Multipath Delay Estimation for Frequency Hopping Systems

Prashanth Hande, Lang Tong and Ananthram Swami

Abstract

The multipath delay estimation problem for a slow frequency hopping system is studied. High resolution delay estimation algorithms are proposed by exploiting invariance structures in the data packet. The proposed approach converts the problem of delay estimation using temporally received packets to one of estimating directions-of-arrival in array processing. Two closed-form estimators are developed. The first algorithm is based on the use of a single invariance and applies the ESPRIT algorithm. The second approach utilizes multiple invariances, and enforces the Cayley-Hamilton constraint in the signal subspace. It is shown, via an analysis of acquisition time, that the use of multiple invariances significantly shortens the number of hops required for parameter identifiability. Simulation examples also demonstrate the advantage of exploiting multiple invariances.

Keywords- *Delay estimation. Multiple Invariance. Eigenstructures.*

Prashanth Hande and Lang Tong are with the School of Electrical Engineering, Cornell University, Ithaca, NY 14853, e-mail: `hande,ltong@ee.cornell.edu`.

Ananthram Swami is with the Army Research Lab, AMSRL-CI-CN, 2800 Powder Mill Rd., Adelphi, MD 20783, email: `a.swami@ieee.org`

Address of corresponding author:

Lang Tong
School of Electrical Engineering
384 Rhodes Hall
Cornell University
Ithaca, NY 14853

Telephone: (607) 255-3900
Fax: (607) 255-9072
Email: ltong@ee.cornell.edu
http://www.ee.cornell.edu / ~ltong

July 14, 2001

1 Introduction

Among the two major spread spectrum techniques, frequency hopping (FH) techniques have been extensively studied for many years [1]. The hopping period can either span several bits for Slow Frequency Hopping (SFH) or can be a fraction of the bit period for Fast Frequency Hopping(FFH). The receiver for a FH system is conceptually a bank of filters arranged so that each filter in the bank is responsible for a portion of the total bandwidth [2]. Acquisition, the process of achieving time synchronization, has been the main focus of receiver design for many of the frequency hopping systems [3, 4, 5], with matched filter or the decorrelator being the usual detection schemes.

Conventionally, FH systems are considered narrow-band and the so called flat-fading model is invoked thereby ignoring the effect of the multipath delays to a considerable extent. The characterization of a signal as narrow-band is channel dependent and depends upon the coherence bandwidth of the channel. It was shown in [6] that the flat-fading model is not valid for a FH system whose bandwidth is comparable to the coherence bandwidth of the multipath channel. Thus the received signal in an SFH system at high data rates is frequency selective and estimation of delays becomes important. This paper addresses the delay estimation problem for an SFH system.

When the propagation channel can be modeled by discrete scatterers, the optimal receiver requires the knowledge of multipath delays. However, delay estimation for an SFH system is nontrivial in many aspects. One possible solution is the transmission of training bits. This, unfortunately, requires a multi-dimensional search, and has no closed-form solution. Although this problem can be circumvented by quantizing the delays into a finite number of levels, the quantization either limits the resolution of the estimator or requires high-dimensional matrix computations.

In this paper, we exploit the packet structure of the transmitted data to obtain high resolution* closed-form delay estimators. The estimators are blind in the sense that there is no training bits embedded in the data packet. We assume instead that all packets contain a fixed (although unknown) segment which may correspond to the packet identifier, synchronization bits, or the address segment. It is this packet structure that enables us to relate the problem of high resolution delay estimation to high resolution direction-of-arrival (DOA) estimation in array processing. *The key to this connection is the translation of SFH packets to the array elements in the problem of DOA estimations.* In this paper, we present two approaches: one based on the celebrated ESPRIT algorithm [7], the other on the more recent SPECC algorithm [9] that appears to be especially suitable for the delay estimation problem because it requires a much smaller number of packets.

The paper is organized as follows. The model and the invariance structures are presented in Section 2. The delay estimation algorithm based on ESPRIT, that exploits a single invariance structure, is presented in Section 3. Also included is an analysis of the acquisition time for the ESPRIT approach. In Section 4, a general version of the SPECC algorithm for delay estimation is presented. Simulation results and discussions are given in Section 5.

2   The Model

The baseband model of the FH received signal in a discrete multipath environment is of the form

$$y_n(t) = \sum_{i=1}^{P} \beta_i e^{-j2\pi f_n \tau_i} s_n(t - \tau_i; \mathbf{b}_n) + w_n(t), \tag{1}$$

---

*By high resolution we mean that, in the absence of noise, the algorithm has the ability to distinguish arbitrarily small relative delays.

where n is the hop index, $y_n(t)$ is the received signal, f_n is the hop frequency, $\mathbf{b}_n$ is the vector of transmitted symbols in the nth hop, $s_n(t; \mathbf{b}_n)$ is the baseband transmitted signal, and $w_n(t)$ is the white Gaussian noise component. Gain and delay parameters $\{\beta_i, \tau_i\}$ are unknown, and assumed independent of hop index. The hop frequency $\{f_n\}$ is known. We assume that the physical multipath channel is quasi time-invariant, with gains and delays constant across a few hops. In general, delays vary with carrier frequency, as do the amplitudes. The assumption of hop-independent gains is not required, as we shall see later.

A widely used approach, one that often leads to asymptotic optimality, is maximum likelihood detection based on following minimization

$$\widehat{\mathbf{b}}_n = \arg \min_{\mathbf{b}_n, \beta_i, \tau_i} \int_0^{T_h} |y_n(t) - \sum_{i=1}^{P} \beta_i e^{-j2\pi f_n \tau_i} s_n(t - \tau_i; \mathbf{b}_n)|^2 dt, \tag{2}$$

where T_h is the hop period (or dwell time), and $\widehat{\mathbf{b}}_n$ the estimated symbol vector in the nth hop. This minimization is difficult to accomplish particularly due to the non-linear dependence of the received signal on the delay parameters $\{\tau_i\}$. However, the ML-minimization problem is tractable if the delays are known, or can be constantly estimated first. Since the dependence on the channel gain parameters $\{\beta_i\}$ is linear, the minimization over $\{\beta_i\}$ and $\mathbf{b}_n$ can be separated. This motivates us to seek robust algorithms for the estimation of the delay parameters. It should be noted that, without joint optimization, this approach is no longer optimal, and tradeoffs between the simplicity of the algorithm and its performance need to be made.

We stress that, because of the nonlinear dependence of $s_n(\cdot; \cdot)$ on τ_i, the problem of delay estimation for the model in (1) is not trivial even if training is available, $i.e.$, $\mathbf{b}_n$ is known.

The discrete-time version of (1) can be written as

$$y_n[k] = \sum_{i=1}^{P} \beta_i e^{-j2\pi f_n \tau_i} s_{n,i}[k; \mathbf{b}_n] + w_n[k], \tag{3}$$

where, with a sampling interval T that satisfies the Nyquist criterion,

$$s_{n,i}[k; \mathbf{b}_n] \stackrel{\Delta}{=} s_n(kT - \tau_i; \mathbf{b}_n), \ y_n[k] \stackrel{\Delta}{=} y_n(kT), \ w_n[k] \stackrel{\Delta}{=} w_n(kT) \ .$$

Note that we do not assume a particular form for the modulation. The task at hand is to obtain estimates of the channel delays $\{\tau_i\}$ based on the received signal samples $y_n[k]$, without knowledge of the channel gains or the transmitted bit sequences.

2.1   The Invariance Structure

The key to the proposed approach is the invariance in the structure of the packet and the received signal. A typical SFH system [11] employs packet transmission with hopping from packet to packet. Some (consecutive) symbols in each packet form the address headers or the synchronization bits and are identical in every packet intended for the same recipient. We use these facts to recast the delay estimation problem as one of direction-of-arrival estimation in array processing.

We assume that the first[†] K symbols are fixed for all hops. This is a realistic assumption in some practical systems. For example, the SFH based PCS system described in [11] has a total of 16 QPSK symbols or 32 bits as sync and guard bits. Similarly, the Bluetooth standard [12, 13] specifies the use of 72 bits as access code and 54 bits as header information in each packet. The access code is used for synchronization and is the same for all packets originating in the same group (called a *piconet* in Bluetooth terms). The header consists of the member address, link flow bits and other control information. Thus it makes sense to assume that the first few bits of each packet received at a terminal are essentially the same.

[†]If the K symbols are located in the middle of the packet, our approach remains valid by restricting the estimation window to those parts of the signal not interfered by the data part of the packet.

This implies that

$$s_{n,i}[k; \mathbf{b}_n] = s_i[k], \qquad k = 0, 1, ..., K - 1. \tag{4}$$

The discrete time version of the received signal can be written as [c.f., eq (3)]

$$y_n[k] = \sum_{i=1}^{P} \beta_i e^{-j2\pi f_n \tau_i} s_i[k] + w_n[k], \quad k = 0, 1, ..., K - 1 \tag{5}$$

where, by exploiting (4), we have suppressed the dependency of $s_{n,i}[k, \mathbf{b}_n]$ on n and $\mathbf{b}_n$. Equation (5) is essentially in the DOA framework; here, k is the snapshot number, and n is the sensor index. We have P sources at bearings τ_i, with source waveforms given by $\beta_i s_i[k]$; f_n is the sensor displacement. Thus, each hop gives us a suite of K snapshots at one sensor (in contrast with the usual DOA problem where conceptually simultaneous sampling of the array elements, gives us one snapshot across all the sensors).

The K samples in (5) can be stacked into a row vector as

$$\mathbf{y}_n = [y_n[0] \ y_n[1] \ \ldots \ y_n[K - 1]], \tag{6}$$

so that

$$\mathbf{y}_n = \mathbf{a}_n \mathbf{S} + \mathbf{w}_n, \tag{7}$$

where

$$\mathbf{w}_n = [w_n[0] \ w_n[1] \ \ldots \ w_n[K - 1]], \tag{8}$$

$$\mathbf{a}_n = [e^{-j2\pi f_n \tau_1} \ e^{-j2\pi f_n \tau_2} \ \ldots \ e^{-j2\pi f_n \tau_P}], \tag{9}$$

$$\mathbf{S} = \begin{pmatrix} \beta_1 & & \\ & \ddots & \\ & & \beta_P \end{pmatrix} \begin{pmatrix} s_1[1] & \ldots & s_1[K] \\ \vdots & \vdots & \vdots \\ s_P[1] & \ldots & s_P[K]. \end{pmatrix}. \tag{10}$$

Suppose next that a sequence of hop frequencies $\{f_n\}$ can be put into a set of (ordered) $M + 1$-tuples $\mathcal{F} = \{(f_{0i}, f_{1i}, \cdots, f_{Mi}), i = 1, \cdots N\}$ such that, for some Δ_f, k_i, and all i

$$f_{0i} = f_{1i} + k_1 \Delta_f = f_{2i} + k_2 \Delta_f = \cdots = f_{Mi} + k_M \Delta_f. \tag{11}$$

Note that a particular hop frequency can appear in multiple positions in the M-tuple. For purposes of exposition, let us consider a simple example. Let the hop frequencies be chosen from integers $\{1, 2, \cdots, 5\}$, and suppose that we have received packets from the hop sequence

$$f_1 = 1, \ f_2 = 4, \ f_3 = 2, \ f_4 = 3, \ f_5 = 5, \ f_6 = 2.$$

There are many ways to form the partitions. If we choose $M = 1$ and $\Delta_f = 1$, we then have

$$\mathcal{F} = \{(f_3, f_1), (f_6, f_1), (f_5, f_2), (f_4, f_3), (f_2, f_4), (f_4, f_6)\}$$

corresponding to $k_1 = 1$; here $N = 6$. Choosing $M = 1$ and $k_1 = 2$ yields $N = 4$,

$$\mathcal{F} = \{(f_4, f_1), (f_2, f_3), (f_2, f_6), (f_5, f_4)\}.$$

If we choose $M = 4$ and $\Delta_f = 1$ (so that $k_i = 1, i = 1, ..., 4$), we have

$$\mathcal{F} = \{(f_5, f_2, f_4, f_3, f_1), (f_5, f_2, f_4, f_6, f_1)\}.$$

Let $\mathbf{y}_{ij}$ be the received data (row) vector corresponding to the header part of the packet associated with frequency f_{ij}, with $i = 0, ..., M$, and $j = 1, ..., N$. We can then form the following $N \times K$ matrices

$$\mathbf{Y}_0 = \begin{bmatrix} \mathbf{y}_{01} \\ \mathbf{y}_{02} \\ \vdots \\ \mathbf{y}_{0N} \end{bmatrix}, \ \mathbf{Y}_1 = \begin{bmatrix} \mathbf{y}_{11} \\ \mathbf{y}_{12} \\ \vdots \\ \mathbf{y}_{1N} \end{bmatrix}, \ \cdots \ \mathbf{Y}_M = \begin{bmatrix} \mathbf{y}_{M1} \\ \mathbf{y}_{M2} \\ \vdots \\ \mathbf{y}_{MN} \end{bmatrix}. \tag{12}$$

In the DOA framework, we have created $(M+1)$ possibly overlapping subarrays, each with N "sensors", and have collected data from K snapshots. If k_i are not all the same, then the different subarrays have different displacements. From (7-10), the data matrices defined above impose the following invariance structure

$$\mathbf{Y}_i = \mathbf{A}\Phi^{k_i}\mathbf{S} + \mathbf{W}_i, \quad i = 0, 1, \cdots, M, \tag{13}$$

where $k_0 := 0$, and $\{\mathbf{W}_0, \cdots, \mathbf{W}_M\}$ are matrices obtained by stacking the noise vectors. Matrices $\mathbf{A}$ and Φ defined by

$$\mathbf{A} = \begin{pmatrix} e^{-j2\pi f_{01}\tau_1} & \dots & e^{-j2\pi f_{01}\tau_P} \\ \vdots & \vdots & \vdots \\ e^{-j2\pi f_{0N}\tau_1} & \dots & e^{-j2\pi f_{0N}\tau_P} \end{pmatrix}, \quad \Phi = \begin{pmatrix} e^{j2\pi\Delta_f\tau_1} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & e^{j2\pi\Delta_f\tau_P} \end{pmatrix} \tag{14}$$

contain information about the delays. The task now is to estimate Φ given that the data matrices $\{\mathbf{Y}_i\}$ are available. This, of course, is exactly the same formulation used in ESPRIT [7] and Multiple Invariance (MI) ESPRIT [8].

3   Delay Estimation via Single Invariance

We first show how the ESPRIT algorithm can be applied to the multipath time-delay estimation problem. Next, we derive a closed-form expression for the average number of hops (sensors!) required to ensure identifiability.

3.1   The ESPRIT Algorithm

Suppose that we partition the set of hop frequencies into only two subsets, *i.e.*, $M = 1$ in (11). Assume for simplicity that $k_1 = 1$ (the general case is discussed at the end of the

section). Consider the auto-covariance matrix $\mathbf{R}_{yy}$ of the $(2N \times K)$ received signal matrix, $\mathbf{Y} = [\mathbf{Y}_0^T \ \mathbf{Y}_1^T]^T$ obtained by stacking together all the received data

$$\mathbf{R}_{yy} = E[\mathbf{Y}\mathbf{Y}^H] = \begin{bmatrix} \mathbf{R}_{00} & \mathbf{R}_{01} \\ \mathbf{R}_{10} & \mathbf{R}_{11} \end{bmatrix} ,$$

where

$$\begin{aligned} \mathbf{R}_{00} &= E[\mathbf{Y}_0\mathbf{Y}_0^H] = \mathbf{A}\mathbf{R}_{ss}\mathbf{A}^H + \mathbf{R}_{w_0,w_0} \\ \mathbf{R}_{10} &= E[\mathbf{Y}_1\mathbf{Y}_0^H] = \mathbf{A}\mathbf{\Phi}\mathbf{R}_{ss}\mathbf{A}^H + \mathbf{R}_{w_1,w_0} , \end{aligned} \quad (15)$$

where $\mathbf{R}_{w_0,w_0}$ and $\mathbf{R}_{w_1,w_0}$ are the noise covariance matrices, and $\mathbf{R}_{ss}$ is the auto-covariance of the transmitted signal, assumed to have full rank. Since the columns of the $\mathbf{S}$ matrix consist of delayed versions of the same signal, the assumption translates to the condition that the various delayed versions of the signal are not fully correlated. To be more precise, we can write

$$R_{ss}(m,n) = \sum_{k=1}^{K} E\{\beta_m \beta_n^* s(kT - \tau_m) s^*(kT - \tau_n)\} ,$$

where the expectation is with respect to the gains and the delays. A common assumption in multipath modeling is that the delays are independent of the gains, and that for discrete multipaths, the path gains are statistically independent of one another. In the case of zero-mean scattering, we have

$$R_{ss}(m,n) = \delta_{mn} \ E\{|\beta|^2\} \sum_{k=1}^{K} E\{|s(kT - \tau_m)|^2\}$$

which is a diagonal matrix.

The signal and noise subspace separation is achieved by considering the eigen-decomposition of $\mathbf{R}_{yy}$ as

$$\mathbf{R}_{yy} = \mathbf{U}_s \Sigma_s \mathbf{U}_s^H + \mathbf{U}_n \Sigma_n \mathbf{U}_n^H, \quad (16)$$

where $\mathbf{U}_s = [\mathbf{u}_1|\cdots|\mathbf{u}_P]$ and $\mathbf{U}_n = [\mathbf{u}_{P+1}|\cdots|\mathbf{u}_{2N}]$ represent the signal and noise subspaces respectively. We have assumed that $K \geq P$ (number of snapshots must be no smaller than the number of sources), and that $N \geq P$ (number of sensors must be no smaller than the number of sources); both assumptions are reasonable. Note that the signal space is represented by the eigenvectors corresponding to the P largest eigenvalues of the covariance matrix while the noise space is represented by the rest. The signal space corresponding to the two data sets can be obtained by splitting the signal space matrix $\mathbf{U}_s$ as

$$\mathbf{U}_s = \begin{bmatrix} \mathbf{U}_0 \\ \mathbf{U}_1 \end{bmatrix}, \tag{17}$$

where $\mathbf{U}_0$ and $\mathbf{U}_1$ are $N \times P$ unitary matrices representing the signal subspaces of the two data sets respectively. With the covariance matrices related as in (15), $\mathbf{U}_0$ and $\mathbf{U}_1$ represent the same subspace and this subspace is well known to be the same as span$\{\mathbf{A}\}$[14]. Assuming that the matrix $\mathbf{A}$ is full rank, this implies the existence of a $P \times P$ full rank matrix $\mathbf{T}$ satisfying

$$\mathbf{U}_0 = \mathbf{AT}, \qquad \mathbf{U}_1 = \mathbf{A\Phi T}. \tag{18}$$

The matrix Φ can be estimated if $\mathbf{U}_0$ and $\mathbf{U}_1$ are known. In practice, the presence of noise and a finite data set makes it possible for us to only obtain estimates, $\hat{\mathbf{U}}_0$ and $\hat{\mathbf{U}}_1$, of these subspaces. In this case we obtain an estimate of Φ as

$$\hat{\Phi} = \arg\min_{\Phi,\mathbf{A},\mathbf{T}} \left\| \begin{bmatrix} \hat{\mathbf{U}}_0 \\ \hat{\mathbf{U}}_1 \end{bmatrix} - \begin{bmatrix} \mathbf{AT} \\ \mathbf{A\Phi T} \end{bmatrix} \right\|_F^2. \tag{19}$$

By defining $\mathbf{B} = \mathbf{AT}$ and $\Psi = \mathbf{T}^{-1}\Phi\mathbf{T}$, we can rewrite the minimization as

$$\hat{\Psi} = \arg\min_{\Psi,\mathbf{B}} \left\| \begin{bmatrix} \hat{\mathbf{U}}_0 \\ \hat{\mathbf{U}}_1 \end{bmatrix} - \begin{bmatrix} \mathbf{B} \\ \mathbf{B\Psi} \end{bmatrix} \right\|_F^2. \tag{20}$$

Note that the parameters of interest are the diagonal elements of Φ which are the same as the eigenvalues of Ψ. Denoting the estimates of these eigenvalues by $\hat{\lambda}_i$, the delay parameters are estimated as

$$\hat{\tau}_i = \frac{\ln \hat{\lambda}_i}{j 2\pi \Delta_f}. \tag{21}$$

In order to avoid ambiguity of the estimates, we must have $\Delta_f \tau_i < 1 \; \forall i$; equivalently, we need $\tau_{\max} < 1/\Delta_f$. The minimization problem in (20) can be solved by either considering that only one of the two data sets $(\mathbf{Y}_0, \mathbf{Y}_1)$ is corrupted by noise or that both the data sets are corrupted by noise. The first assumption leads to the least squares (LS) ESPRIT implementation [15] while the second assumption leads to total least squares (TLS) ESPRIT [16].

We assumed $k_1 = 1$ in the preceding development; in the general case, matrix Φ will be replaced by Φ^{k_1}, and the identifiability conditions becomes $\tau_{\max} < 1/k_1 \Delta_f$. The frequency-spacing Δ_f can only be an integer multiple of the hop spacing. Since the hop spacing is determined by the signal bandwidth (approximately the reciprocal of the symbol time duration), we see that the above approach allows us to resolve only fractional time delays. This is both an advantage and a limitation. It is an advantage because in conventional approaches, fractional time delays are considered unresolvable; our algorithm exploits invariances in the data, and allows us to estimate fractional delays. It is a limitation, because it does not allow us to resolve true multipath delays causing ISI (in the usually understood sense of the term, with delays larger than the symbol period).

3.2 Acquisition Time for the ESPRIT Approach

An important consideration in the ESPRIT algorithm that has been underplayed until now is the selection of the separation frequency Δ_f. The algorithm requires frequencies separated by a fixed Δ_f but the received frequencies are pseudo-random. We note that packet reuse

is allowed in the algorithm as long as pairs of frequencies differ by a constant Δ_f. For the parameters to be identifiable by ESPRIT, the number of distinct frequency pairs N must be no less than the number of parameters $P^\ddagger$. The problem is to pick pairs of received frequencies whose difference occurs the maximum number of times. In other words, we choose Δ_f so as to minimize the average number of packets one must wait until the identifiability condition for ESPRIT is satisfied.

Once again consider the previous example but with a different realization of a random hop sequence

$$f_1 = 1, f_2 = 3, f_3 = 5, f_4 = 1, f_5 = 4, f_6 = 2, \cdots$$

Suppose that the channel is a two-ray multipath channel, and delays τ_1 and τ_2 need to be estimated. To apply ESPRIT with $M = 1$, if we choose $\Delta_f = 2$, we obtain the following set

$$\mathcal{F}_1 = \{(f_2, f_1), (f_3, f_2), (f_2, f_4), (f_5, f_6)\} \ ,$$

for which $N = 4$. The corresponding matrix $\mathbf{A}$ is a 4×2 matrix with full (column) rank of 2. The identifiability condition for ESPRIT is satisfied after receiving the third packet. If, on the other hand, $\Delta_f = 1$ is chosen, we then have

$$\mathcal{F} = \{(f_6, f_1), (f_5, f_2), (f_3, f_5), (f_6, f_4), (f_2, f_6)\} \ ,$$

where $N = 5$, This leads to a 5×2 matrix $\mathbf{A}$ with full column rank of 2. However, the identifiability condition for ESPRIT is satisfied only after the fifth packet is received, which means that one has to wait longer.

We now analyze the average number of packets one has to wait until the identifiability condition is satisfied. We note that the identifiability condition of ESPRIT is satisfied when

---

$^\ddagger N$ corresponds to the number of sensors and P the number of sources in the DoA estimation problem.

the number of distinct pairs in $\mathcal{F}$ reaches the number of parameters to be estimated. In the following we compute the average number of distinct pairs in $\mathcal{F}$.

Since the set of T hop frequencies usually constitutes an arithmetic series, we let the set $G = \{1, 2, \cdots, T\}$ represent the hop frequencies. The minimum separation between hop frequencies is represented by 1 in this formulation (without loss of generality, since the constant factors can be absorbed into the delays). Consider the received frequency to be a realization of a random sequence $f_n \in G$. Next, for an integer k, we form the set $\mathcal{F}_k = \{(f_{0i}, f_{1i})\}$ from the L-element sequence $\{f_n\}$ such that, for all i,

$$f_{0i} = f_{1i} + k. \tag{22}$$

Let H_k be the number of *distinct* pairs in $\mathcal{F}_k$. Assume that frequency hops are randomly chosen from G with equal probability. Then there are $T - k$ distinct pairs ($\{1, k+1\}, \{2, k+2\}$ and so on) that can possibly be members of $\mathcal{F}_k$. We can then obtain $E[H_k]$ by calculating, for each possible pair, the probabilities of all trials in which the pair occurs at least once. Since the probability of the occurrence of any pair is the same as any other pair, we need to evaluate this trial probability for any one pair and then sum it over all possible $T - k$ pairs. With $Pr(.)$ representing the probability of an event and $\{a, b\}$ representing any possible pair, we have

$$E[H_k] = (T - k)[Pr(\{a, b\} \text{occurs exactly once}) + Pr(\{a, b\} \text{occurs exactly twice}) + \cdots].$$

Each probability can be evaluated by calculating the probability that a occurs i times and b occurs j times. Let P_{ij} denote the probability that a occurs exactly i times and b occurs exactly j times. Then we have

$$E[H_k] = (T - k) \sum_{i=1}^{L-1} \sum_{j=1}^{L-i} P_{ij},$$

where

$$P_{ij} = \frac{1}{T^L} \binom{L}{i} \binom{L-i}{j} (T-2)^{L-i-j}.$$

So the expectation can be written as

$$E[H_k] = \frac{(T-k)}{T^M} \sum_{i=1}^{L-1} \sum_{j=1}^{L-i} \binom{L}{i} \binom{L-i}{j} (T-2)^{L-i-j}$$

This can be shown to evaluate to

$$E[H_k] = \frac{(T-k)}{T^L} [T^L - 2(T-1)^L + (T-2)^L].$$

This shows that the expected number of pairs differing by k decreases linearly with k making the choice of $k = 1$ optimal. For large T, the expected number of distinct pairs that differ by k can be approximated as

$$E[H_k] \approx (T-k)L(L-1)/T^2.$$

As expected, the expected wait time increases as the size of the hopset (T), or the number of invariances (L) increases; it decreases as the difference k increases. For evaluation of channel parameters in the P multipath case, we require a minimum of P pairs of frequencies for either the LS-ESPRIT or the TLS-ESPRIT. If we choose $k = 1$ as the frequency pair difference, then on an average we need to wait for $\hat{M}$ hops where

$$P \approx (T-1)\hat{M}(\hat{M}-1)/T^2$$

which implies that $\hat{M} \approx \sqrt{TP}$. For typical values like $T = 75$, $P = 3$ we need to wait on an average for $\sqrt{TP} = 15$ hops before we have a set of P pairs to which the algorithm can be applied. The waiting time is increased further if more than P pairs need to be used.

The above analysis brings out some of the drawbacks associated with the ESPRIT algorithm for channel estimation. For large T and M, there are nearly as many pairs that differ

by $k = 2, 3, \cdots$ as those that differ by $k = 1$. Since we are exploiting only the $k = 1$ structure, information provided by invariances associated with $k = 2, k = 3, \cdots$ remain unexploited indicating the possibility of better algorithms. The drawback of the ESPRIT approach arises from the fact that the ESPRIT algorithm is capable of exploiting only a single invariance. The uniform linear variation in frequencies of packets received, as assumed in our simulations, favors the ESPRIT algorithm. But in reality, this structure is rarely encountered and the received frequencies are more or less random. This prompts us to search for an algorithm that can work well when we have a small number of packets sharing multiple invariances. One such possibility is shown in the next section.

4   Delay Estimation by Multiple Invariance

4.1   The SPECC Algorithm

SPECC stands for Signal Parameter Estimation via the Cayley-Hamilton Constraint. The SPECC algorithm [9, 10] was developed for solving the parameter estimation problem involving multiple invariances exactly in the form given in (13). The original SPECC algorithm assumes $k_i = i$. Here, we present a more general version that allows non-consecutive k_i.

The task now is to estimate Φ assuming that data matrices $\mathbf{Y}_0, \cdots, \mathbf{Y}_M$, in (12)-(13), are the only measurements available. This is accomplished by separating the complex vector space to which the received signals belong into orthogonal signal and noise spaces. Proceeding as before, it can be shown that the signal subspaces ($\mathbf{U}_i$) of the data in each set ($\mathbf{Y}_i$) share the following structure

$$\mathbf{U}_0 = \mathbf{AT}, \quad \mathbf{U}_1 = \mathbf{A}\Phi^{k_1}\mathbf{T}, \quad \cdots, \quad \mathbf{U}_M = \mathbf{A}\Phi^{k_i}\mathbf{T}. \tag{23}$$

Defining $\mathbf{B} = \mathbf{AT}$ and $\Psi = \mathbf{T}^{-1}\Phi\mathbf{T}$, the structure can be re-written as

$$\mathbf{U}_0 = \mathbf{B}, \quad \mathbf{U}_1 = \mathbf{B}\Psi^{\mathbf{k_1}}, \quad \cdots, \quad \mathbf{U}_M = \mathbf{B}\Psi^{\mathbf{k_M}}. \tag{24}$$

In practice, the presence of noise and only a finite data set implies that only estimates of the signal subspaces, $\{\hat{\mathbf{U}}_0, \cdots, \hat{\mathbf{U}}_M\}$ are available. The search for Φ now assumes the form

$$\hat{\Phi} = \arg\min_{\Phi, \mathbf{A}, \mathbf{T}} \left\| \begin{bmatrix} \hat{\mathbf{U}}_0 \\ \vdots \\ \hat{\mathbf{U}}_M \end{bmatrix} - \begin{bmatrix} \mathbf{AT} \\ \vdots \\ \mathbf{A}\Phi^{\mathbf{k_M}}\mathbf{T} \end{bmatrix} \right\|_F^2. \tag{25}$$

An equivalent representation is

$$\hat{\Psi} = \arg\min_{\mathbf{B}, \Psi} \left\| \begin{bmatrix} \hat{\mathbf{U}}_0 \\ \vdots \\ \hat{\mathbf{U}}_M \end{bmatrix} - \begin{bmatrix} \mathbf{B} \\ \vdots \\ \mathbf{B}\Psi^{\mathbf{k_M}} \end{bmatrix} \right\|_F^2. \tag{26}$$

Note that the eigenvalues of Ψ are the diagonal elements of Φ.

This minimization has been dealt with in [8] where the proposed solution was a multidimensional search for the minimizing Ψ. A closed form solution, SPECC, based on an alternate constraint on the signal subspaces was proposed in [9]. If $k_i = i$, the problem reduces to the one considered in [9].

Let $p(\lambda) = \sum_{i=0}^{P} b_i\lambda^i$ be the characteristic polynomial of the $P \times P$ square matrix Ψ. The Cayley-Hamilton theorem enforces the following constraint on Ψ

$$p(\Psi) = b_0\mathbf{I}_P + b_1\Psi + \ldots + b_P\Psi^P = \mathbf{0}, \tag{27}$$

where $\mathbf{I}_P$ is the identity matrix of size P. The polynomial $p(\lambda)$ belongs to the class of polynomials called the annihilating polynomials of Ψ [20, page 221] all of which reduce the matrix Ψ to the zero matrix. The minimal polynomial is defined to be the annihilating

polynomial of minimum degree. It turns out that the characteristic polynomial is the minimal polynomial for a square matrix with distinct eigenvalues [20, page 240]. On the other hand, it is possible to have a minimal polynomial of degree less than that of the characteristic polynomial for matrices with non-distinct eigenvalues. Nevertheless the minimal polynomial has all the eigenvalues of the matrix as roots. Only that the multiplicity of the roots of the minimal polynomial is not the same as that of the eigenvalues of the corresponding square matrix. The minimal polynomial is unique and any annihilating polynomial has the minimal polynomial as a factor.

Consider an annihilating polynomial of Ψ, $c(\lambda) = \sum_{i=0}^{M} c_{k_i} \lambda^{k_i}$ with zero coefficients at all positions except maybe those at $\{k_i, i = 0, \cdots, M\}$. By definition of an annihilating polynomial we have

$$c(\Psi) = c_{k_0} \Psi^{k_0} + \cdots + c_{k_M} \Psi^{k_M} = \mathbf{0}. \tag{28}$$

This, along with (24), imposes the following constraint on the signal space

$$c_{k_0} \mathbf{U}_{k_0} + \cdots + c_{k_M} \mathbf{U}_{k_M} = \mathbf{0}. \tag{29}$$

Since the eigenvalues of Ψ occur as roots of $c(\lambda)$, it remains to be shown that the coefficients c_{k_i} can be determined from (29). In fact, we show that an annihilating polynomial of Ψ of the form given in (28) can always be determined from (29).

Let $\mathscr{Z}(g)$ denote the set of roots of a polynomial $g(\lambda)$. Let $m(\lambda) = \sum_{i=0}^{d} m_i \lambda^d$ be the minimal polynomial of Ψ. By definition of minimal polynomials, $\mathscr{Z}(m) = \{\lambda_1, \ldots, \lambda_P\}$ is the set of eigenvalues of the size P square matrix Ψ.

Theorem 1 *Let* $\mathbf{U}_{k_i}$, $\mathbf{B}$ *and* Ψ *be as defined before. Without loss of generality, we assume that the integers* k_i *are strictly increasing i.e.,* $0 \leq k_0 < k_1 < \cdots < k_M$. *Let* $m(\lambda)$ *be the*

minimal polynomial of Ψ. *Consider the minimization*

$$\{a_{k_i}\} = \arg \min_{\alpha_{k_i}, \sum \alpha_{k_i}^2 = 1} \|\alpha_{k_0}\mathbf{U}_{k_0} + \ldots + \alpha_{k_M}\mathbf{U}_{k_M}\|_F. \tag{30}$$

Let $a(\lambda) = a_{k_0}\lambda^{k_0} + a_{k_1}\lambda^{k_1} + \ldots + a_{k_M}\lambda^{k_M}$. *Then we have the following*

1. *If* $M \geq d$ *and* $\mathbf{B}$ *is full column rank, then* $\mathcal{Z}(m) \subseteq \mathcal{Z}(a)$.

2. *If* Ψ *is diagonalizable as* $\Psi = \mathbf{MDM}^{-1}$ *and* $\mathbf{BM}$ *has no zero columns, then* $\mathcal{Z}(p) \subseteq \mathcal{Z}(a)$.

3. *If* Ψ *has distinct eigenvalues and* $\mathbf{BM}$ *has no zero columns, then* $\mathcal{Z}(p) = \mathcal{Z}(a)$.

Remarks: Notice that the characteristic polynomial is minimal ($d = P$) when the eigenvalues are distinct. The theorem implies that $P + 1$ data sets structured as in (23) for some distinct set of integers $\{k_0, k_1, \cdots\}$ are sufficient for the estimation of the distinct eigenvalues of Ψ. Note that there is no obvious way to recognize the eigenvalues from $\mathcal{Z}(a)$, the set of roots of $a(\lambda)$. For the particular case of channel estimation, the eigenvalues are of unit amplitude and hence identifiable from the set of elements in $\mathcal{Z}(a)$ provided that no other root has unit amplitude.

Notice also, that if Ψ is diagonalizable, then the full rank condition on $\mathbf{B}$ can be relaxed. It suffices to have the row space of $\mathbf{B}$ orthogonal to no eigenvectors of Ψ which translates to the no zero column condition on $\mathbf{BM}$. This implies that the theorem holds even when the rank of the matrix $\mathbf{B}$ is less than P, the number of multipaths. Since the Ψ matrix for the estimation problem turns out to be diagonalizable with $\mathbf{M} = \mathbf{T}^{-1}$, this removes the ESPRIT restriction that the invariance be present among at least P data packets in each set and makes it possible to exploit invariances even among data sets with just one packet.

Proof: Since all eigenvalues of Ψ occur as roots of any annihilating polynomial, we need to show first that $a(\lambda)$ is an annihilating polynomial of Ψ. What concerns us now

is the particular form of $a(\lambda) = \sum_{i=0}^{M} a_{k_i} \lambda^{k_i}$. Is there any annihilating polynomial $c(\lambda) = \sum_{i=0}^{M} c_{k_i} \lambda^{k_i}$ with zero coefficients at all positions except maybe those at $\{k_i, i = 0, \cdots, M\}$? We show below that such a $c(\lambda)$ exists and that this results in $a(\lambda)$ being an annihilating polynomial.

Recall that any annihilating polynomial has the minimal polynomial $m(\lambda)$ as a factor. To prove that an annihilating polynomial $c(\lambda)$ of the given form exists, we need to show that a polynomial $r(\lambda) = \sum_{i=0}^{l} r_i \lambda^i$ exists such that $c(\lambda) = m(\lambda) r(\lambda)$. Comparing the largest power on both sides, we see that $k_M = d + l$. The polynomial $c(\lambda) = \sum_{i=0}^{M} c_{k_i} \lambda^{k_i}$ has $k_M - M$ zero coefficients and we have a total of $l + 1$ variables $\{r_0, \cdots, r_l\}$ with which we can set these coefficients to zero. This is always possible if $K_M - M = d + l - M < l + 1$ which is true if $M \geq d$. So an annihilating polynomial $c(\lambda)$ of the same form as $a(\lambda)$ always exists under the given conditions.

That $a(\lambda)$ is also an annihilating polynomial is now easily seen. Since $c(\lambda)$ is an annihilating polynomial of Ψ and since the subspaces are related as in (23), we have

$$c_{k_0} \mathbf{U}_{k_0} + \cdots + c_{k_M} \mathbf{U}_{k_M} = \mathbf{0}. \tag{31}$$

Since $\{a_{k_i}, i = 0, \cdots, M\}$ is a minimizer of (30) we have

$$||a_{k_0} \mathbf{U}_{k_0} + \cdots + a_{k_M} \mathbf{U}_{k_M}||_F \leq ||c_{k_0} \mathbf{U}_{k_0} + \cdots + c_{k_M} \mathbf{U}_{k_M}||_F = \mathbf{0}.$$

Since the norm is non-negative, we have

$$||a_{k_0} \mathbf{U}_{k_0} + \cdots + a_P \mathbf{U}_{k_M}|| = ||a_{k_0} \mathbf{B} \Psi^{k_0} + \cdots + a_{k_M} \mathbf{B} \Psi^{k_M}|| = \mathbf{0},$$

which implies

$$\mathbf{B}(a_{k_0} \Psi^{k_0} + \cdots + a_{k_M} \Psi^{k_M}) = \mathbf{0}.$$

Under full rank assumption on $\mathbf{B}$, it can be concluded that

$$a_{k_0} \Psi^{k_0} + \cdots + a_{k_M} \Psi^{k_M} = \mathbf{0}. \tag{32}$$

This indicates that $a(\lambda)$ is an annihilating polynomial for the matrix Ψ and the first part of the theorem stands proved.

If Ψ is diagonalizable, then the full rank condition on $\mathbf{B}$ is not required to show that the polynomial $a(\lambda)$ is an annihilating polynomial. To see this, diagonalize Ψ in (32) to obtain

$$\mathbf{BM}(a_{k_0}\mathbf{D}^{k_0} + \cdots + a_{k_M}\mathbf{D}^{k_M})\mathbf{M}^{-1} = \mathbf{0}.$$

Since $\mathbf{M}$ is full rank, we can write this as

$$\mathbf{BM}\tilde{\mathbf{D}} = \mathbf{0},$$

where $\tilde{\mathbf{D}} = (a_{k_0}\mathbf{D}^{k_0} + \cdots + a_{k_M}\mathbf{D}^{k_M})$ is also diagonal. If $\mathbf{BM} = [\mathbf{c}_1 \ \ldots \mathbf{c}_P]$ and $\{\beta_i, i = 1, \ldots, d\}$ are the diagonal elements of $\tilde{\mathbf{D}}$ then we have

$$[\beta_1\mathbf{c}_1 \ \cdots \ \beta_P\mathbf{c}_P] = \mathbf{0}, \tag{33}$$

which indicates that if none of the columns of $\mathbf{BM}$ is zero, then we have

$$\tilde{\mathbf{D}} = (a_{k_0}\mathbf{D}^{k_0} + \cdots + a_{k_M}\mathbf{D}^{k_M}) = \mathbf{0}. \tag{34}$$

From this, we conclude that $a(\lambda)$ is an annihilating polynomial of Ψ and hence includes all the eigenvalues of Ψ among its roots. This completes the proof of the theorem. $\square$.

The SPECC algorithm based on Theorem 1 can be described as follows. Collect frequencies into sets $\{f_{k_0 i}, \cdots, f_{k_M i}\}$ such that

$$f_{k_0 i} + k_0\Delta_f = f_{k_1 i} + k_1\Delta_f = f_{k_2 i} + k_2\Delta_f = \cdots = f_{k_M i} + k_M\Delta_f. \tag{35}$$

The header data samples of packets corresponding to frequencies in each set is collected into $\{\mathbf{Y}_{k_i}, i = 0, \cdots, M\}$. The signal subspace estimates $\{\hat{\mathbf{U}}_{k_i}, i = 0, \cdots, M\}$ corresponding to each data set is obtained by eigen-decomposing an estimate of the covariance matrix of the data as before.

We solve the minimization problem,

$$\{a_{k_0}, \ldots, a_{k_M}\} = \arg \min_{\{\alpha_{k_0}, \cdots, \alpha_{k_M}\}} \|\alpha_{k_0}\hat{\mathbf{U}}_{k_0} + \ldots + \alpha_{k_M}\hat{\mathbf{U}}_{k_M}\|_F, \tag{36}$$

assuming a normalization constraint on $\{\alpha_{k_i}\}$ and find the roots of the resulting polynomial $a(x) = a_{k_0}x^{k_0} + a_{k_1}x^{k_1} + \ldots + a_{k_M}x^{k_M}$. Note that the minimization is quadratic in nature and hence has a unique solution which can be obtained by taking the SVD of the appropriate matrix. Next, we estimate the eigenvalues as the P roots nearest to the unit circle. Finally, we estimate the delays via (21).

4.2   The Acquisition Time for the SPECC Approach

The SPECC algorithm was shown to require a minimum of $P+1$ data sets sharing P invariances among themselves. On the other hand, the ESPRIT algorithm requires a minimum of two data sets and can do no better when more are available. ESPRIT requires that a minimum of P packets be available in each data set whereas the SPECC algorithm works even with just one packet in each of the $P+1$ data sets. This factor contributes towards decreasing the waiting time quite drastically when compared to the ESPRIT algorithm.

In Section 3.2 it was shown that we need to wait for $\sqrt{TP}$ number of packets, on an average, before the ESPRIT algorithm can be employed for delay estimation. We seek a similar result for the SPECC algorithm in this section. The problem can be cast in the integer domain as follows. Consider M realizations of a random variable X which can take values in the domain $G = \{1, 2, \cdots, T\}$. We need to find the M for which the average number of distinct numbers among the M realizations of X is at least as large as $P+1$. The average number of distinct numbers in the M realizations of the random variable X can be

shown to be given by

$$E = T(1 - \frac{(T-1)^M}{T^M}).$$ (37)

For large T, we can approximate the average number as $E \approx M$ so that when $T = 75$ and $P = 3$, we have $E = P + 1$ when $\hat{M} = 4$. The waiting time is reduced from that for 15 packets in the ESPRIT case to 4 packets in the SPECC case.

The occurrence of P pairs differing by the same constant and the occurrence of $P + 1$ distinct frequencies, respectively are the two events that need to be satisfied before the ESPRIT and SPECC algorithm can be applied. The reduction in the waiting time manifests itself more clearly when we compare the probabilities of the two events against the number of received packets. The probabilities are plotted in Fig. 1 and it is immediately obvious that the SPECC algorithm requires fewer packets than does ESPRIT. In particular, we considered $T = 75$ and $P = 3$. The SPECC algorithm requires about 6 packets before estimation can start with probability one. The ESPRIT algorithm requires 20 packets before estimation can proceed with probability 0.8.

5   Simulations and Discussion

The simulation setup is first described in detail; next, we present typical simulation results, and finally we discuss these results, leading to insights into the performance of the various algorithms considered here.

5.1   Simulation Setup

We compare the performance of the ESPRIT- and SPECC- based algorithms. The algorithms are applied to a frequency hopping system similar to the one described in [11].

The transmission frequency was confined to the range $1899 - 1929\text{MHz}$, the uplink frequency range for the PCS system described in [11]. A total of 75 frequencies are considered to be the 'hop' frequencies in this range of 30MHz thus providing a frequency separation of 400kHz. The symbol period considered for the system is $T_s = 4\mu s$, the minimum required to achieve 500kbps using QPSK. The number of multipaths in all our simulations was assumed to be $P = 3$, but results tend to be similar for larger number of multipaths. This system allows a minimum Δ_f of 400kHz. Thus delays larger than $2.5\mu s$ require additional processing for estimation, one possible technique was outlined in Section 3. We consider here the harder problem of separation of delays less than the symbol period and assume the multipath delays to be $0.18\mu s$, $0.43\mu s$ and $0.90\mu s$ respectively in all our simulations.

Frequency is hopped from packet to packet with each packet consisting of the same header symbols and varying payloads. For our purposes, we generate a set of K random symbols which are used as header symbols in all packets. The header parts of the packets are thus identical but generated randomly for every Monte Carlo run. Since we never operate on the payload part of the packets, we never actually generate the payloads. Thus each packet consists of K QPSK symbols transmitted at a hop frequency.

The algorithm requires culling of packets with frequencies sharing a certain structure. This aspect of the algorithm was studied in Sections 3.2 and 4.2, and we re-emphasize the significance of this step in the algorithm. At this moment, however, we are interested in studying the performance of the algorithm provided that the frequency structure is readily available. With this in mind, the hop frequencies of the packets are assigned definite values rather than assigning random values in the given range. In particular, we assume a total of N packets transmitted at frequencies that form an arithmetic progression, a structure that immediately leads to the "linear array" [7] formulation of the ESPRIT algorithm.

The frequency structure described above is far from reality but extracts the best performance from the ESPRIT algorithm. The SPECC algorithm and the MI-ESPRIT algorithm were presented as an alternative to this highly constrained structure. For purposes of fair comparison, however, we continue to assume the same "linear-array" structure of the frequencies for the SPECC and the MI-ESPRIT algorithm as well. Our simulations assumed a total of $N = 20$ packets with "linear-array" frequency structure. If the packet frequencies are labeled with the set of integers $\{1, 2, 3, \cdots, N\}$, the frequency structure exploited in the three algorithms are as follows:

- ESPRIT: $\{1, 2, \cdots, 19\}$, $\{2, 3, \cdots, 20\}$

- SPECC: $\{1 - 5\}$, $\{6 - 10\}$, $\{11 - 15\}$, $\{16 - 20\}$

- MI-ESPRIT: $\{1 - 5\}$, $\{6 - 10\}$, $\{11 - 15\}$, $\{16 - 20\}$

The SPECC algorithm requires a minimum of 4 groupings to estimate $P = 3$ multipath delays and hence the above grouping. Note that this grouping corresponds to $\{k_0 = 0, k_1 = 1, k_2 = 2, k_3 = 3\}$. MI-ESPRIT can incorporate a more general grouping but we assume the same grouping as assumed for SPECC for comparison purposes. The delays are expressed as fractions of the symbol period and the mean squared error calculated for this fraction is referred to as the 'normalized' MSE. Performance of the three algorithms is compared by plotting normalized MSE against SNR and against the number of header symbols. SNR is measured at the receiver front end as the signal to noise ratio where the signal constitutes the sum of all multipaths. MSE calculation is accomplished by running a total of 500 Monte Carlo runs.

Before we proceed with comparison of results, we note that the SPECC and ESPRIT based algorithms are the only viable options for channel delay estimation. The MI-ESPRIT

algorithm involves a tedious multi-dimensional search and is displayed on the plots only as a benchmark; it is not a practical on-line algorithm. Note also, that in each case, we confine our comparison to one of three delays involved. Performance tends to be similar for typical values of the delays.

5.2 Results and Discussions

A comparison of the behavior of the normalized mean square against SNR for the three algorithms is shown in Fig. 2. The SPECC algorithm clearly outperforms the TLS-ESPRIT algorithm in estimating the delay parameters. The MSE achieved by the SPECC algorithm is considerably lower than that of the ESPRIT algorithm at low SNRs. The two MSEs approach each other at high SNRs. The MI-ESPRIT algorithm, with optimal exploitation of the frequency structure performs the best as expected. However, the SPECC algorithm performs the best at low SNRs and appears to be a worthy alternative to ESPRIT for low power applications. At low SNRs, SPECC offers a 5 to 7 dB advantage over TLS (which has comparable computational complexity) and MI-ESPRIT (whose computational complexity is extremely high).

Fig. 3 is a comparison plot between the three algorithms with varying number of header symbols. Since the number of signal samples available is limited by the number of header symbols, signal space estimation tends to be inaccurate. It is clear from the plots that the SPECC algorithm is more robust against inaccurate subspace estimation when compared to the ESPRIT algorithm. Applications that employ the ESPRIT scheme require a sufficient number of samples to be available for a good estimation of the signal subspace. The subspace estimation is further hindered by the presence of noise. The SPECC algorithm exploits a linear constraint on the subspaces as opposed to the rotational invariance exploited by

the ESPRIT algorithm and hence is better suited to handle inaccurate estimates of signal subspaces. Detailed performance analyses of the suggested approaches are issues of current interest.

The plot of the MSE against SNR for a more general case is shown in Fig. 4. A total of $N = 30$ packets was assumed available at the receiver. The packet hop frequencies continue to have a uniform linear arrangement. Labeling the packet hop frequencies as $\{1, 2, \cdots, 30\}$, the frequency partition considered for the two algorithms are

- ESPRIT: $\{1, 2, \cdots, 19\}$, $\{2, 3, \cdots, 20\}$

- SPECC: $\{1 - 5\}$, $\{11 - 15\}$, $\{16 - 20\}$, $\{21 - 25\}$,

- MI-ESPRIT: $\{1 - 5\}$, $\{11 - 15\}$, $\{16 - 20\}$, $\{21 - 25\}$,

The total number of packets used by SPECC is still 20 and hence the ESPRIT scheme was made available the same number of packets. Note that this partition corresponds to $M = P$ and $k_0 = 0, k_1 = 2, k_2 = 3, k_3 = 4$. The performance is similar to the basic SPECC algorithm and provides better estimation than the ESPRIT algorithm especially at low SNRs.

Fig. 5 is a comparison of the MSEs with the following frequency partition.

- ESPRIT: $\{1, 2, \cdots, 19\}$, $\{2, 3, \cdots, 20\}$

- SPECC: $\{1 - 5\}$, $\{11 - 15\}$, $\{21 - 25\}$, $\{26 - 30\}$,

- MI-ESPRIT: $\{1 - 5\}$, $\{11 - 15\}$, $\{21 - 25\}$, $\{26 - 30\}$,

Notice again that the three algorithms are provided with the same number of packets. The SPECC partition corresponds to $M = P$ and $k_0 = 0, k_1 = 2, k_2 = 4, k_3 = 5$. A significant degradation in the performance of the SPECC algorithm is observed. While the SPECC

algorithm, in theory, can exploit a general structure, performance degrades as the value of k_i strays away from the optimal value. Note that $k_i = i$ is optimal in the sense that the resulting polynomial has the same number of roots as the number of parameters estimated. For $k_i > i$, we have more roots than required and hence the degradation in the performance.

A crucial factor to be considered for ESPRIT estimation is the availability of at least as many packets in each data set as the number of parameters to be estimated. The estimation is substantially better as the number of packets available increases. Fig. 6 displays MSE vs. number of packets. The assumption of uniform linearity among the hop frequencies of the packets available makes it possible for us to have all but one packet in each data set. It is clear from the plot that most of the gain in the performance is achieved at a certain optimal number of packets and any increase in the number of packets beyond this does not add much to the performance.

6   Conclusions

In this paper, we have considered the multipath delay estimation problem for a slow frequency hopping system. The key idea was to exploit the invariant structures in the data packet. Two approaches were proposed; neither requires training bits; both lead to closed-form estimates; which can be used to initialize high-complexity optimization algorithms, such as MI-ESPRIT The use of multiple invariances can significantly improve the performance; thus the SPECC algorithm has the advantage. Furthermore, the SPECC algorithm also has shorter acquisition time than the ESPRIT approach. We also demonstrated that the SPECC algorithm is more robust to inaccurate estimation of the signal subspaces that results as a consequence of limitations in the number of header symbols available for estimation. One limitation of the proposed algorithm is that the multipath delay is limited to one symbol

interval due to the ambiguity of the phase-wrapping.

References

[1] R.L. Pickholtz, D.L. Schilling, and L.B. Milstein. "Theory of Spread spectrum Communication - a Tutorial". *IEEE Trans. Commun.*, COM-30:855–884, 1982.

[2] M.K. Simon, J.K. Omura, R.A. Scholtz, and B.K. Levitt. *Spread Spectrum Communication Handbook*. McGraw-Hill Inc, New York, 1994. Revised Edition.

[3] Y.A. Chau and J.-K. Wang. "Spectral Estimation based acquisition for frequency hopping spread spectrum communications in a nonfading or Rayleigh fading channel". *IEEE Trans. Communications*, 45(4):445–55, April 1997.

[4] P.-T. Sun and C.-Y. Chu. "Hidden preamble detector for acquisition of frequency hopping multiple-access communication system". *IEE Proceedings-Communications*, 144(3):161–5, June 1997.

[5] L. Aydm and A. Polydoros. "Joint hop-timing estimation for FH signals using a coarsely channelized receiver". In *Proc. MILCOM 95*, pp. 769–73, San Diego, CA, November 1995.

[6] P.H. Hande, L. Tong, A. Swami. "Channel Estimation for Frequency Hopping Systems" In *MILCOM 99*, Vol. 2, pp. 1323 - 1327, November 1999.

[7] R. Roy, A. Paulraj, and T. Kailath. "ESPRIT - a subspace rotation approach to estimation of parameters of cisoids in noise". *IEEE Trans. Acoust. Speech, Signal Proc.*, ASSP-34(10):1340–1342, October 1986.

[8] A.L. Swindlehurst, B. Ottersten, R. Roy, T. Kailath. "Multiple invariance ESPRIT". *IEEE Trans. on Signal Processing*, 40(4):867–881, April 1992.

[9] P.H. Hande and L. Tong "Signal Parameter Estimation via the Cayley-Hamilton Theorem" *to appear in IEEE Signal Processing Letters*, April, 2001.

[10] P.H. Hande, L. Tong and A. Swami "Channel estimation for frequency hopping systems via multiple invariances" *Proc. IEEE Wireless Comm. and Networking Conf.,* Sept. 2000.

[11] P.D. Rasky, G.M. Chiasson, D.E. Borth and R.L. Peterson. "Slow frequency-hop TDMA/CDMA for macrocellular personal communications". *IEEE Personal Communications*, Vol. 1, pp. 26–35, 2nd Quarter, 1994.

[12] Bluetooth Special Interest Group. "The Bluetooth Specification". `http://www.bluetooth.com`.

[13] J.C. Haartsen, "The Bluetooth radio system", *IEEE Personal Communications*, Vol 7, 28–36, Feb 2000.

[14] Lang Tong and S. Perreau. "Multichannel blind identification: from subspace to maximum likelihood methods". *Proceedings of the IEEE*, Vol. 86:1951 - 1968, October 1998.

[15] A. Paulraj, R. Roy and T. Kailath. "Estimation Of Signal Parameters Via Rotational Invariance Techniques-ESPRIT". *Asilomar Conference on Circuits, Systems and Computers*, pp.83–89, November 1985.

[16] R. Roy and T. Kailath. "ESPRIT - Estimation of signal parameters via rotational invariance techniques". *IEEE Trans. Acoust. Speech, Signal Proc.*, ASSP-37(7):984–995, July 1989

[17] B. Ottersten, M. Viberg, T. Kailath. "Performance analysis of the total least squares ESPRIT algorithm". *IEEE Trans. on Signal Processing*, SP-39:1122–1135, May 1991.

[18] G.H. Golub and C.F. Van Loan. *Matrix Computations.* Johns Hopkins University Press, Baltimore,MD, 1983.

[19] K.T. Wong and M.D. Zoltowski. "Closed-form multi-dimensional multi-invariance ES-PRIT". *IEEE International Conference on Acoustics, Speech, and Signal Processing, 1997. ICASSP-97*, vol. 5:3489 - 3492.

[20] P. Lancaster and M. Tismenetsky. *The Theory of Matrices.* Academic Press, Inc. 1985
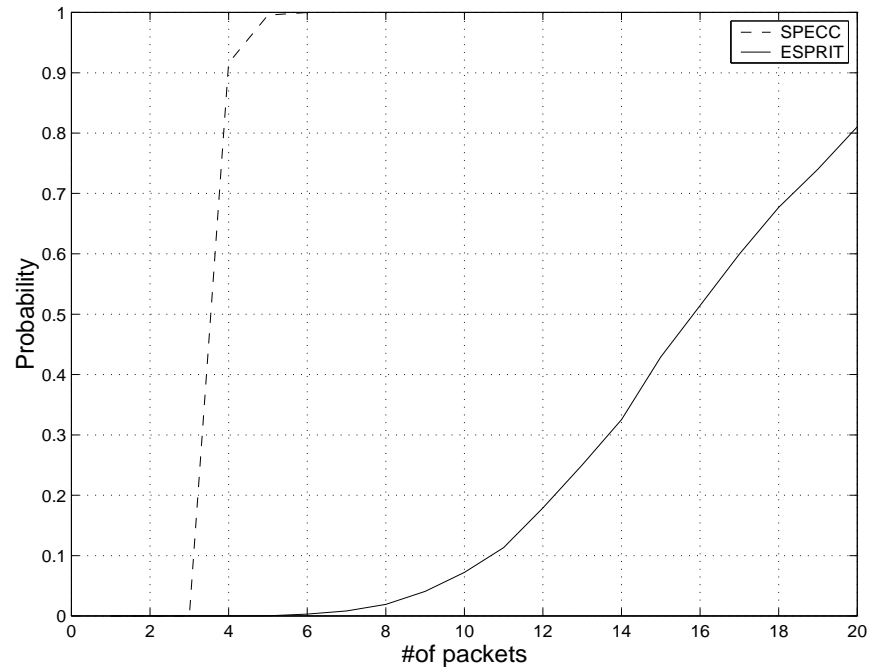
Figure 1: *Comparison of the Probabilities of the two events that lead to Channel estimation*
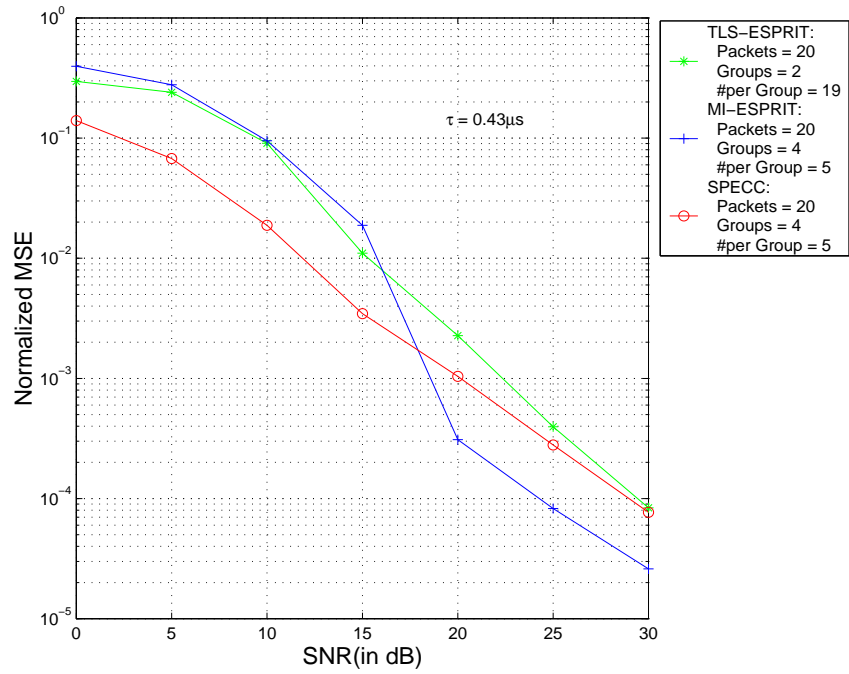
Figure 2: *Comparison of the MSE for three algorithms, Delays:* $\tau = [0.18 \; 0.43 \; 0.90]\mu s$, $\{k_0 = 0, k_1 = 1, k_2 = 2, k_3 = 3\}$
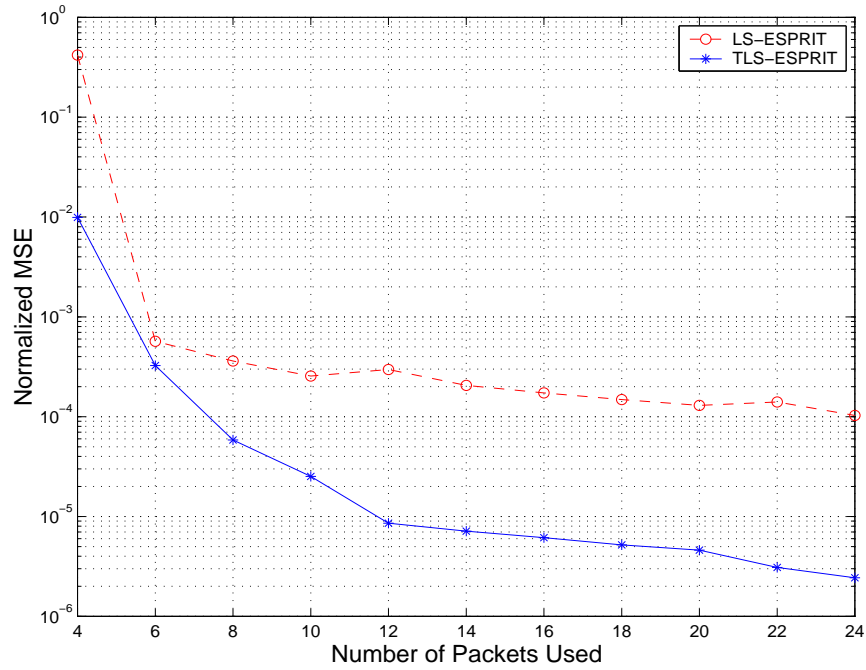
Figure 3: *Normalized MSE plot for SPECC and TLS-ESPRIT against number of header symbols , Delays: $\tau = [\ 0.18\ 0.43\ 0.90]\mu s$, SNR=20dB, $\{k_0 = 0, k_1 = 1, k_2 = 2, k_3 = 3\}$*
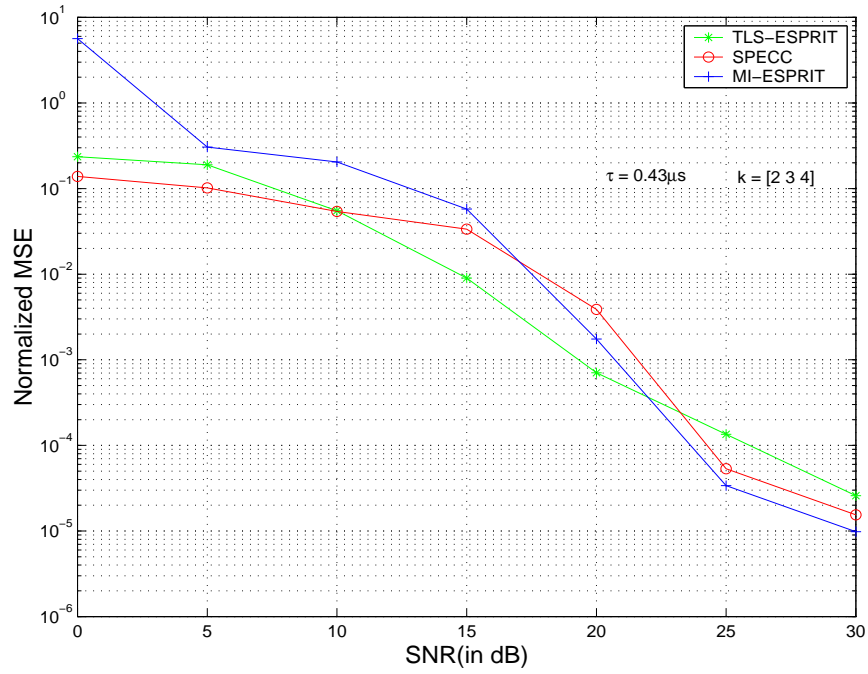
Figure 4: *Comparison of the MSE for the three Algorithms, Delays:* $\tau =$ $[0.18\ 0.43\ 0.90]\mu s, k_0 = 0, k_1 = 2, k_2 = 3, k_3 = 4$
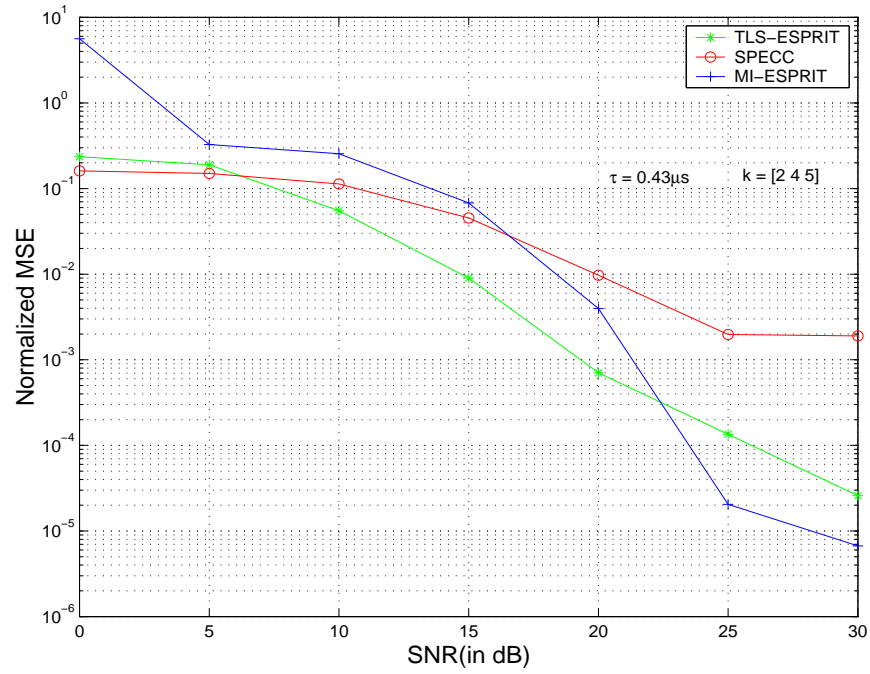
Figure 5: *Comparison of the MSE for the three Algorithms, Delays:* $\tau =$ $[0.18\ 0.43\ 0.90]\mu s, k_0 = 0, k_1 = 2, k_2 = 4, k_3 = 5$
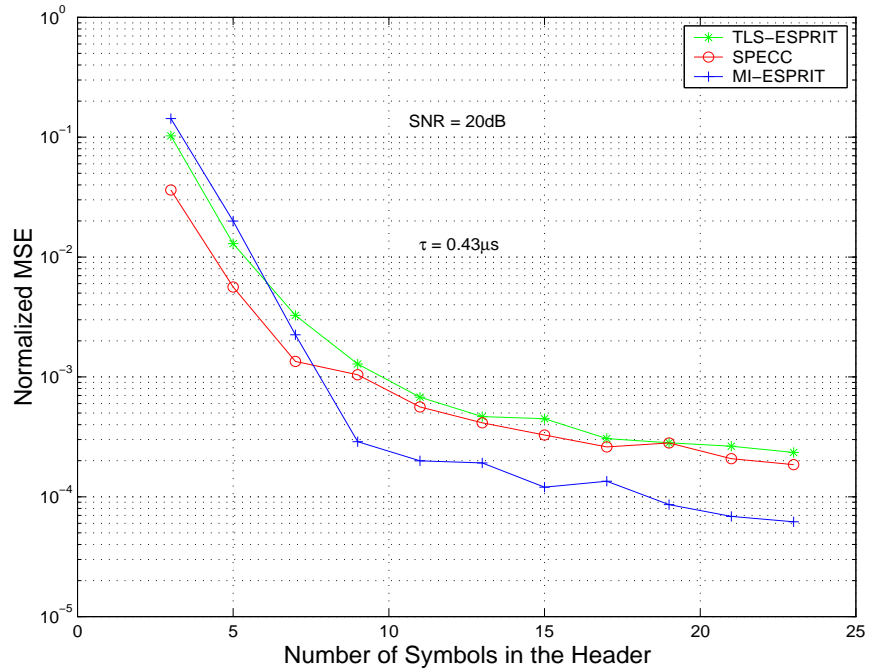
Figure 6: *Normalized MSE plot for LS-ESPRIT and TLS-ESPRIT against number of packets, Delays: $\tau = [0.18\ 0.43\ 0.90]\mu s$, SNR=30dB.*