PARAMETRIC ESTIMATION OF HARMONICALLY RELATED SINUSOIDS

A Thesis

by

RICHA DIXIT

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

May 2010

Major Subject: Electrical Engineering

PARAMETRIC ESTIMATION OF HARMONICALLY RELATED SINUSOIDS

A Thesis

by

RICHA DIXIT

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Approved by:

Chair of Committee,	Tie Liu
Committee Members,	Jean-Francois Chamberland-Tremblay
	Srinivas Shakkottai
	Donald K. Friesen
Head of Department,	Costas N. Georghiades

May 2010

Major Subject: Electrical Engineering

ABSTRACT

Parametric Estimation of Harmonically Related Sinusoids. (May 2010)

Richa Dixit, B. Tech., Institute of Technology, Banaras Hindu University, India

Chair of Advisory Committee: Dr. Tie Liu

Mud-pulse telemetry is a method used for measurement-while-drilling (MWD) in the oil industry. The telemetry signals are corrupted by spurious mud pump noise consisting of a large number of harmonically related sinusoids. In order to denoise the signal, the noise parameters have to be tracked accurately in real time. There are well established parametric estimation techniques for determining various parameters of independent sinusoids. The iterative methods based on the linear prediction properties of the sinusoids provide a computationally efficient way of solving the non linear optimization problem presented by these methods. However, owing to the large number of these sinusoids, incorporating the harmonic relationship in the problem becomes important.

This thesis is aimed at solving the problem of estimating parameters of harmonically related sinusoids. We examine the efficacy of IQML algorithm in estimating the parameters of the telemetry signal for varying SNRs and data lengths. The IQML algorithm proves quite robust and successfully tracks both stationary and slowly varying frequency signals. Later, we propose an algorithm for fundamental frequency estimation which relies on the initial harmonic frequency estimate. The results of tests performed on synthetic data that imitates real field data are presented. The analysis of the simulation results shows that the proposed method manages to remove noise causing sinusoids in the telemetry signal to a great extent. The low computational complexity of the algorithm also makes for an easy implementation on field where computational power is limited.

To My Master, whose love and grace makes everything possible

ACKNOWLEDGMENTS

I would like to thank my advisor, Dr. Tie Liu, for accepting me to be a part of his research team. As a mentor and critic, he guided me in the right direction throughout the course of this research. He motivated me to challenge myself and ask the right questions. Like every member of his team I cannot thank him enough for his unceasing support on both professional and personal front. I would also like to extend special thanks to Dr. Jean-Francois Chamberland-Tremblay whose courses on linear systems analysis and estimation and detection theory laid foundation of my research. I am grateful to my committee members Dr. Donald K. Friesen and Dr. Srinivas Shakkottai for lending their help whenever I needed it. I am indebted to my fellow graduate students, all the professors and staff of the ECE Department at Texas A&M University who made this journey more fun and meaningful.

I owe my deepest gratitude to my mother, who encouraged me to follow my dreams and lent her support every step of the way. I am also grateful to my younger brother for always believing in my abilities. Most of all I would like to thank my husband, Ketan, for being a constant source of love and support. Not only was he a dutiful friend, but he also gave me invaluable professional advice from time to time. Finally, I would like to thank my late father whose memories and ideals continue to inspire me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

A. Motivation and Background

The problem of estimating parameters of multiple sinusoids buried in noise has been of great interest to the signal processing community for many years. The specific applications include time series analysis and system identification and antennae array processing. However, little attention has been paid to the special case of harmonic sinusoids. The harmonic frequency estimation has important applications in speech signal processing, automotive control systems as well as instrumentation and measurement. This thesis was motivated by a practical problem of noise cancelation while performing measurements, where parameter estimation of harmonically related sinusoids was required. A greater part of this section is devoted to describing the problem background with emphasis on the practical challenges which dictate the selection of an estimation procedure.

1. Mud Pump Noise Cancelation

Mud-pulse telemetry is a method used for measurement-while-drilling (MWD) in the oil industry. MWD systems provide drilling operators great control over the construction of a well by providing information about conditions at the bottom of a wellbore in real time as the wellbore is being drilled. The information includes directional drilling variables such as inclination and direction (azimuth) of the drill bit, and geographical formation data such as natural gamma radiation levels and electrical resistivity of the rock formation. MWD systems measure parameters(such

---

The journal model is *IEEE Transactions on Automatic Control.*

as the previously mentioned examples) and transmit the acquired data to the earth's surface from within the wellbore. Fig. 1 shows a typical mud-pulse telemetry system.

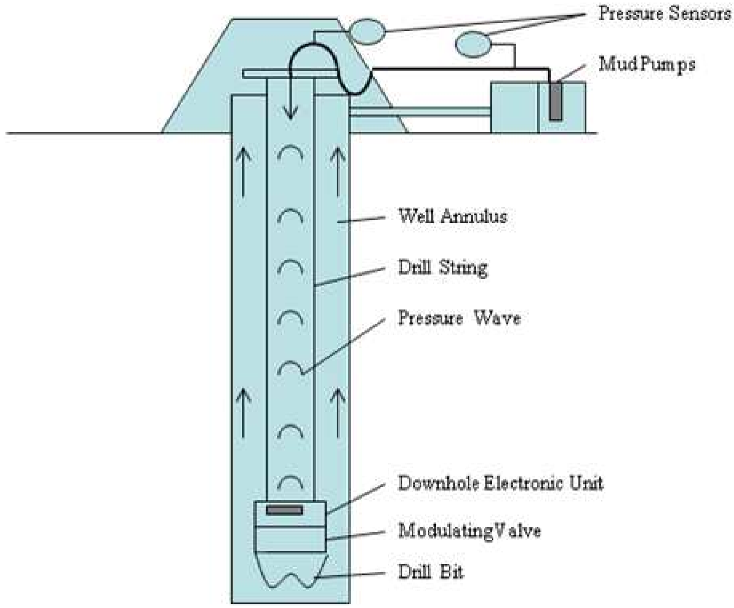


Fig. 1. A typical mud-pulse telemetry system

One of the methods used to send MWD data to the surface is by means of pressure waves in the drilling fluid("Drilling mud") that is pumped through the drill string by pumps on the surface. The downhole electronics, control a valve that is in the stream of mud. A modulated pressure wave is generated by controlling the interruption of the valve with the drilling mud, thus causing the pressure in the drill string to change in a manner related to the downhole measurement data. The pressure waves generated by the valve travel through the drilling mud to the surface, where they are measured by one or more pressure sensors. The outputs of the pressure sensors can be digitized in analog-to-digital converters and processed by a signal processing module, which recovers the information data from the pressure variations and then sends the data to a computer. The transmitted information is then assessed by the drilling operator.

Other methods for communicating MWD data to the surface exist, such as acoustic wave propagation through the drill string, electromagnetic radiation through the ground formation, and electrical transmission through an insulated cable. However, in terms of the overall cost and effectiveness, mud-pulse telemetry remains the most commercially viable method as of today.

Mud-pulse telemetry is confined within a very low frequency band(below 50 Hz, depending on the characteristics of the drilling mud being used and the drilling depth), due to the attenuation mechanism of the drilling mud. At such low frequencies, the pressure wave that carries information from the transmitter to the pressure sensors is subjected to severe noise corruptions. On most occasions, the higher power pressure wave from the surface mud pumps contributes the most significant amounts of noise. The mud pump noise is mainly the result of reciprocating motion of mud pump pistons and hence is harmonic in nature. The pressure waves from the surface mud pumps travel in the opposite direction from the main information carrying wave, and the pressure sensors detect pressure variations representative of the sum of signal and noise waves. In case that the harmonic components of the noise are present within the frequency range used for transmission of the telemetry wave, severe distortions are introduced to the received signal making correct detection of the transmitted data extremely difficult.

In the past, this problem was circumvented by transmitting only over a small portion of the entire spectrum available, so that no more than a couple of mud pump harmonics are present within the telemetry band. Such a passive approach reduces interfering effects of the mud pump noise at the expense of under utilizing the bandwidth. As a result, the data rate of communication is severely limited, and the drilling operator must be selective about what data are transmitted. This lack of information makes its difficult to optimize the drilling wells. Telemetry is a major limiting factor for the application of the MWD technology.

2. Objective and Outline

The goal of this thesis is to develop a robust and efficient algorithm which can estimate the parameters of multiple mud pump harmonics. The knowledge of noise harmonic parameters can then be suitably used to remove them from the received pressure waves, thus making the entire spectrum available for data communication. At the potential rates that are supported by the use of the entire spectrum, the extra information available to drilling operator not only can reduce the total drilling cost of a typical well by a substantial amount, but can also turn many currently marginal gas reservoir into commercial targets.

The rest of this thesis is organized as follows. In the next section the estimation problem is mathematically formulated and a brief overview of the popular methods to solve such problem is presented. In Chapter II, an iterative parameter estimation algorithm is introduced which is not only computationally efficient, but also apt in incorporating the prior knowledge of the unknown parameters. The *linear prediction* properties of sinusoids are used here to give an accurate estimate of pump noise frequencies. In Chapter III, the goal is to exploit harmonic relationship of sinusoids to improve estimation procedure. Here a novel method for the detection of the fundamental frequencies of the noisy sinusoids is presented, which builds on the harmonic estimate obtained previously. The knowledge of fundamental frequencies drastically reduces the size of the set of unknown parameters, thereby making it possible to eliminate a large number of noise causing signals. The details of both the estimation and detection methods, the algorithm, and the software implementation are given in the respective sections along with the simulation results obtained on the set of synthetic data.

B. Parameter Estimation

1. Problem Statement

Mathematically, the received pressure wave at the j^{th} pressure sensor can be written as:

$$y_j[m] = \sum_{l=0}^{L-1} h_{l,j} x[m-l] + \sum_{k=0}^{K/2-1} A_{k,j} \cos(\omega_k m + \phi_{k,j}) + w_j[m], \quad m = 0, \ldots, M-1, \quad j = 0, \ldots, J-1 \tag{1.1}$$

where, $\sum_{l=0}^{L-1} h_{l,j} x[m-l]$ represents the information signal, $\sum_{k=0}^{K/2-1} A_{k,j} \cos(\omega_k m + \phi_{l,k})$ represents the mud pump noise which is the sum of $K/2$ real sinusoids with frequencies ω_k, amplitudes $A_{k,j}$ and phases $\phi_{l,k}$ and $w_j[m]$ represents the noise from the other sources. The exact structure of the information carrying wave is not important in this problem formulation.

In complex notation the pump noise can be rewritten as:

$$z_j[m] = \sum_{k=0}^{K-1} s_{k,j} \lambda_k^m \tag{1.2}$$

where, the mud pump noise $z_j[m]$ is a sum of K complex exponentials with complex frequencies λ_k and complex amplitudes $s_{k,j}$.

Here,

$$\lambda_{K-k}^* = \lambda_k = \exp\{j\omega_k\} \tag{1.3}$$

and

$$s_{K-k,j}^* = s_{k,j} \tag{1.4}$$

where K is an even number. The frequency parameters $\{\omega_k\}$ form a union of harmonic sets, each corresponding to the harmonic components originated from one pump. The stroke rate of each pump may drift over time due to the mechanical instability. The key issue for effective mud pump noise cancelation is to design an estimation algorithm which can faithfully track the frequency parameters.

The set of unknown parameters here consists of the amplitude, phase and frequency of the sinusoids i.e. $\theta = \{A_{k,j}, \phi_{k,j}, \omega_k\}$. As the number of sinusoids grow, this set becomes larger. Most of the estimation techniques hinge on first finding a correct estimate of frequency. The other linear parameters are then determined by using separable regression techniques described in section 2. There are numerous frequency estimation schemes, but they can be broadly put into two categories: non-parametric and parametric techniques. The non-parametric estimators , including the periodogram and correlogram methods, are based directly on the Fourier transform. Although no assumptions are made about the observed data sequence, the resolution, or ability to resolve closely spaced frequencies using non-parameteric approach is fundamentally limited by the length of the data available. Alternatively, the parametric approach assumes a signal model with known functional form and gives better resolution. The maximum likelihood, non-linear least squares, Prony's method, iterative filtering and subspace methods like MUSIC are examples of parametric estimation. A detailed comparative analysis of the two types of frequency estimators can be found in [1]. The mud pump frequency estimation involves resolving closely spaced frequencies for a wide range of SNR, so in this thesis the focus will be on parametric estimation, particularly the methods that utilize the linear prediction property of sinusoidal signals.

In order to find a reliable parametric estimator for a given class of problems the simplest and the most straight forward technique is to look for optimal estimators. They comprise of a class of unbiased estimators exhibiting minimum variance, the so-called MVU estimator. There are well established methods based on determination of the sufficient statistics and the Cramer Rao Lower Bound (CRLB) on the variance of such estimators. In cases where such estimators do not exist the popular approach is to find practical estimators. The maximum likelihood estimator(MLE), defined as the value of θ that maximizes the likelihood function, gives an asymptotically optimal solution. The MLE has a Gaussian PDF and turns out to be unbiased while achieving CRLB for large data records. For certain signal in noise problems the MLE achieves CRLB at high SNRs. However, finding an exact MLE becomes tedious in cases, as in current problem, where a large set of parameters are to be determined with an unknown underlying noise structure. The numerical approaches like grid search (Newton Raphson method) or iterative maximization (EM algorithm) simplify the problem, when a closed form expression cannot be found, but they do not guarantee convergence to the MLE.

Another approach is to use an estimator such as least square estimator(LSE) which does not have any optimality properties but still gives good performance in many cases. The LSE chooses θ to minimize the least square error criterion while making no probabilistic assumptions on the data. It is usually applied in situations where a precise statistical characterization of the data is unknown or where an optimal estimator cannot be found or may be too complicated to apply in practice.

2. Nonlinear Least Squares

A signal model that is non linear in the unknown parameter θ is said to generate a non linear least squares problem. Let $\mathbf{x}(\theta)$ is the signal model for observation $\mathbf{y}$ [2]. The LS procedure estimates model parameters θ by minimizing:

$$J = (\mathbf{y} - \mathbf{x}(\theta))^{\mathbf{T}}(\mathbf{y} - \mathbf{x}(\theta)) \tag{1.5}$$

If $(\mathbf{y}-\mathbf{x}(\theta)) \sim \mathcal{N}(\mathbf{0}, \sigma^{\mathbf{2}}\mathbf{I})$, the LSE is also the MLE. When $\mathbf{x}(\theta)$ is N-dimensional non linear function of θ the minimization of J becomes difficult. The first step to solving this problem is to reduce its complexity. Two most common ways of doing that are: 1. Transformation of parameters 2. Separability of parameters

In the first case a one-to one transformation of θ is made, that produces a linear signal model in the new space. If an invertible function g can be found such that if $\alpha = \mathbf{g}(\theta)$, then

$$\mathbf{x}(\theta(\alpha)) = \mathbf{x}(\mathbf{g}^{-\mathbf{1}}(\alpha)) = \mathbf{H}\alpha \tag{1.6}$$

The signal model then becomes linear in α and its linear LSE is given by pseudo-inverse:

$$\hat{\alpha} = (\mathbf{H^T H})^{-\mathbf{1}}\mathbf{H^T y} \tag{1.7}$$

In second case the signal model is non linear but it is linear in some of the parameters. Generally, a separable signal model has the form:

$$\mathbf{x} = \mathbf{H}(\alpha)\beta \tag{1.8}$$

where,

$$\theta = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \tag{1.9}$$

and $\mathbf{H}(\alpha)$ is a matrix dependent on α. This model is linear in β but non-linear in α. As a result, the LS error may be minimized with respect to $\beta$and thus reduced to a function of α only. Since,

$$\mathbf{J}(\alpha, \beta) = (\mathbf{y} - \mathbf{H}(\alpha)\beta)^{\mathbf{T}}(\mathbf{y} - \mathbf{H}(\alpha)\beta) \tag{1.10}$$

the β that minimizes J for a given α is:

$$\hat{\beta} = (\mathbf{H}(\alpha)^{\mathbf{T}}\mathbf{H}(\alpha))^{-\mathbf{1}}\mathbf{H}(\alpha)^{\mathbf{T}}\mathbf{y} \tag{1.11}$$

and the resulting LS error is:

$$\mathbf{J}(\alpha, \hat{\beta}) = \mathbf{x^T}[\mathbf{I} - \mathbf{H}(\alpha)(\mathbf{H}(\alpha)^\mathbf{T}\mathbf{H}(\alpha))^{-\mathbf{1}}\mathbf{H}(\alpha)^\mathbf{T}]\mathbf{x} \tag{1.12}$$

The problem now reduces to maximization of:

$$\mathbf{x^T}\mathbf{H}(\alpha)(\mathbf{H}(\alpha)^\mathbf{T}\mathbf{H}(\alpha))^{-\mathbf{1}}\mathbf{H}(\alpha)^\mathbf{T}\mathbf{x} \tag{1.13}$$

over α.

Minimizing 1.13 could still be a difficult non-linear optimization problem which is known to be plagued by local minima in the error surface. So instead of solving this non-linear problem, many methods were developed to solve the deterministic linear prediction problem, which is related to, but not the same as, the true least-squares optimization. The following section introduces the basic concept of linear prediction which is rigourously used in Chapter II to develop a solution based on the Iterative Quadratic Maximum Likelihood (IQML)estimation approach.

3. Linear Prediction

Resolving closely spaced sinusoids in the presence of noise can be challenging particularly when number of data is small and the SNR is low. Various methods based on the linear prediction approach have been developed to solve this problem including Prony's method, autoregressive (AR) and autoregressive moving average modeling (ARMA), Pisarenko method and principle eigenvector (PE) approach.

The basic idea behind linear prediction (originally introduced by Prony) is that, instead of determining the unknown frequencies directly, the coefficients of a prediction polynomial are estimated, the roots of which are the exponentials with those frequencies. Consider the exponential function:

$$y[m] = \sum_{k=0}^{K-1} s_k e^{j\omega_k m}, m = 0, \ldots, M-1 \tag{1.14}$$

where, s_k and ω_k are the complex amplitude and frequency of the k-th sinusoid, respectively. There are $2K$ unknowns so it will need atleast $2K$ independent observation samples to determine these values. Here $M \geq 2K$. Viewing 1.14 as a linear system in s_k's, a Vandermonde matrix can be observed. The coefficients s_k's can be separated from the exponentials e^{ω_k}'s by the introduction of auxiliary polynomial:

$$b(z) = (z - e^{j\omega_0})(z - e^{j\omega_1}) \ldots (z - e^{j\omega_(k-1)}) = \sum_{k=0}^{K} b_k z^{K-k} \tag{1.15}$$

If the coefficients of b are known then $e^{j\omega_k}$'s are found by computing its roots. In order to find b_k's, a Hankel system of equation is developed as follows. $b(e^{j\omega_k}) = 0$ for all values of k, so:

$$\sum_{k=0}^{K-1} s_k b(e^{j\omega_k}) = 0 \tag{1.16}$$

Simplifying above equation and collecting terms in b_k's:

$$\begin{aligned} \sum_{k=0}^{K-1} s_k \sum_{i=0}^{K} b_i e^{j\omega_k K-i} &= \sum_{i=0}^{K} b_i \sum_{k=0}^{K-1} s_k e^{j\omega_k K-i} \\ &= \sum_{i=0}^{K} b_i y[K-i] = 0 \end{aligned}$$

$$\sum_{i=1}^{K} b_i y[K-i] = -b_0 y[K] \tag{1.17}$$

The following sets of equation can similarly be defined:

$$\sum_{k=0}^{K-1} s_k e^{j\omega_k p} b(e^{j\omega_k}) = 0, p = 1, \ldots, M-K-1 \tag{1.18}$$

A set of linear equations is obtained by simplifying 1.18 to the form expressed in 1.17.

Upon combining, the problem can be written as:

$$\begin{pmatrix} y[0] & y[1] & \cdots & y[K-1] \\ y[1] & y[2] & \cdots & y[K] \\ \vdots & \vdots & & \vdots \\ y[M-K-1] & y[M-K] & \cdots & y[M-2] \end{pmatrix} \begin{pmatrix} b_K \\ b_{K-1} \\ \vdots \\ b_1 \end{pmatrix}, = \begin{pmatrix} y[K] \\ y[K+1] \\ \vdots \\ y[M-1] \end{pmatrix}, \tag{1.19}$$

or

$$\mathbf{Ab} = \mathbf{y} \tag{1.20}$$

where $\mathbf{A}$ is linear prediction (LP) matrix, $\mathbf{b}$ is LP vector and $\mathbf{y}$ is observation vector. The order of LP vector is generally greater than K as the number of signals are over estimated in such procedures.

Unfortunately, Prony's method is well known to perform poorly when the signal is embedded in noise; Kahn et al (1992) show that it is actually inconsistent. On the other hand the Pisarenko method based on same LP property, though consistent proves inefficient for estimating sinusoidal signals and inconsistent for estimating damped sinusoids or exponential signals.

A modified Prony algorithm, extended to the least squares context, was proposed by Osborne (1975). It was generalized in Smyth (1985) and Osborne and Smyth (1991) [3] to estimate any function which satisfies a difference equation with coefficients linear and homogeneous in the parameters. Osborne and Smyth (1991) considered in detail the special case of rational function fitting, and proved that the algorithm is asymptotically stable in that case.

The next chapter describes in detail the Iterative Quadratic Maximum Likelihood (IQML) estimation procedure which is an iterative method based on LP property of sinusoids, for solving the NLS problem.

CHAPTER II

ITERATIVE QUADRATIC MAXIMUM LIKELIHOOD ESTIMATION

In this chapter the focus is develop an algorithm to reliably estimate the frequency parameters without directly exploiting the harmonic relationship among different sinusoidal components originated from the same pump. Treating different harmonic components from the same pump as sinusoids with independent frequencies is, a suboptimal strategy. However, algorithms based on this suboptimal strategy are much less demanding in terms of computational complexity and, will perform reasonably well in many engineering situations. A separate method which attempts to further improve the estimation performance by exploiting the harmonic relationship among different sinusoidal components originated from the same pump is described in Chapter III.

The method for estimating the pump noise frequencies is based on the maximum-likelihood estimation framework. Once reliable estimates of the frequency parameters are obtained, properly designed notch filters can be deployed to cancel the pump noise without overly damaging the telemetry signal.

A. The Deterministic Maximum Likelihood Problem

Iterative methods for solving the least squares, or maximum likelihood, spectral parameter estimation problem have been developed by Kumaresan et al. [4], and Bresler and Macovski [5]. These methods are collectively known as iterative quadratic maximum likelihood(IQML). The application of the IQML method for noise cancelation only requires the estimation of the time domain signal, which is a superposition of exponentially signals in noise. The model for the IQML method is the same as that for the NLS method described in the last chapter and is restated below:

$$y_j[m] = \sum_{k=0}^{K-1} s_{k,j}\lambda_k^m + n_j[m], \quad m = 0, \ldots, M-1, \quad j = 0, \ldots, J-1 \tag{2.1}$$

where, $n_j[m]$ represents the sum of the information-carrying wave and the noise from the other sources.

Let $\lambda = (\lambda_0, \ldots, \lambda_{K-1})^t$ and $\mathbf{s_j} = (s_{0,j}, \ldots, s_{K-1,j})^t$ be the collections of the complex frequency and amplitude parameters, respectively. The least-square estimate is given by:

$$\min_{\lambda,\mathbf{s_j}} \sum_{j=0}^{J-1} \sum_{m=0}^{M-1} \left[y_j[m] - \sum_{k=0}^{K-1} s_{k,j}\lambda_k^m \right]^2 . \tag{2.2}$$

Using the vector-matrix notation to compactly describe the model:

$$\mathbf{y_j} = (y_j[0], \ldots, y_j[M-1])^t \tag{2.3}$$

Here each vector $\mathbf{y_j}$ of length M given by a sum of $K < M$ exponential signal vectors corrupted by additive noise. Also, define a $M \times K$ Vandermonde matrix $\mathbf{A}(\lambda)$:

$$\mathbf{A}(\lambda) = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ \lambda_0 & \lambda_1 & \cdots & \lambda_{K-1} \\ \vdots & \vdots & & \vdots \\ \lambda_0^{M-1} & \lambda_1^{M-1} & \cdots & \lambda_{K-1}^{M-1} \end{pmatrix}, \tag{2.4}$$

The least-square estimation problem 2.2 can be written as

$$\min_{\lambda,\mathbf{s_j}} \sum_{j=0}^{J-1} \|\mathbf{y_j} - \mathbf{A}(\lambda)\mathbf{s_j}\|^2 . \tag{2.5}$$

For any given λ, the least-square estimate of $\mathbf{s_j}$ can be found using the separability argument given in Chapter II:

$$\hat{\mathbf{s}}_\mathbf{j} = \mathbf{A}^\dagger(\lambda)\mathbf{y_j} = [\mathbf{A}^*(\lambda)\mathbf{A}(\lambda)]^{-\mathbf{1}}\mathbf{A}^*(\lambda)\mathbf{y_j} \tag{2.6}$$

where $\mathbf{A}^{\dagger}(\lambda)$ is the pseudoinverse of $\mathbf{A}(\lambda)$ Thus, the problem of finding the least-square estimate of λ reduces to

$$\hat{\lambda} = \arg\min_{\lambda} \sum_{j=0}^{J-1} \|\mathbf{y_j} - \mathbf{A}^{\dagger}(\lambda)\mathbf{y_j}\|^2 = \arg\min_{\lambda} \mathrm{Tr}(\mathbf{P}_{\mathbf{A}}^{\perp}(\lambda)\hat{\mathbf{R}}_{\mathbf{y}}) \tag{2.7}$$

where $\mathbf{P}_{\mathbf{A}}^{\perp}(\lambda)$ is the projection matrix onto the null space of $\mathbf{A}(\lambda)$, given by:

$$\mathbf{P}_{\mathbf{A}}^{\perp}(\lambda) = \mathbf{I} - \mathbf{A}(\lambda)[\mathbf{A}^*(\lambda)\mathbf{A}(\lambda)]^{\mathbf{-1}}\mathbf{A}^*(\lambda) \tag{2.8}$$

and $\hat{\mathbf{R}}_{\mathbf{y}}$ is the measurement sample correlation given by:

$$\hat{\mathbf{R}}_{\mathbf{y}} = \sum_{j=0}^{J-1} \mathbf{y_j}\mathbf{y_j^*} \tag{2.9}$$

Consider the *linear prediction* polynomial:

$$b(z) = \sum_{k=0}^{K} b_k z^{K-k} = b_0 \prod_{k=0}^{K-1} (z - \lambda_k). \tag{2.10}$$

where the collection of the polynomial coefficients $\mathbf{b} = (b_0, \ldots, b_K)^t$ is a nonzero vector. Through some simple manipulations, it can be shown that the projection matrix

$$\mathbf{P}_{\mathbf{A}}^{\perp}(\lambda) = \mathbf{B}(\mathbf{B}^*\mathbf{B})^{\mathbf{-1}}\mathbf{B}^* \tag{2.11}$$

where

$$\mathbf{B} = \begin{pmatrix} b_K^* & & \\ \vdots & b_K^* & \\ b_0^* & \vdots & \\ & b_0^* & b_K^* \\ & & \vdots \\ & & b_0^* \end{pmatrix}. \tag{2.12}$$

Thus the least-square estimate of the linear prediction polynomial coefficients can be written as

$$\hat{\mathbf{b}} = \arg\min_{b_0=1} \operatorname{Tr}\left[\mathbf{B}(\mathbf{B}^*\mathbf{B})^{-1}\mathbf{B}^*\hat{\mathbf{R}}_{\mathbf{y}}\right]. \tag{2.13}$$

Once an estimate $\hat{\mathbf{b}}$ of the linear prediction polynomial coefficients is obtained, the complex frequency parameters λ can be obtained by finding the roots of the estimated linear prediction polynomial

$$\hat{b}(z) = \sum_{k=0}^{K} \hat{b}_k z^{K-k}. \tag{2.14}$$

B. The Algorithm

The least square estimation problem 2.13 is a very complex nonlinear optimization problem. In order to make it amenable for numerical solution, define:

$$\mathbf{Y_j} = \begin{pmatrix} y_j[k] & y_j[k-1] & \cdots & y_j[0] \\ y_j[k+1] & y_j[k] & \cdots & y_j[1] \\ \vdots & \vdots & & \vdots \\ y_j[M] & y_j[M-1] & \cdots & y_j[M-k] \end{pmatrix}. \tag{2.15}$$

Rewriting 2.13 as

$$\hat{\mathbf{b}} = \arg\min_{b_0=1} \left\{ \mathbf{b}^* \left[\sum_{j=0}^{J-1} \mathbf{Y_j^*}(\mathbf{B}^*\mathbf{B})^{-1}\mathbf{Y_j} \right] \mathbf{b} \right\}. \tag{2.16}$$

This form of the least-square estimation problem can be solved using the following iterative algorithm for estimating the complex frequency parameters λ. It only requires the solution of a quadratic minimization problem at each step and converges in a small number of steps (normally less than 5).

a) Initialization: $n = 0$ and $\mathbf{b}^{(0)} = \mathbf{b_0}$;

b) Calculate

$$\mathbf{B}^{(n)} = \begin{pmatrix} b_K^{(n)^*} & & \\ \vdots & b_K^{(n)^*} & \\ b_0^{(n)^*} & \vdots & \\ & b_0^{(n)^*} & b_K^{(n)^*} \\ & & \vdots \\ & & b_0^{(n)^*} \end{pmatrix}$$

and

$$\mathbf{C}_y^{(n)} = \sum_{j=0}^{J-1} \mathbf{Y}_\mathbf{j}^* (\mathbf{B}^{(\mathbf{n})^*} \mathbf{B}^{(\mathbf{n})})^{-1} \mathbf{Y}_\mathbf{j};$$

c) Solve

$$\mathbf{b}^{(n+1)} = \arg \min_{b_0=1} \{\mathbf{b}^* \mathbf{C}_y^{(n)} \mathbf{b}\}; \tag{2.17}$$

d) Check convergence: If $\|\mathbf{b}^{(n+1)} - \mathbf{b}^{(n)}\| \leq \varepsilon$, go to Step e); otherwise, go back to Step b);

e) $\hat{\lambda}$ is given by the roots of

$$\hat{b}(z) = \sum_{i=0}^{K} b_i^{(n)} z^{K-i} \tag{2.18}$$

Amplitudes and phases can be obtained by substituting the frequencies obtained above into 2.6.Without directly exploiting the harmonic relationship among different sinusoidal components originated from the same mud pump, the most computationally intensive procedure in each iteration is to solve a *quadratic* optimization problem in Step c).

1. Implementation of Constraints in Real Sinusoids

The described algorithm was implemented using Matlab (standard package). To find the inverse and eigenvectors of the matrices, the Matlab built-in function `inv` and `eig`, were used respectively. The initialization was chosen as

$$\mathbf{b}_0 = \arg\min_{b_0=1} \left\{ \mathbf{b}^* \left[\sum_{j=0}^{J-1} \mathbf{Y_j^* Y_j} \right] \mathbf{b} \right\}, \tag{2.19}$$

which is known as the *Prony* estimate in the literature. The vector $\mathbf{b}$ must be constrained to avoid the trivial all-zero solution. In the quadratic minimization problem 2.19, it is usually difficult to decide between imposing linear ($b_0 = 1$) or quadratic ($\|\mathbf{b}\| = 1$, $\|.\|$ where denotes the Euclidean norm) constraint. While the former constraint leads to an analytical solution, the latter requires a numerical solution for which good iterative computer algorithms exist today. The conditions for appropriate implementation of these constraints are discussed in [6]. In our experiments we use the linear non-triviality constraint $b_0 = 1$. The symmetric condition in 1.3 implies the symmetric condition for the linear prediction polynomial coefficients:

$$b_{K-k} = b_k. \tag{2.20}$$

With this symmetry, the size of the quadratic optimization in each iteration step is effectively reduced by a half. Specifically, by letting

$$\mathbf{b} = \mathbf{P}\tilde{\mathbf{b}} \tag{2.21}$$

where $\tilde{\mathbf{b}} = (b_0, b_1, \ldots, b_{K/2})^t$ and

$$\mathbf{P} = \begin{pmatrix} 1 & & & & 0 \\ & 1 & & & 0 \\ & & \ddots & & \vdots \\ & & & 1 & 0 \\ 0 & 0 & \cdots & 0 & 1 \\ & & & 1 & 0 \\ & & & & \vdots \\ & 1 & & & 0 \\ 1 & & & & 0 \end{pmatrix}, \tag{2.22}$$

the quadratic optimization problem in Step c) can be equivalently written as

$$\tilde{\mathbf{b}}^{(n+1)} = \arg \min_{\|\tilde{\mathbf{b}}\|=1} \{\tilde{\mathbf{b}}^* \tilde{\mathbf{C}}_{\mathbf{y}}^{(\mathbf{n})} \tilde{\mathbf{b}}\} \tag{2.23}$$

where $\tilde{\mathbf{C}}_y^{(n)} = \mathbf{P}^* \mathbf{C}_y^{(n)} \mathbf{P}$. In this case, it is well known that the solution is given by the *eigenvector* corresponding to the *minimum* eigenvalue of $\tilde{\mathbf{C}}_y^{(n)}$. The computational complexity of IQML is analyzed in [7]. The expensive matrix inversion operations here has been made simpler resulting in the reduction of the computational complexity of IQML.

2. Simulation Results and Discussion

The iterative aspect of IQML raises concerns about its convergence and consistency. These issues are addressed in detail in [8]. Our experiments show that IQML converges in a small number of steps (usually less than 10). The initial experiments on IQML in [5] showed that the MSE at SNRs higher than 3dB coincided with the CRLB. Below this threshold, the MSE increased rapidly due to algorithm converging to a local rather than a global minimum of the ML criterion. The theoretical asymptotic performance results for a large M obtained for IQML in [9] show that the MSE is different from CRLB and the difference between the two increase with M since the IQML estimates are almost always biased. Several results are presented in [10] which show that IQML is much more robust than TLS and LP methods over a broad range of conditions. However, due to the greater computational complexity of the IQML method [9], it can prove slower than the other techniques.

The IQML formulation also has the advantage that its frequency resolution does not depend on the absence of spectral smoothness but rather on the detection performance of the algorithm estimating the correct number of signals present. If the number of signals present are estimated correctly using any of the techniques outlined in [11], [12] or [13] then estimation accuracy is the only issue. Both detection and estimation accuracy will depend on the SNR and the specific method employed.

In the current problem, due to the presence of a large number of sinusoids it becomes difficult to directly estimate all of them. In fact the algorithm fails to converge in certain cases when $K \geq 6$. Thus, the algorithm is applied to only those out-of-band pressure waves that give the best estimation performance. The estimation result can then be mapped to rest of the bands through the harmonic relationship among different sinusoidal components originated from the same pump. This method has the advantage of estimating the frequencies of the harmonic components without suffering from the interference from the telemetry signal, but the mapping to the signal band might be a serious issue when the stroke rates of the mud pumps drift significantly over the time.

The simulations were performed on both the synthetic as well as real engineering data. Each frequency estimate is based on a data block of length 2 seconds and reinitialized across different data blocks. The results show that as long as the harmonic components are relatively strong when compared with the background noise, the proposed algorithm can faithfully track their frequencies. The algorithm also manages to resolve closely spaced frequencies (within a separation of ±0.005). The results presented here are obtained from synthetic data sets. The experiments were performed on two separate sets of synthetic data. The first set consists of the signals corrupted with noise from stationary pumps. The frequency of such pumps remains constant over a substantial period of time. The second set comprises of signal corrupted by chirp pump noise. The frequency of these pumps drifts with varying rates over the period of time. These simulations mimic those scenarios where pumps are switched on or off or when their rpms change abruptly. Table I gives the number of sinusoids estimated using IQML in each case, the frequency band over which simulations were performed and the strongest frequencies estimated in case of stationary pumps.

Table I. Summary of simulation parameters and results

Signal Set	**K**	**Frequency Band (Hz)**	**Strongest Frequency (Hz)**
One Stationary Pump	3	0 - 25	6, 12, 18
Two Stationary Pumps	2	17 - 20	18, 18.9
One Chirp Pump	3	0 - 25	Time varying
Two Chirp Pumps	2	5 - 7	Time varying

Fig. 2 to 5 give the spectral content and the IQML estimates of the stationary pumps. The chirp pump results are given in Fig. 6 to 9.

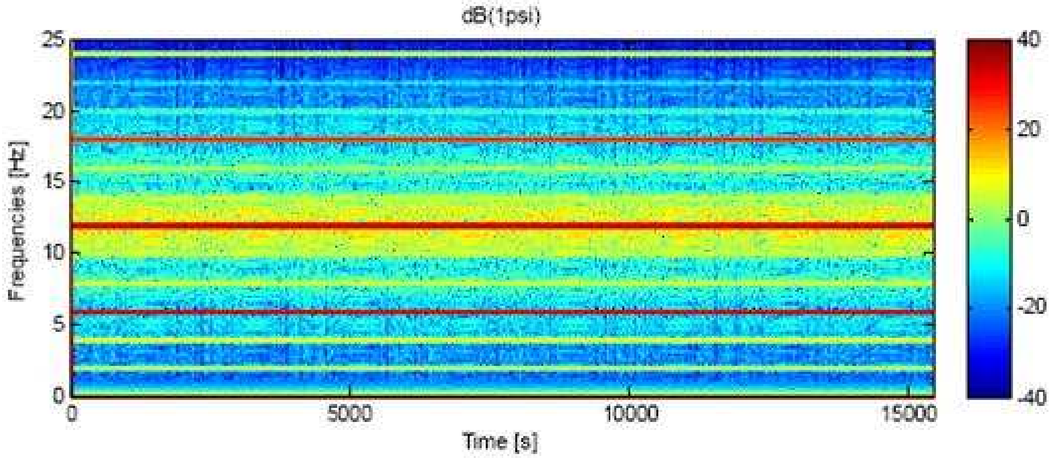


Fig. 2. Time frequency spectrum with one stationary pump

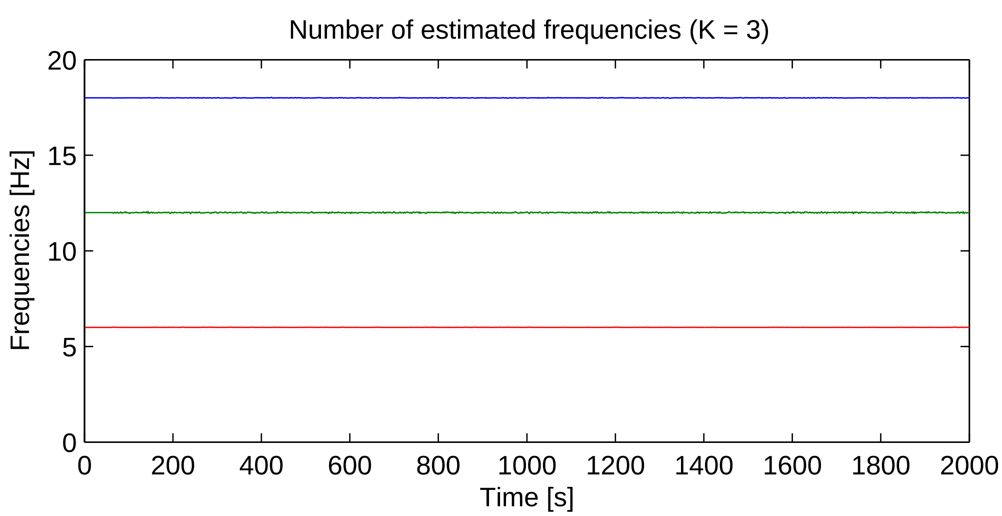


Fig. 3. Estimated frequencies with one stationary pump

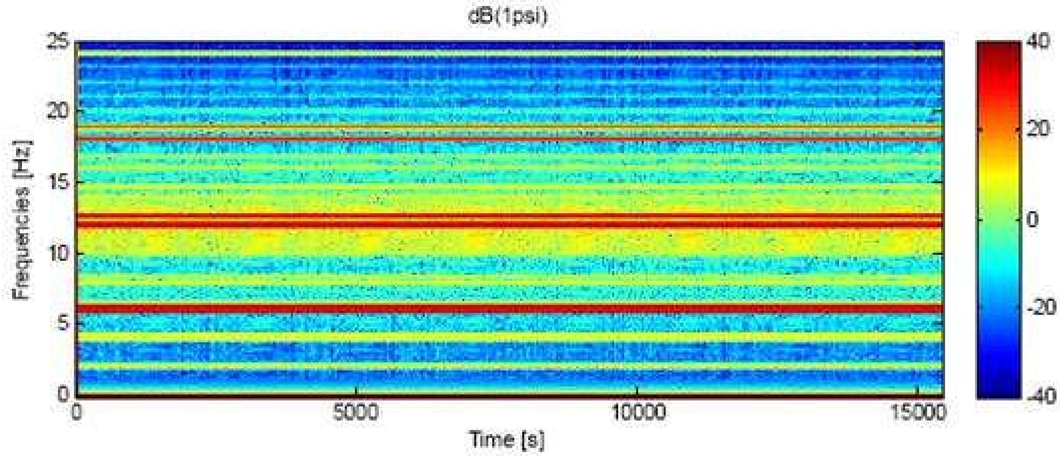


Fig. 4. Time frequency spectrum with two stationary pumps

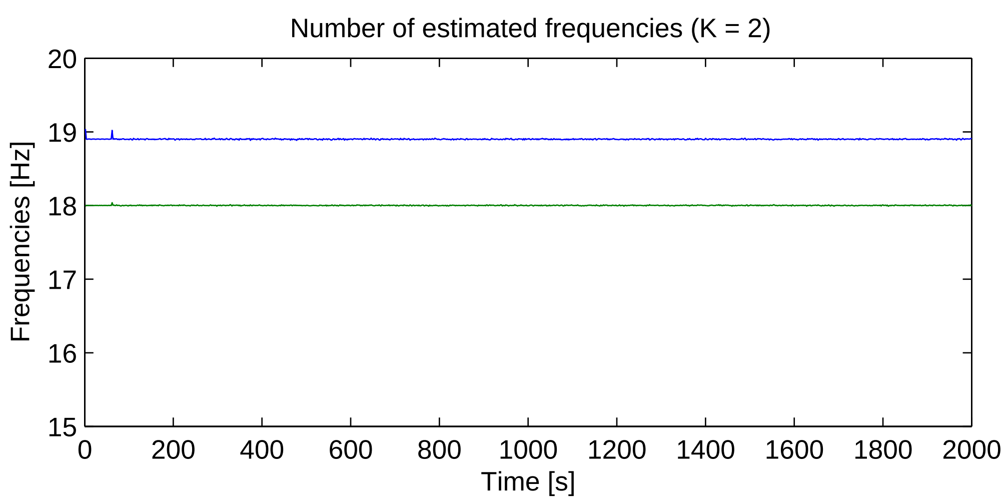


Fig. 5. Estimated frequencies with two stationary pumps

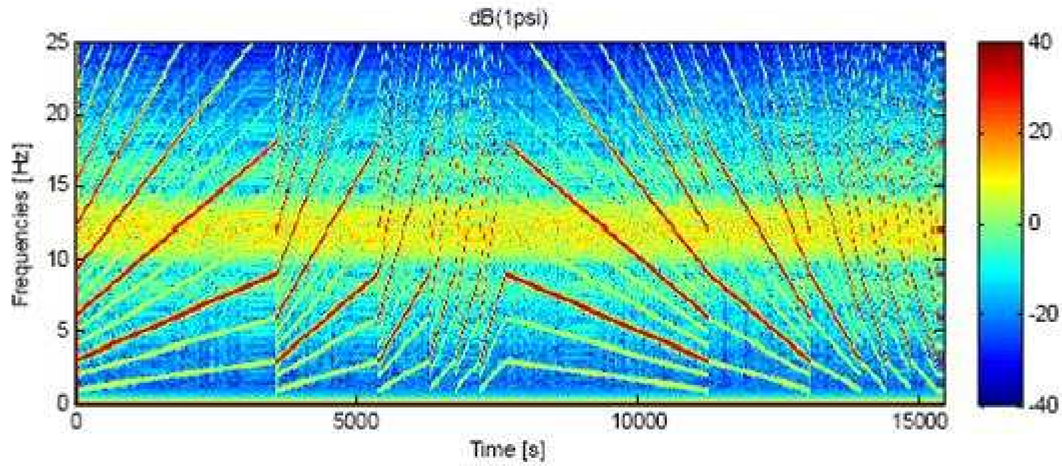


Fig. 6. Time frequency spectrum with one chirp pump

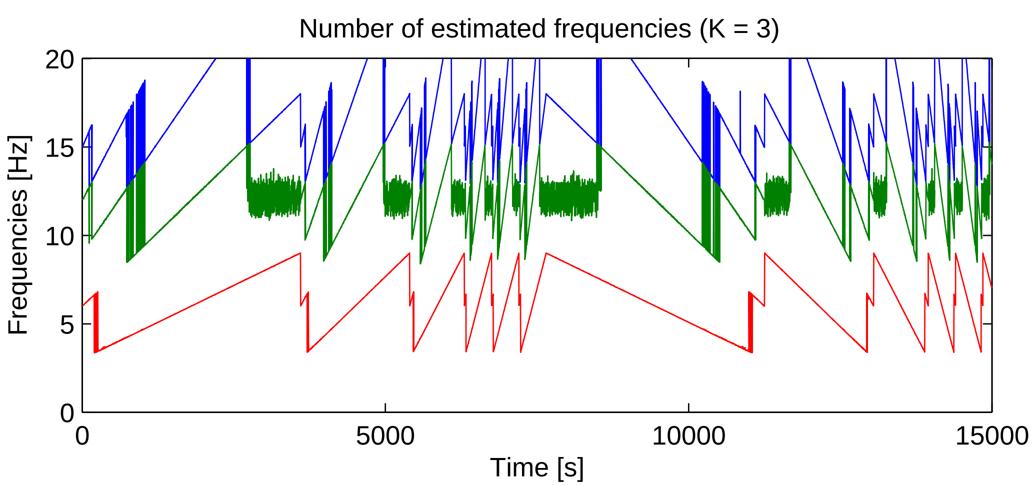


Fig. 7. Estimated frequencies with one chirp pump

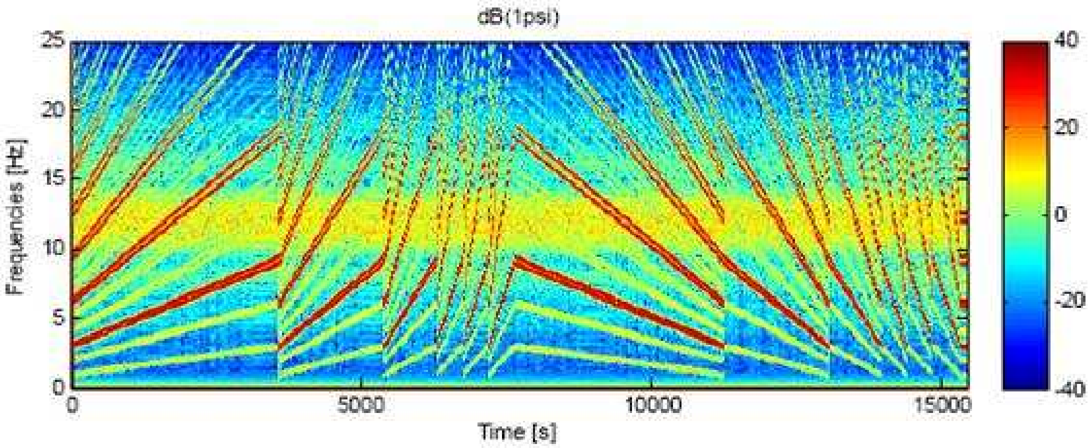


Fig. 8. Time frequency spectrum with two chirp pumps

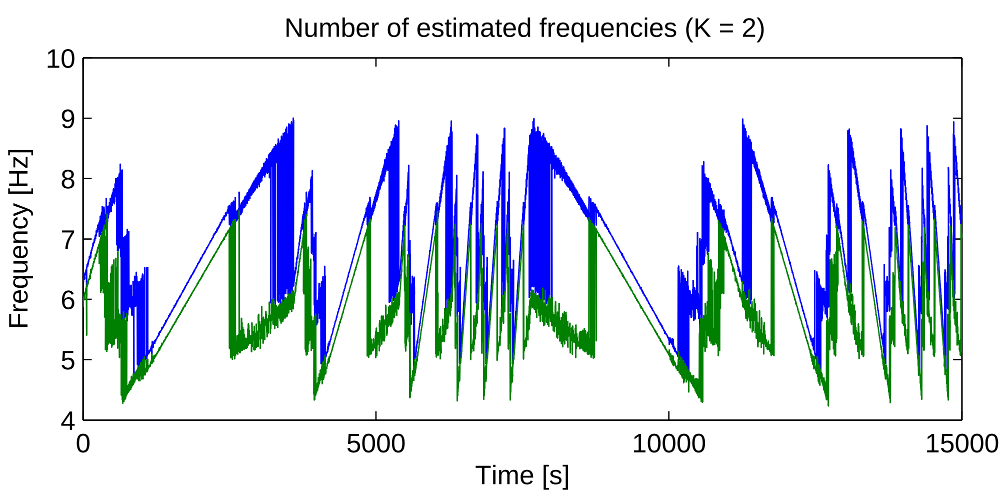


Fig. 9. Estimated frequencies with two chirp pumps

CHAPTER III

HARMONIC SET ESTIMATION

The mud pump noise is composed of several harmonics of one or more fundamental frequencies. If the harmonic relationship of the sinusoids can be worked into the problem the parameter estimates can be significantly improved as compared to the case where they are treated as independent sinusoids. The harmonic set estimation finds application in the field of musical sound analysis [14], voice recognition, power systems [15], time series analysis [16] etc. In most of the previous works the emphasis is on using the harmonic structure to improve the initial estimates obtained by solving standard least squares problem. In [17], complex harmonics are considered and accurate frequency estimation is achieved via weighted least squares(WLS) where the weighting matrix is given by the Markov estimate. An initial estimate of all harmonic parameters is required to construct the Markov estimate which is obtained from applying MUSIC approach described in [18]. The algorithm described in [19] extends the similar idea to real harmonic sinusoidal frequencies.

A less intensive approach is to first determine the fundamental frequency of the harmonic sinusoidal signals present and then obtain the amplitudes and phase of all the harmonics of this fundamental by solving the least squares problem. The problem of fundamental frequency estimation is the same as that of pitch estimation in sound signals. It is one of the classic speech processing problems that is still a hot topic. A number of techniques have been proposed for pitch estimation, mostly aiming at the measurement of periodicity in the time or frequency domain. The telemetry signal is usually less complex than a voice or music signal and thus the extensive techniques used of pitch estimation will unnecessarily increase the complexity of computation. A more computationally efficient method proposed in [14] estimates fundamental frequency based on the estimate of a small set of *partials* obtained using short time fourier transform (STFT) as opposed to [17], [19] where an initial estimate of the sinusoids or the fundamental frequency was required. In this chapter we first attempt to exploit the harmonic relationship to rewrite the optimization step in IQML approach. Solving this optimization problem directly produces an estimate for fundamental frequency but it involves complicated computations. A simpler method on the lines of [14] for fundamental frequency estimation is then proposed which requires preprocessing on the data to exploit the inherent diversity in the system. The results obtained after testing the proposed algorithm on the synthetic set of data are discussed and its performance is evaluated.

A. Direct Fundamental Frequency Estimation

The harmonic relationship can be worked into the original problem such that the optimization step is reduced to a single dimensional minimization. The optimization in case of real sinusoids is given as

$$\tilde{\mathbf{b}}^{(n+1)} = \arg \min_{\|\tilde{\mathbf{b}}\|=1} \{\tilde{\mathbf{b}}^{*}\tilde{\mathbf{C}}_{\mathbf{y}}^{(\mathbf{n})}\tilde{\mathbf{b}}\} \tag{3.1}$$

For simplicity, assume that there is only one active pump in the system so the frequency parameters are given by $\omega_k = k\omega_0,\ k = 1, \ldots, K$. Also, assume that the number of harmonics of the fundamental frequency present in the signal can be accurately estimated. The fundamental frequency ω_0 is related to the linear prediction polynomial coefficients $\mathbf{b}$ via the equation

$$b(z) = \sum_{k=0}^{K} b_k z^{K-k} = b_0 \sum_{k=1}^{\frac{K}{2}} (z^2 - 2\cos(k\omega_0)z + 1). \tag{3.2}$$

On further simplification $\cos(k\omega_0)$ can be expressed in terms of $\cos(\omega_0)$ by using

Tchebychev polynomials of the first kind

$$\cos(k\omega_0) = T_k(\cos(\omega_0)) \tag{3.3}$$

where,

$$T_k(x) = \frac{k}{2}\sum_{r=0}^{\lfloor k/2 \rfloor} \frac{(-1)^r}{k-r} C_r^{k-r} (2x)^{k-2r} \tag{3.4}$$

Equivalently, $\tilde{\mathbf{b}}$ can be written as

$$\tilde{\mathbf{b}} = \mathbf{C} \begin{pmatrix} 1 \\ \cos\omega_0 \\ \vdots \\ \cos^R \omega_0 \end{pmatrix} \tag{3.5}$$

for $R = \frac{K^2+2K}{8}$ and some fixed $(\frac{K}{2}+1) \times (R+1)$ matrix C. For example, with $K = 4$ we have

$$\mathbf{C} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & -2 & -4 & 0 \\ 2 & -4 & 0 & 8 \end{pmatrix}. \tag{3.6}$$

With this relationship, the quadratic optimization problem in 3.1 can be written as

$$\begin{aligned} \widehat{\cos\omega_0}^{(n+1)} &= \arg\min_{\cos\omega_0} \left\{ \begin{pmatrix} 1 & \cos\omega_0 & \cdots & \cos^R\omega_0 \end{pmatrix} \mathbf{C}^* \tilde{\mathbf{C}}_y^{(n)} \mathbf{C} \begin{pmatrix} 1 \\ \cos\omega_0 \\ \vdots \\ \cos^R\omega_0 \end{pmatrix} \right\} && (3.7) \\ &= \arg\min_{\cos\omega_0} \mathrm{Tr} \left(\mathbf{C}^* \tilde{\mathbf{C}}_y^{(n)} \mathbf{C} \begin{pmatrix} 1 & \cos\omega_0 & \cdots & \cos^R\omega_0 \\ \cos\omega_0 & \cos^2\omega_0 & \cdots & \cos^{R+1}\omega_0 \\ \vdots & \vdots & & \vdots \\ \cos^R\omega_0 & \cos^{R+1}\omega_0 & \cdots & \cos^{2R}\omega_0 \end{pmatrix} \right) && (3.8) \end{aligned}$$

which can be solved by setting the derivative

$$\frac{\partial}{\partial \cos\omega_0}\mathrm{Tr}\left(\mathbf{C}^*\tilde{\mathbf{C}}_y^{(n)}\mathbf{C}\begin{pmatrix} 1 & \cos\omega_0 & \cdots & \cos^R\omega_0 \\ \cos\omega_0 & \cos^2\omega_0 & \cdots & \cos^{R+1}\omega_0 \\ \vdots & \vdots & & \vdots \\ \cos^R\omega_0 & \cos^{R+1}\omega_0 & \cdots & \cos^{2R}\omega_0 \end{pmatrix}\right) = 0. \quad (3.9)$$

Note that this derivative with respect to $\cos\omega_0$ is a *polynomial* of $\cos\omega_0$ and hence amenable for numerical solution. Once an estimate of $\cos\omega_0$ is obtained, an estimate of $\tilde{\mathbf{b}}$ can be obtained via 3.5, and the iteration can continue to the next round. Such an approach will be more computationally efficient as no searching procedure is involved. However, as the number of the fundamental frequencies and harmonics present increase the complexity of the computations increases. [19] follows a similar approach but the fundamental frequency estimation is based on WLS technique.

B. Proposed Method

The main steps of the proposed method are shown in Fig. 10.

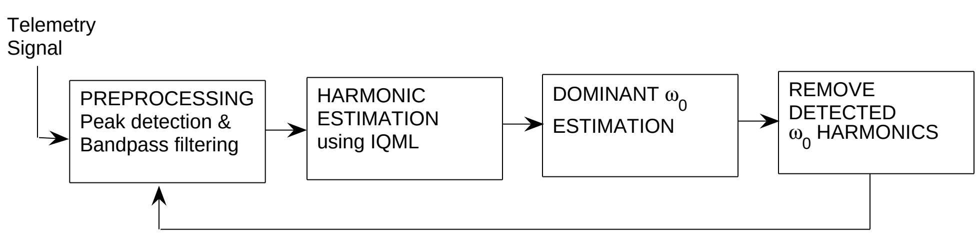


Fig. 10. Flowchart of the proposed method

1. Preprocessing

The presence of multiple harmonics of the same fundamental provides selection diversity in the system which can be exploited during the estimation of fundamental frequency. In order to harness this diversity, preprocessing of the spectrum before the actual f_0 analysis is required. It is an important factor in the performance of the system as it provides robustness in additive noise and ensures that signal with varying spectral shapes can be handled. The goal is to reliably estimate only the strongest harmonics of a particular fundamental frequency and then run the proposed algorithm on these estimates.

The preprocessing module generates the frequency spectrum of each signal frame (comprising of 2 seconds of data sampled at 120Hz) using the Fast Fourier Transform (FFT) method. Under reasonable assumptions, each harmonic in the input signal produces a local maximum in the magnitude spectrum. Several heuristics were proposed to discriminate local maxima induced by harmonics from those induced by noise. For the operational SNR values, an iterative peak detector can be used to determine all the major peaks above a certain threshold. Once the strongest harmonics are located in the spectrum an aptly designed bandpass filter can be used to isolate them. The iterative NLS estimation scheme described in chapter II can then be applied to the filtered signal. The strongest harmonic frequency f is estimated from the given data.

2. Dominant Fundamental Frequency Estimation

The goal here is to estimate the dominant fundamental frequency ω_0 in signal using hypothesis testing procedure to estimate harmonic number k first. The harmonic estimated by the IQML procedure is given by $\hat{\omega}$. Since the fundamental frequency is closely related to the speed (in rpm) of the mud-pump, ω_0 can b e assumed to be restricted to the range $[\omega_{0,min}, \omega_{0,max}]$. Then, the k-th hypothesis is given as:

$$\hat{\omega} = k\omega_0, \; k \in \left\{ \left\lceil \frac{\hat{\omega}}{\omega_{0,max}} \right\rceil, \left\lceil \frac{\hat{\omega}}{\omega_{0,max}} \right\rceil + 1, \ldots, \left\lfloor \frac{\hat{\omega}}{\omega_{0,min}} \right\rfloor \right\} \quad (3.10)$$

Under the k-th hypothesis, the fundamental frequency is given by:

$$\omega_{0,k} = \frac{\hat{\omega}}{k} \quad (3.11)$$

The next step is to determine a testing criterion. The mud pump noise under the k-th hypothesis is reconstructed as:

$$z_k[m] = \sum_{l=1}^{L_k} \hat{A}_{l,k} \cos(l\omega_{0,k}m + \hat{\phi}_{l,k}) \quad (3.12)$$

where, L_k is the total number of harmonics of $\omega_{0,k}$, for a certain sampling rate. By Nyquist theorem it is given by:

$$L_k = \left\lfloor \frac{\pi}{\omega_{0,k}} \right\rfloor \quad (3.13)$$

In the absence of a certain fundamental frequency, the $z_k[m]$ will just consist of the background noise components. The energy per harmonic of the fundamental frequency $\omega_{0,k}$ can be defined as:

$$D_k = \frac{\|z_k\|^2}{L_k}. \quad (3.14)$$

The average harmonic energy D_k is similar to the continuous time domain transform defined in [20] for the instantaneous frequency estimation. It can be used as a metric to test different hypotheses. More precisely, the correct estimate of k can be obtained by:

$$\hat{k} = \arg\max_k D_k \quad (3.15)$$

Then, the fundamental frequency is given by $\hat{\omega_0} = \hat{\omega}/\hat{k}$

The accuracy of fundamental frequency estimate depends on the accuracy of the "first" estimate i.e. the harmonic estimate and also, on the apriori knowledge of the range of fundamental frequency.

The algorithm can be summarized as follows:

a) Initialization: $k = \left\lceil \frac{\hat{\omega}}{\omega_{0,max}} \right\rceil$.

b) Calculate $z_k[m] = \sum_{l=1}^{L_k} \hat{A}_{l,k} \cos(l\omega_{0,k}m + \hat{\phi}_{l,k})$ and $D_k = \frac{\|z_k\|^2}{L_k}$.

c) If $k < \left\lfloor \frac{\hat{\omega}}{\omega_{0,min}} \right\rfloor$, then $k = k + 1$ and repeat step b.

d) Finally, $\hat{k} = \arg\max_k D_k$ and fundamental frequency $\hat{\omega_0} = \hat{\omega}/\hat{k}$.

The exact value of the estimated ω_0 is based on the frequency estimate of a single strongest harmonic. However, the ω_0 estimate can be improved by considering the frequency estimates of all the harmonics in harmonic series of the winner candidate during peak detection in the preprocessing stage. The ω_0 refinement can be thought of as a weighted average of the local ω_0 estimates, where the local estimate for the k^{th} harmonic is $\omega/\hat{k}$. The weights are assigned according to the SNR and the stability of the absolute frequency and are obtained during the preprocessing of the data.

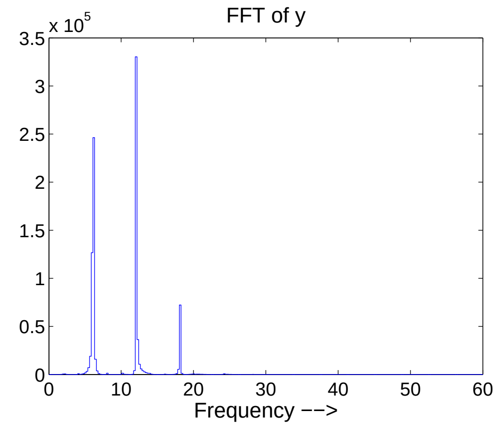
In the presence of more than one fundamental frequency the process can be repeated. Multiple-ω_0 estimation accuracy can be improved by an iterative estimation and cancelation scheme where each detected harmonics series is canceled from the signal before estimating the next ω_0. The basic cancelation mechanism described here is similar to that presented in [21] for a mixture of sound signals.

3. Simulation Results

The simulations were performed both on synthetic sets as well as real field data provided. Presented here are the results of simulations run on synthetic data from stationary pump 1 and 2. The stationary pump 1 operates at 120RPM giving a fundamental frequency of 2Hz. The strongest harmonics are the multiples of three i.e $3^{rd}, 9^{th}, 15^{th}$ and so on. Based on this observation, the fundamental frequency is assumed to lie in the range $[\frac{\omega_0}{2}, \frac{3\omega_0}{2}]$. So, for the stationary pump 1 the analog frequency $F_0 \in [1, 3]Hz$. The frequency spectrum of this pump is shown in Fig. 11. It also shows the distribution of the energy per harmonic of the fundamental frequency $\omega_{0,k}$, given by D_k against the hypothesis k. The correct estimate of harmonic number $\hat{k}$ is the one for which D_k is maximum. In this case 9^{th} harmonic was estimated giving a fundamental frequency of 2.001 Hz. The experiments ar repeated for different filter parameters and range of fundamental frequency. Similar, simulations were run on stationary pump 2. It contained two different fundamental frequencies corresponding to two mud pumps running at 120RPM and 126RPM. The fundamental frequencies were picked up to be 2.001Hz and 2.1025HZ, in two successive runs of the detection algorithm. The results are presented in Fig. 12 and Fig. 13. These results conform with the available knowledge of fundamental frequency and hence prove the accuracy of detection routine.

The experiments are repeated on real engineering data. The range of fundamental frequencies for each signal set is selected based on the information on pump stroke rates. The precision of proposed method has the same order of magnitude as that of the sinusoid estimator employed.The method is also robust against weak or absent fundamentals as well as incomplete series e.g. only multiples of 3, harmonics present.

One major drawback of this technique is that it is highly sensitive to the range of fundamental frequency provided, especially in the signals where a regular pattern of harmonics are absent. For example, in stationary pump 2 if the range fundamental frequency is increased to $[\frac{\omega_0}{2}, 2\omega_0]$ i.e. $F_0 \in [1, 4]Hz$. Then D_k can achieve a false maximum for $F_0 = 3Hz$ and still cover all the major harmonics.



(a) FFT of signal

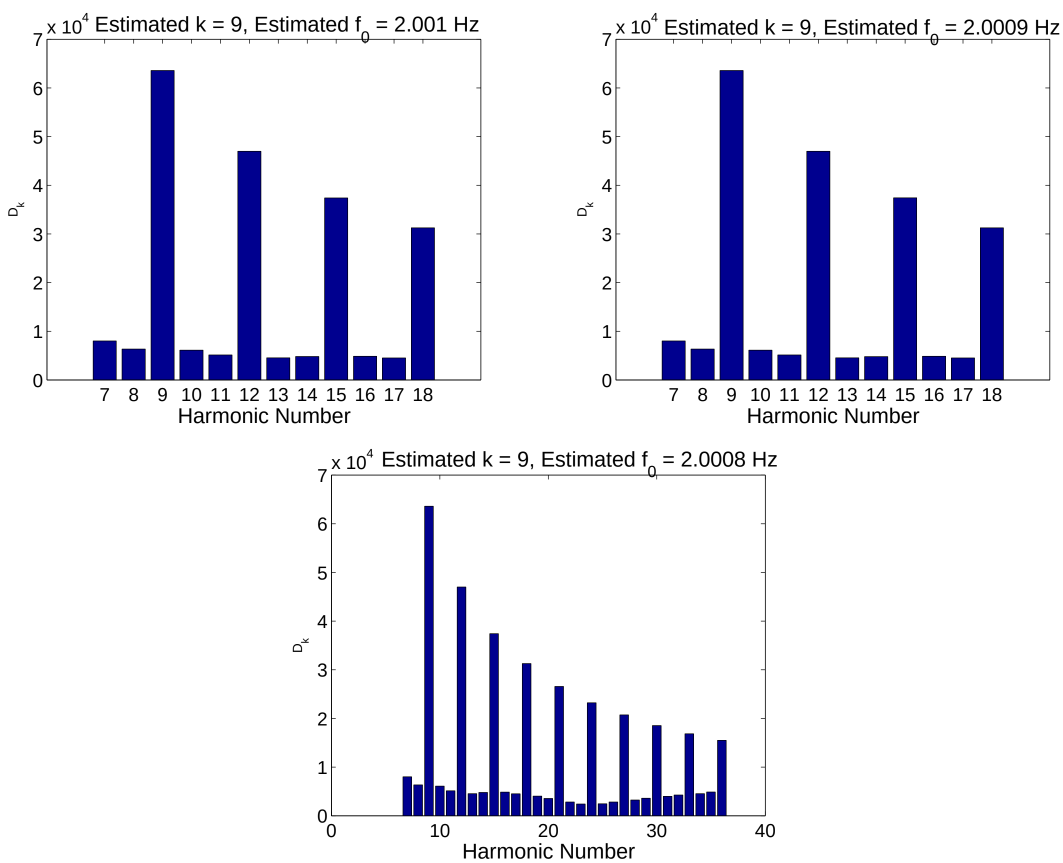
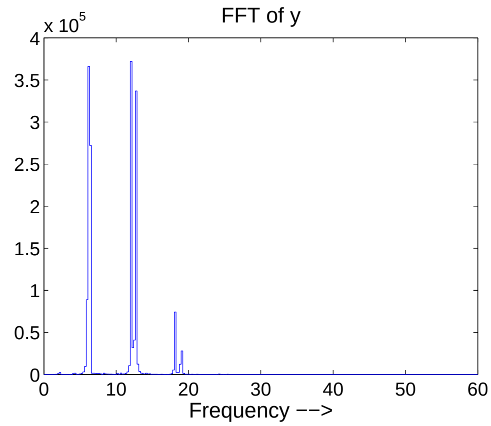


Fig. 11. Energy per harmonic with one pump



(a) FFT of signal

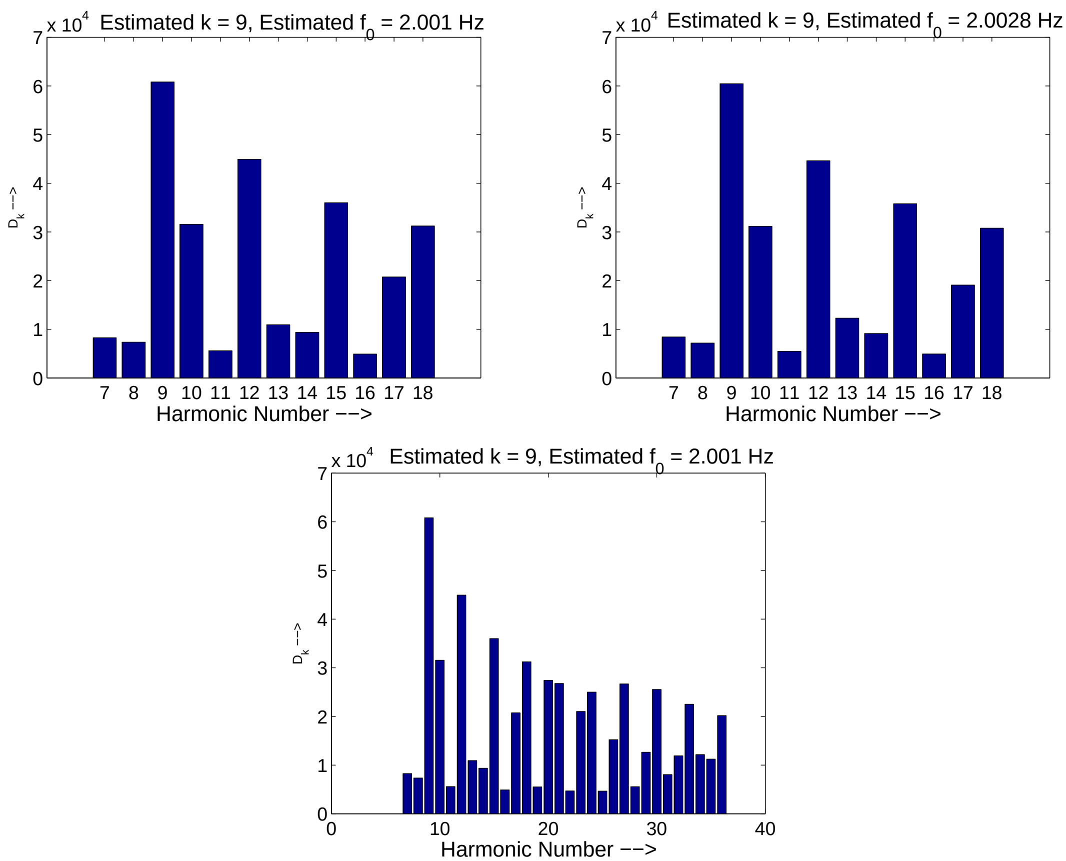
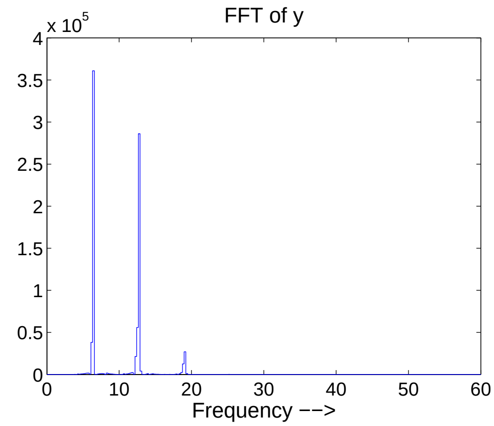


Fig. 12. Energy per harmonic with two pumps



(a) FFT of signal

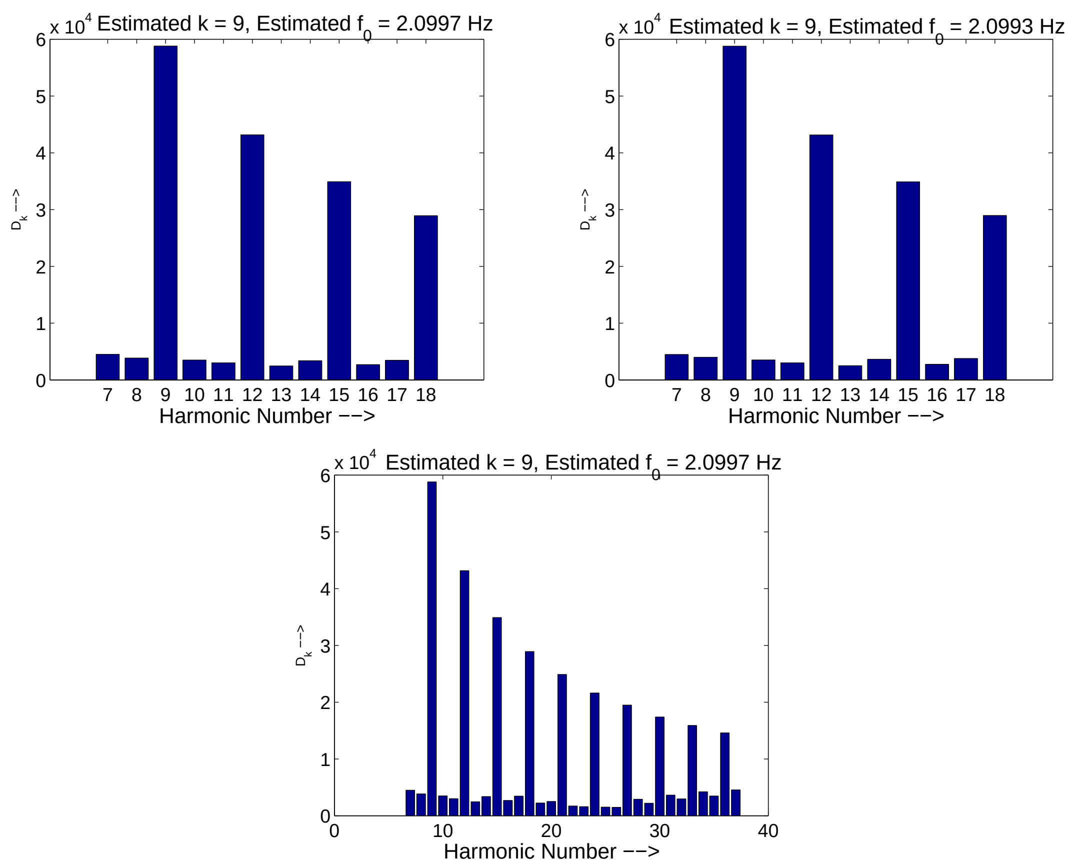


Fig. 13. Energy per harmonic after cancellation

CHAPTER IV

CONCLUSION

The problem of wideband noise cancelation over MWD mud-pulse telemetry channels was solved to a great extent by the use of aforementioned techniques. The principal difficulty of overlapping spectral components due to use of multiple mud pumps was overcome using state-of-the art IQML based estimation algorithm. For appropriately selected band limited signals it provides enough resolution to pick up these overlapping components separately. In addition to this, apriori knowledge of number noise causing sources i.e.mud pumps and their stroke rate is utilized to establish the harmonic relationship between estimated frequencies and their corresponding fundamental frequency. An attempt is made to build the harmonic constraint into the original IQML optimization routine, but later a more straightforward hypothesis testing procedure is developed to estimate the fundamental frequencies. The knowledge of fundamental frequency is used to build the entire harmonic set and once the remaining parameters of noisy sinusoids are estimated they can be fed as input to a specifically designed noise canceler .

Thus far, the described procedures have been successfully tested over the synthetic as well as the real field data and the results are encouraging. As mentioned in Chapter III, there are some harmonic patterns which can produce ambiguous results in the hypothesis testing procedure, but they can be avoided by carefully selecting the fundamental frequency ranges.

REFERENCES

[1] C. Offelli, C. Narduzzi, D. Petri, and F. E. Zanin, "Comparative analysis of non-parametric and parametric frequency estimation methods," in *IEEE Instrumentation and Measurement Technology Conference, 1993. IMTC/93. Conference Record.*, 1993, pp. 603–607.

[2] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, Prentice Hall PTR, Upper Saddle River, NJ, USA, 1993.

[3] M. H. Kahn, M. S. Mackisack, M. R. Osborne, and G. K. Smyth, "On the consistency of prony's method and related algorithms," *Journal of Computational and Graphical Statistics*, vol. 1, no. 4, pp. 329–349, 1992.

[4] R. Kumaresan, L. Scharf, and A. Shaw, "An algorithm for pole-zero modeling and spectral analysis," *IEEE Transactions on Acoustics Speech and Signal Processing*, vol. 34, no. 3, pp. 637–640, 1986.

[5] Y. Bresler and A. Macovski, "Exact maximum likelihood parameter estimation of superimposed exponential signals in noise," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 34, no. 5, pp. 1081–1089, 1986.

[6] V. Nagesha, S. Kay, D. S. Inc, and M. A. Newton, "On frequency estimation with the IQML algorithm," *IEEE Transactions on Signal Processing*, vol. 42, no. 9, pp. 2509–2513, 1994.

[7] M. P. Clark and L. L. Scharf, "On the complexity of IQML algorithms," *IEEE Transactions on Signal Processing*, vol. 40, no. 7, pp. 1811–1813, 1992.

[8] P. Stoica, J. Li, and T. Söderström, "On the inconsistency of IQML," *Signal Processing*, vol. 56, no. 2, pp. 185–190, 1997.

[9] J. Li, P. Stoica, and Z. S. Liu, "Comparative study of IQML and MODE for direction-of-arrival estimation," in *1997 IEEE International Conference on Acoustics, Speech, and Signal Processing, 1997. ICASSP-97.*, 1997, vol. 5.

[10] G. Zhu, W. Y. Choy, and B. C. Sanctuary, "Spectral parameter estimation by an iterative quadratic maximum likelihood method," *Journal of Magnetic Resonance*, vol. 135, no. 1, pp. 37–43, 1998.

[11] H. Akaike, "A new look at the statistical model identification," *IEEE Transactions on Automatic Control*, vol. 19, no. 6, pp. 716–723, 1974.

[12] S. Bakamidis, M. Dendrinos, and G. Carayannis, "SVD analysis by synthesis of harmonic signals," *IEEE Transactions on Signal Processing*, vol. 39, no. 2, pp. 472–477, 1991.

[13] J.Q. Zhang, "An eigenvalue residuum-based criterion for detection of the number of sinusoids in white gaussian noise," *Digital Signal Processing*, vol. 13, no. 2, pp. 275–283, 2003.

[14] A. Mitre, M. Queiroz, and R. R. A. Faria, "Accurate and efficient fundamental frequency determination from precise partial estimates," in *4th AES Brazil Conference*, 2006, pp. 113–118.

[15] T. Lobos, T. Kozina, and H.J. Koglin, "Power system harmonics estimation using linear least squares method and SVD," *IEE Proceedings-Generation, Transmission and Distribution*, vol. 148, no. 6, pp. 567–572, 2001.

[16] A. M. Walker, "On the estimation of a harmonic component in a time series with stationary independent residuals," *Biometrika*, vol. 58, no. 1, pp. 21, 1971.

[17] H. Li, P. Stoica, and J. Li, "Computationally efficient parameter estimation for harmonic sinusoidal signals," *Signal Processing*, vol. 80, no. 9, pp. 1937–1944, 2000.

[18] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Transactions on Antennas and Propagation*, vol. 34, no. 3, pp. 276–280, 1986.

[19] K. W. Chan and H. C. So, "Accurate frequency estimation for real harmonic sinusoids," *IEEE Signal Processing Letters*, vol. 11, no. 7, pp. 609–612, 2004.

[20] T. Abe, T. Kobayashi, and S. Imai, "Robust pitch estimation with harmonics enhancement in noisy environments based on instantaneous frequency," in *Fourth International Conference on Spoken Language Processing*, 1996, pp. 1277–1280.

[21] A.P. Klapuri, "Multiple fundamental frequency estimation based on harmonicity and spectral smoothness," *IEEE Transactions on Speech and Audio Processing*, vol. 11, no. 6, pp. 804–816, 2003.

VITA

Richa Dixit is a graduate student in the Dept. of Electrical and Computer Engineering at Texas A&M University. She received her Bachelor of Technology degree in electronics engineering from the Institute of Technology, Banaras Hindu University, India in 2007. During the course of her undergraduate education, she worked briefly as an intern in the National Institute of Oceanography, India and Siemens AG, India. She obtained her M.S. in electrical engineering at TAMU in 2010. Her research work on the parametric estimation of harmonic sinusoids was motivated by the project funded by Schlumberger. She can be contacted through the following address: Department of Electrical and Computer Engineering, Texas A&M University, College Station, Texas 77843-3128. Her email-id is richadixit.ece@gmail.com.