \dots & 0 \\ \gamma_I(2.5T_s) & \gamma_I(1.5T_s) & \dots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \dots & \gamma_I(0.5T_s) \end{bmatrix},$$

$$\mathbf{F} = \begin{bmatrix} 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix},$$

$$\mathbf{P} = \frac{1}{2} \begin{bmatrix} 1 & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & 1 & \dots & 0 & 0 \\ \vdots & & & & & \vdots & \\ 0 & 0 & 0 & 0 & \dots & 1 & 1 \end{bmatrix},$$

\mathbf{I} is the identity matrix, \mathbf{B} is defined in (27), and the noise vector \mathbf{v} is defined in (13). (The fact that both \mathbf{v} and its conjugate \mathbf{v}^* are required to describe \mathbf{v}_2 is a consequence of the fact that the noise samples are non-proper complex-valued Gaussian random variables [12]. Readers familiar with non-proper complex-valued Gaussian random variables will be familiar with this form.) The matrix-vector formulation of the Ungerboeck observation model may be expressed as

$$\mathbf{y}_2 = \mathbf{H}_1 \mathbf{s} + \mathbf{H}_2 \mathbf{s}^* + \mathbf{v}_2 \quad (36)$$

and the ML sequence estimate is

$$\hat{\mathbf{s}} = \underset{\mathbf{s}}{\operatorname{argmin}} \left\{ (\mathbf{y}_2 - \mathbf{H}_1 \mathbf{s} - \mathbf{H}_2 \mathbf{s}^*)^\dagger \Phi^{-1} (\mathbf{y}_2 - \mathbf{H}_1 \mathbf{s} - \mathbf{H}_2 \mathbf{s}^*) \right\}$$

where

$$\Phi = \mathbb{E}[\mathbf{v}_2 \mathbf{v}_2^\dagger] = N_0 \left(\mathbf{P} \Gamma_1 \mathbf{P}^\top + \mathbf{P} \mathbf{B} \Gamma_1^\top \mathbf{B} \mathbf{P}^\top \right) \quad (37)$$

is the autocorrelation matrix of the noise vector \mathbf{v}_2 [the matrix Γ_1 is defined in (13)]. A noise-whitened system is obtained by writing $\Phi = \Upsilon \Upsilon^\dagger$ based on the Cholesky decomposition and using $\tilde{\mathbf{y}}_2 = \Upsilon^{-1} \mathbf{y}_2$ for detection. The result is

$$\tilde{\mathbf{y}}_2 = \Upsilon^{-1} (\mathbf{H}_1 \mathbf{s} + \mathbf{H}_2 \mathbf{s}^*) + \tilde{\mathbf{v}}_2 \quad (38)$$

and the corresponding ML sequence estimate is

$$\hat{\mathbf{s}} = \underset{\mathbf{s}}{\operatorname{argmin}} \left\{ \left| \tilde{\mathbf{y}}_2 - \Upsilon^{-1} (\mathbf{H}_1 \mathbf{s} + \mathbf{H}_2 \mathbf{s}^*) \right|^2 \right\}. \quad (39)$$

This is clearly not a Forney observation model in the traditional sense because it is not clear that the resulting ISI model $\Upsilon^{-1} (\mathbf{H}_1 \mathbf{s} + \mathbf{H}_2 \mathbf{s}^*)$ has any merit from a computational point of view — that is, it could very well be the case that the ISI model cannot be written in a recursive form suitable for use with the Viterbi algorithm.

IV. CONCLUSIONS

We have derived the Ungerboeck observation models for offset QPSK operating in an ISI channel using three different points of view for OQPSK. In each case, the trellis and modified Viterbi algorithm corresponding to the Ungerboeck observation model was identified and described. The trellises and the associated Viterbi algorithm have the same complexity.

The Forney observation model is entirely different situation. The bit-based representations of OQPSK produce an *linear* discrete-time system for their corresponding Ungerboeck observation models. The Forney observation models are obtained from the Ungerboeck observation models in a straight-forward way. The symbol-based representation of OQPSK produces a *non-linear* discrete-time system for the Ungerboeck observation model. We show that the resulting discrete-time system can be whitened, but it is not clear if the resulting system possesses an ISI model that can be written recursively. Consequently, it is not clear if a Forney observation model in the traditional sense exists for this case.

ACKNOWLEDGMENTS

This work was supported by the Test Resource Management Center (TRMC) Test and Evaluation Science and Technology (T&E/S&T) Program through a grant from the Army PEO STRI Contracting Office under contract W900KK-09-C-0016.

REFERENCES

- [1] G. D. Forney, "Maximum-likelihood sequence estimation of digital sequences in the presence of intersymbol interference," *IEEE Transactions on Information Theory*, vol. 18, no. 3, pp. 363–378, May 1972.
- [2] G. Ungerboeck, "Adaptive maximum-likelihood receiver for carrier-modulated data-transmission systems," *IEEE Transactions on Communications*, vol. 22, no. 5, pp. 624–636, May 1974.
- [3] A. Hafeez and W. Stark, "Decision feedback sequence estimation for unwhitened ISI channels with applications to multiuser detection," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 9, pp. 1785–1795, December 1998.
- [4] R. Schober, H. Chen, and W. Gerstacker, "Decision-feedback sequence estimation for time-reversal space-time block coded transmission," *IEEE Transactions on Vehicular Technology*, vol. 53, no. 4, pp. 1273–1278, July 2004.
- [5] M. Lončar and F. Rusek, "On reduced-complexity equalization based on Ungerboeck and Forney observation models," *IEEE Transactions on Signal Processing*, vol. 56, no. 8, pp. 3784–3789, August 2008.
- [6] P. Bello and K. Pahlavan, "Adaptive equalization for SQPSK and SQPR over frequency selective microwave LOS channels," *IEEE Transactions on Communications*, vol. 32, pp. 609 – 615, May 1984.
- [7] J. Tu, "Optimum MMSE equalization for staggered modulation," in *Proceedings of the IEEE Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, 1–3 November 1993, pp. 1401 – 1405.
- [8] M. Simon, *Bandwidth-Efficient Digital Modulation with Applications to Deep Space Communications*. Hoboken, NJ: John Wiley & Sons, 2003.
- [9] L. Li and M. Simon, "Performance of coded OQPSK and MIL-STD SOQPSK with iterative decoding," vol. 52, no. 11, pp. 1890–1900, November 2004.
- [10] M. Simon, "Multiple-bit differential detection of offset QPSK," vol. 51, no. 6, pp. 1004–1011, June 2003.
- [11] J. Proakis and M. Salehi, *Digital Communications, Fifth Edition*. Boston, MA: McGraw-Hill, 2008.
- [12] P. Schreier and L. Scharf, *Statistical Signal Processing of Complex-Valued Data*. New York: Cambridge University Press, 2010.