Residual carrier frequency offset and symbol timing offset in narrow-band OFDM modulation

Master's thesis
Niko Ohukainen
2264750
Department of Mathematical Sciences
University of Oulu
Spring 2017

Contents

Introduction

The technological progress in wireless communication has enabled us to be reachable anywhere without the help of electrical conductors, such as wires. The modern cellular network offers almost ubiquitous connectivity and wireless local area networks are found in homes, offices and commercial spaces. The development has led to the point where each thing is equipped with wireless connectivity. This expanding of Internet and networks to cover thing like meters, buildings and health care appliances is called the *Internet of Things* (IoT).

The requirements for the connected things are different from the requirements for mobile phones and computers that have fixed power source or are easily recharged. Mobile phones and computers are characterized by the need for high data rates and very small end-to-end transmission delays. These devices have enough computing power to handle complex transmission schemes and can use resources to transmit with high power. For things like temperature meters and location beacons, this is not possible. The devices might be required to run from a small battery for years. In addition, these devices might need to transmit or receive only small amounts of data with long breaks in between. From the device point of view, fast synchronization is vital to be able to save power when transmission is not needed. The possibly low signal strength together with the low computation power lay a challenge for the synchronization process because the parameters need to be inferred from the received signal.

Orthogonal frequency-division multiplexing (OFDM) is a digital modulation method used in existing wireless standards such as Long-Term Evolution (LTE) [1], IEEE 802.11 [2] and Digital Video Broadcasting (DVB) [3]. The 3rd Generation Partnership Project (3GPP) is defining a radio access technology based on OFDM that is targeted to enable IoT devices ubiquitous networking [1, chapter 10]. To reach the requirement of low cost and low power consumption, the transmission bandwidth is narrowed and repetitions are introduced to enable communications in situations where receiver suffers from very low signal strength. This specification is used as a guideline for the numerical simulations in section 4.

In the first section, a few concepts are introduced that give the foundation to the reviews in following sections. The *Fourier transform* is defined together with its discrete version, which gives the basis for OFDM modulation. The *Nyquist-Shannon sampling theorem* proves that a continuous signal can be constructed from its discrete samples without losing information of the original signal under certain requirements. This thesis focuses on the estimation of synchronization parameters in OFDM system and the Cramer-Rao

lower bound is used to give a benchmark for the derived estimator schemes. Lastly the Kalman filtering concept is presented which is used in one of the estimators derived in section 3.

The second section defines the transmission system using OFDM modulation to transmit data. The transmitted signal is derived together with the perturbations caused by the multi-path channel. An OFDM receiver is responsible for demodulating the signal and de-multiplexing the data streams carried on different sub-carriers. In this section the effect of non-idealities in the OFDM system are introduced and their effect is incorporated to the signal model derived in the last part of this section.

The estimation of the synchronization parameters is discussed in the third section. Three estimators are derived according to the signal model presented in the second section and the Cramer-Rao lower bound is derived to give a benchmark for the estimators. The estimators are developed for two purposes, to estimate the parameters from a short signal sample and to track the time-variant change in the estimation parameters during continuous reception.

In the last section the derived estimators' performance is numerically evaluated in a simulation environment. The estimator mean, mean squared error performance and continuous tracking capabilities are studied.

1 Preliminaries

1.1 Fourier transform

The Fourier transform is used to represent a function of time with the frequencies it is composed of. Taking the Fourier transform of a function gives a complex-valued function; where the magnitude represents the amount of that frequency in the original function and the phase represent the phase offset from the basic sinusoid of that frequency.

Definition 1.1. If function $s : \mathbb{R} \to \mathbb{C}$ is such that $\int_{-\infty}^{\infty} |s(t)|dt < \infty$, then its *Fourier transform* $S : \mathbb{R} \to \mathbb{C}$ is the bounded function on $\mathbb{R}$ defined by

$$S(f) = \mathcal{F}(s)(f) = \int_{-\infty}^{\infty} e^{-j2\pi tf} s(t)dt. \tag{1}$$

For a function S, $\int_{-\infty}^{\infty} |S(f)|df < \infty$, the *inverse Fourier transform* $\mathcal{F}^{-1}(S)$ is defined as

$$s(t) = \mathcal{F}^{-1}(S)(t) = \int_{-\infty}^{\infty} e^{j2\pi tf} S(f)df. \tag{2}$$

The Fourier transform and the inverse Fourier transform are, as the naming suggests, inverse operations.

Theorem 1.2. *For a function $s : \mathbb{R} \to \mathbb{C}$, that has a Fourier transform $\mathcal{F}(s) = S : \mathbb{R} \to \mathbb{C}$, it holds that*

$$\mathcal{F}^{-1}\left(\mathcal{F}(s)\right)(t) = s(t). \tag{3}$$

Proof. See Theorem 5.3 in [4]. □

The corresponding operation in discrete domain is the discrete Fourier transform. It transforms a finite, equally spaced sequence of time domain samples to an equally spaced frequency domain sequence. The interval at which the discrete Fourier transform is sampled is the reciprocal of the duration of the time domain sequence.

Definition 1.3. Let $s[n] \in \mathbb{C}^N$ be a finite sequence of equally spaced values of a function s. The *discrete Fourier transform* $S[k] \in \mathbb{C}^N$ is defined as

$$S[k] = \mathcal{F}(s)[k] = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} e^{-j2\pi kn/N} s[n]. \tag{4}$$

The *inverse discrete Fourier transform* is defined as

$$s[n] = \mathcal{F}^{-1}(S)[n] = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} e^{j2\pi kn/N} S[k]. \tag{5}$$

The discrete Fourier transform can be presented as a linear mapping by defining the discrete Fourier transform matrix $\mathbf{F} \in \mathcal{M}_{N\times N}$, where the elements are given by

$$F_{n,k} = \frac{1}{\sqrt{N}} e^{-j2\pi kn/N}. \tag{6}$$

The Fourier transform is then given as

$$\mathbf{S} = \mathbf{F}\mathbf{s}, \tag{7}$$

where $\mathbf{s}$ and $\mathbf{S}$ are the discrete sequences in vector form.

In addition, the discrete Fourier transform and the inverse discrete Fourier transform are each other's inverse operations.

Theorem 1.4. *Let $s[n] \in \mathbb{C}^n$ be a finite sequence. Then*

$$\mathcal{F}^{-1}\left(\mathcal{F}(s)\right)[n] = s[n]. \tag{8}$$

Proof. By definition

$$\begin{aligned}\mathcal{F}^{-1}\left(\mathcal{F}(s)\right)[n] =& \frac{1}{N}\sum_{k=0}^{N-1} e^{j2\pi kn/N}\sum_{m=0}^{N-1} e^{-j2\pi km/n}s[m] \\ =& \frac{1}{N}\sum_{m=0}^{N-1} s[m]\sum_{k=0}^{N-1} e^{j2\pi k(n-m)/N}.\end{aligned}$$

The sum $\sum_{k=0}^{N-1} e^{j2\pi k(n-m)/N}$ evaluates to N when $n-m=0$ and zero otherwise. Then

$$\mathcal{F}^{-1}\left(\mathcal{F}(s)\right)[n] = s[n].$$

□

1.2 Nyquist-Shannon sampling theorem

The following theorem is by Shannon [5]. It states that a continuous signal can be completely recovered from its samples, when they are taken with a rate of twice the signal bandwidth.

Theorem 1.5 (Nyquist-Shannon sampling theorem)**.** *Let $s(t)$ be a band-limited signal such that it's spectrum $S(f) = \mathcal{F}(s)(f)$ is limited to an interval $[-B, B]$. Then the signal can be recovered from it's samples if it is sampled at a frequency of $2B$.*

Proof. Let $S(f)$ be the frequency spectrum of $s(t)$. Then $S(f) = 0$, if $|f| > B$. We know that a signal is completely defined by it's spectrum by

$$s(t) = \int_{-\infty}^{\infty} S(f)e^{j2\pi ft}df = \int_{-B}^{B} S(f)e^{j2\pi ft}df.$$

Then, if we sample the signal at time instances $t = \frac{n}{2B}$, we get

$$s\left(\frac{n}{2B}\right) = \int_{-B}^{B} S(f)e^{j2\pi f\frac{n}{2B}}df = \int_{-B}^{B} S(f)e^{j\pi\frac{nf}{B}}df.$$

On the other hand, the spectrum $S(f)$ can be represented by it's Fourier series as

$$S(f) = \sum_{n=-\infty}^{\infty} c_n e^{-j\pi\frac{nf}{B}},$$

where the coefficients c_n are given by

$$c_n = \frac{1}{2B}\int_{-B}^{B} S(f)e^{j\pi\frac{nf}{B}}df.$$

We see that the integral is exactly a sample of the signal at time index $t = \frac{n}{2B}$ and therefore

$$S(f) = \frac{1}{2B} \sum_{n=-\infty}^{\infty} s\left(\frac{n}{2B}\right) e^{-j\pi \frac{nf}{B}}.$$

Now, we get a representation of the signal $s(t)$ from it's samples as

$$\begin{aligned}
s(t) =& \frac{1}{2B} \int_{-B}^{B} \sum_{n=-\infty}^{\infty} s\left(\frac{n}{2B}\right) e^{-j\pi \frac{nf}{B}} e^{j2\pi ft} df \\
=& \frac{1}{2B} \sum_{n=-\infty}^{\infty} s\left(\frac{n}{2B}\right) \int_{-B}^{B} e^{j2\pi f\left(t-\frac{n}{2B}\right)} df \\
=& \frac{1}{2B} \sum_{n=-\infty}^{\infty} s\left(\frac{n}{2B}\right) \frac{e^{j2\pi f\left(t-\frac{n}{2B}\right)}}{j2\pi\left(t-\frac{n}{2B}\right)}\Big|_{f=-B}^{B} \\
=& \sum_{n=-\infty}^{\infty} s\left(\frac{n}{2B}\right) \frac{e^{j\pi(2Bt-n)} - e^{-j\pi(2Bt-n)}}{j2\pi(2Bt-n)} \\
=& \sum_{n=-\infty}^{\infty} s\left(\frac{n}{2B}\right) \operatorname{sinc}(\pi(2Bt-n)).
\end{aligned}$$

So, the signal $s(t)$ is completely determined by the samples taken at time indices $t = \frac{n}{2B}$. □

1.3 Cramer-Rao lower bound

The Cramer-Rao lower bound gives the lower bound for the variance of an unbiased estimate. An estimator that reaches the Cramer-Rao lower bound has the lowest possible mean squared error amongst unbiased estimators. This fact is given by the following theorem provided by Kay [6].

Theorem 1.6. *Let $f(\mathbf{x};\boldsymbol{\Theta})$ be a probability density function with real valued parameters $\boldsymbol{\Theta} \in \mathbb{R}^p$. Let $\mathbf{X} = [X_1, \ldots, X_n]^T \in \mathbb{C}^n$, be a vector of independent random variables with density function $f(\mathbf{x};\boldsymbol{\Theta})$. Assume that the set $S = \{\mathbf{x} \in \mathbb{C}^n \mid f(\mathbf{x};\boldsymbol{\Theta}) > 0\}$ doesn't depend on $\boldsymbol{\Theta}$. Let $\hat{\boldsymbol{\Theta}} = \hat{\boldsymbol{\Theta}}(\mathbf{x})$ be an unbiased estimator of $\boldsymbol{\Theta}$. Assume that following regularity conditions are satisfied*

1. *Derivatives $\partial f(\mathbf{x};\boldsymbol{\Theta})/\partial\theta_i$ are integrable and*

$$\frac{\partial}{\partial\theta_i} \int_S f(\mathbf{x};\boldsymbol{\Theta}) d\mathbf{x} = \int_S \frac{\partial f(\mathbf{x};\boldsymbol{\Theta})}{\partial\theta_i} d\mathbf{x} \quad \text{for all } i = 1, 2, \ldots, p. \tag{9}$$

2. *Second derivatives $\partial^2 f(\mathbf{x};\boldsymbol{\Theta})/(\partial\theta_i\partial\theta_j)$ are integrable and*

$$\frac{\partial^2}{\partial\theta_i\partial\theta_j} \int_S f(\mathbf{x};\boldsymbol{\Theta}) d\mathbf{x} = \int_S \frac{\partial^2 f(\mathbf{x};\boldsymbol{\Theta})}{\partial\theta_i\partial\theta_j} d\mathbf{x} \tag{10}$$

for all $i = 1, 2, \ldots, p$ *and* $j = 1, 2, \ldots, p$.

3. *We can change the order of integration and differentiation as*

$$\frac{\partial}{\partial \theta_i} \int_S \hat{\boldsymbol{\Theta}}(\mathbf{x}) f(\mathbf{x}; \boldsymbol{\Theta}) d\mathbf{x} = \int_S \hat{\boldsymbol{\Theta}}(\mathbf{x}) \frac{\partial f(\mathbf{x}; \boldsymbol{\Theta})}{\partial \theta_i} d\mathbf{x} \tag{11}$$

for all $i = 1, 2, \ldots, p$.

Then, the covariance matrix of $\hat{\boldsymbol{\Theta}}$ *satisfies*

$$\mathbf{C}_{\hat{\boldsymbol{\Theta}}} - \mathbf{F}^{-1}(\boldsymbol{\Theta}) \geq 0, \tag{12}$$

where ≥ 0 *means that the matrix is positive semi-definite and where the Fisher information matrix* $\mathbf{F}(\boldsymbol{\Theta})$ *is given by*

$$[\mathbf{F}(\boldsymbol{\Theta})]_{ij} = -\mathbb{E}\left\{ \frac{\partial^2 \ln f(\mathbf{x}; \boldsymbol{\Theta})}{\partial \theta_i \partial \theta_j} \right\}. \tag{13}$$

Further, the variance of estimator $\hat{\theta}_i$ *is lower bounded by*

$$\mathbb{E}\{|\hat{\theta}_i - \theta_i|^2\} \geq [\mathbf{F}^{-1}(\boldsymbol{\Theta})]_{ii}. \tag{14}$$

Proof. Because $\hat{\boldsymbol{\Theta}}$ is unbiased we get

$$\boldsymbol{\Theta} = \mathbb{E}\{\hat{\boldsymbol{\Theta}}\} = \int_S \hat{\boldsymbol{\Theta}}(\mathbf{x}) f(\mathbf{x}; \boldsymbol{\Theta}) d\mathbf{x}.$$

Differentiating both sides with respect to $\boldsymbol{\Theta}$ and using the regularity condition (11) yields

$$\frac{\partial}{\partial \boldsymbol{\Theta}} \boldsymbol{\Theta} = \mathbf{I} = \int_S \hat{\boldsymbol{\Theta}}(\mathbf{x}) \frac{\partial f(\mathbf{x}; \boldsymbol{\Theta})}{\partial \boldsymbol{\Theta}}^T d\mathbf{x},$$

where the superscript $(\cdot)^T$ denotes transpose.

Define a random vector

$$\mathbf{Z} = \frac{1}{f(\mathbf{x}; \boldsymbol{\Theta}) + 1_{S^c}(\mathbf{x})} \frac{\partial f(\mathbf{x}; \boldsymbol{\Theta})}{\partial \boldsymbol{\Theta}},$$

where

$$1_{S^c} = \begin{cases} 1, & \mathbf{x} \in S^c \\ 0, & \mathbf{x} \in S. \end{cases}$$

Then, the expectation of $\mathbf{Z}$ is given by

$$\begin{aligned}\mathbb{E}\{\mathbf{Z}\} &= \int_S \frac{1}{f(\mathbf{x};\boldsymbol{\Theta})}\frac{\partial f(\mathbf{x};\boldsymbol{\Theta})}{\partial \boldsymbol{\Theta}} f(\mathbf{x};\boldsymbol{\Theta})d\mathbf{x} \\ &= \int_S \frac{\partial f(\mathbf{x};\boldsymbol{\Theta})}{\partial \boldsymbol{\Theta}} d\mathbf{x} = \frac{\partial}{\partial \boldsymbol{\Theta}} \int_S f(\mathbf{x};\boldsymbol{\Theta})d\mathbf{x} = \mathbf{0},\end{aligned}$$

where the regularity condition (9) was used.

Evaluate the expression $\mathbb{E}\{(\hat{\boldsymbol{\Theta}} - \boldsymbol{\Theta})\mathbf{Z}^T\}$ to get

$$\begin{aligned}\mathbb{E}\{(\hat{\boldsymbol{\Theta}} - \boldsymbol{\Theta})\mathbf{Z}^T\} =& \mathbb{E}\{\hat{\boldsymbol{\Theta}}\mathbf{Z}^T\} - \boldsymbol{\Theta}\mathbb{E}\{\mathbf{Z}^T\} \\ &= \int_S \hat{\boldsymbol{\Theta}}(\mathbf{x})\frac{1}{f(\mathbf{x};\boldsymbol{\Theta})}\frac{\partial f(\mathbf{x};\boldsymbol{\Theta})}{\partial \boldsymbol{\Theta}}^T f(\mathbf{x};\boldsymbol{\Theta})d\mathbf{x} \\ &= \int_S \hat{\boldsymbol{\Theta}}(\mathbf{x})\frac{\partial f(\mathbf{x};\boldsymbol{\Theta})}{\partial \boldsymbol{\Theta}}^T d\mathbf{x} = \mathbf{I}.\end{aligned}$$

Let $\mathbf{a} \in \mathbb{R}^p$ and $\mathbf{b} \in \mathbb{R}^p$ be arbitrary vectors. Now $\mathbf{a}^T\mathbf{b}$ is a scalar and

$$(\mathbf{a}^T\mathbf{b})^2 = (\mathbf{a}^T\mathbf{I}\mathbf{b})^2 = (\mathbf{a}^T\mathbb{E}\{(\hat{\boldsymbol{\Theta}} - \boldsymbol{\Theta})\mathbf{Z}^T\}\mathbf{b})^2 = (\mathbb{E}\{\mathbf{a}^T(\hat{\boldsymbol{\Theta}} - \boldsymbol{\Theta})\mathbf{Z}^T\mathbf{b}\})^2.$$

Applying Cauchy-Schwartz inequality gives

$$\begin{aligned}(\mathbb{E}\{\mathbf{a}^T(\hat{\boldsymbol{\Theta}} - \boldsymbol{\Theta})\mathbf{Z}^T\mathbf{b}\})^2 \leq& \mathbb{E}\{(\mathbf{a}^T(\hat{\boldsymbol{\Theta}} - \boldsymbol{\Theta}))^2\}\mathbb{E}\{(\mathbf{Z}^T\mathbf{b})^2\} \\ =& \mathbb{E}\{\mathbf{a}^T(\hat{\boldsymbol{\Theta}} - \boldsymbol{\Theta})(\hat{\boldsymbol{\Theta}} - \boldsymbol{\Theta})^T\mathbf{a}\}\mathbb{E}\{\mathbf{b}^T\mathbf{Z}\mathbf{Z}^T\mathbf{b}\} \\ =& \mathbf{a}^T\mathbf{C}_{\hat{\boldsymbol{\Theta}}}\mathbf{a}\mathbf{b}^T\mathbf{C}_{\mathbf{Z}}\mathbf{b},\end{aligned}$$

where $\mathbf{C}_{\hat{\boldsymbol{\Theta}}}$ and $\mathbf{C}_{\mathbf{Z}}$ are the covariance matrices of $\hat{\boldsymbol{\Theta}}$ and $\mathbf{Z}$. Since $\mathbf{b}$ was arbitrary we can choose $\mathbf{b} = \mathbf{C}_{\mathbf{Z}}^{-1}\mathbf{a}$ to have

$$\begin{aligned}(\mathbf{a}^T\mathbf{b})^2 = (\mathbf{a}^T\mathbf{C}_{\mathbf{Z}}^{-1}\mathbf{a})^2 \leq& \mathbf{a}^T\mathbf{C}_{\hat{\boldsymbol{\Theta}}}\mathbf{a}\mathbf{a}^T\mathbf{C}_{\mathbf{Z}}^{-1}\mathbf{C}_{\mathbf{Z}}\mathbf{C}_{\mathbf{Z}}^{-1}\mathbf{a} = \mathbf{a}^T\mathbf{C}_{\hat{\boldsymbol{\Theta}}}\mathbf{a}\mathbf{a}^T\mathbf{C}_{\mathbf{Z}}^{-1}\mathbf{a} \\ \text{or} \qquad \mathbf{a}^T\mathbf{C}_{\mathbf{Z}}^{-1}\mathbf{a}\mathbf{a}^T\mathbf{C}_{\mathbf{Z}}^{-1}\mathbf{a} \leq& \mathbf{a}^T\mathbf{C}_{\hat{\boldsymbol{\Theta}}}\mathbf{a}\mathbf{a}^T\mathbf{C}_{\mathbf{Z}}^{-1}\mathbf{a}.\end{aligned}$$

Because covariance matrices are positive semi-definite and symmetric, the term $\mathbf{a}^T\mathbf{C}_{\mathbf{Z}}^{-1}\mathbf{a}$ is a non-negative real number and we have

$$\mathbf{a}^T\mathbf{C}_{\mathbf{Z}}^{-1}\mathbf{a} \leq \mathbf{a}^T\mathbf{C}_{\hat{\boldsymbol{\Theta}}}\mathbf{a}.$$

Then it follows that $\mathbf{C}_{\hat{\boldsymbol{\Theta}}} - \mathbf{C}_{\mathbf{Z}}^{-1}$ is positive semi-definite.

To prove that $\mathbf{C}_{\mathbf{Z}} = \mathbf{F}(\boldsymbol{\Theta})$ we use the regularity conditions (9) and (10) to find that

$$0 = \frac{\partial}{\partial \theta_i}\frac{\partial}{\partial \theta_j}1 = \frac{\partial}{\partial \theta_i}\frac{\partial}{\partial \theta_j}\int_S f(\mathbf{x};\boldsymbol{\Theta})d\mathbf{x}$$

$$= \frac{\partial}{\partial \theta_i} \int_S \frac{\partial \ln f(\mathbf{x};\mathbf{\Theta})}{\partial \theta_j} f(\mathbf{x};\mathbf{\Theta}) d\mathbf{x}$$
$$= \int_S \left(\frac{\partial^2 \ln f(\mathbf{x};\mathbf{\Theta})}{\partial \theta_i \partial \theta_j} f(\mathbf{x};\mathbf{\Theta}) + \frac{\partial \ln f(\mathbf{x};\mathbf{\Theta})}{\partial \theta_i} \frac{\partial \ln f(\mathbf{x};\mathbf{\Theta})}{\partial \theta_j} f(\mathbf{x};\mathbf{\Theta}) \right) d\mathbf{x}.$$

So

$$\mathbb{E}\{Z_i Z_j\} = -\mathbb{E}\left\{ \frac{\partial^2 \ln f(\mathbf{x};\mathbf{\Theta})}{\partial \theta_i \partial \theta_j} \right\} \qquad \text{or} \qquad [\mathbf{C_Z}]_{ij} = [\mathbf{F}(\mathbf{\Theta})]_{ij}.$$

The last part of this theorem states that the variance of the estimate $\hat{\theta}_i$ is lower bounded by $[\mathbf{F}^{-1}(\mathbf{\Theta})]_{ii}$. This is obtained by the fact that diagonal elements of a positive semi-definite matrix are non-negative and

$$[\mathbf{C}_{\hat{\mathbf{\Theta}}} - \mathbf{F}^{-1}(\mathbf{\Theta})]_{ii} \geq 0 \qquad \text{or} \qquad \mathbb{E}\{|\hat{\theta}_i - \theta_i|^2\} \geq [\mathbf{F}^{-1}(\mathbf{\Theta})]_{ii}.$$

□

1.4 Kalman filtering

Kalman filter is an algorithm to find the solution for a linear quadratic estimation problem. The Kalman filter recursively calculates new estimates of a systems state by using knowledge of previous state and observations. The system model and noise covariances are assumed to be known. In a situation where the model is linear and accurate the Kalman filter is shown to be the optimal estimator in the sense that it minimizes the mean-square error [7].

Consider a stochastic process $\mathbf{X}_k$ that evolves according to the following state transition equation

$$\mathbf{x}_k = f(\mathbf{x}_{k-1}, \mathbf{u}_k, \mathbf{v}_k), \tag{15}$$

where the current state $\mathbf{x}_k$ depends only on the state $\mathbf{x}_{k-1}$ of the previous time step, known control input $\mathbf{u}_k$ and zero mean white Gaussian noise $\mathbf{v}_k$ with covariance $\mathbf{Q}_k$. Observations of the system are provided by

$$\mathbf{y}_k = h(\mathbf{x}_k, \mathbf{n}_k), \tag{16}$$

where $\mathbf{n}_k$ is zero mean white Gaussian noise with covariance matrix $\mathbf{R}_k$.

The Kalman filter is used to estimate the state $\mathbf{x}$. First, the knowledge from previous time step is used to predict the state vector and then new observation is used to make the estimate more accurate. We denote the estimate of state $\mathbf{x}$ at time index n with observations up to time index m as $\hat{\mathbf{x}}_{n|m}$. The prediction of the new state is given by

$$\hat{\mathbf{x}}_{k|k-1} = \mathbb{E}\{f(\mathbf{x}_{k-1}, \mathbf{u}_k, \mathbf{v}_k)\}, \tag{17}$$

with a prediction error covariance

$$\mathbf{P}_{k|k-1} = \mathbb{E}\{(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})^T\}. \tag{18}$$

When an observation of the system is made available, it is used to improve the estimate. The update is given by

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k(\mathbf{y}_k - \hat{\mathbf{y}}_k), \tag{19}$$

where $\mathbf{K}_k$ is the *Kalman gain* and the observation is predicted by

$$\hat{\mathbf{y}}_k = \mathbb{E}\{h(\hat{\mathbf{x}}_{k|k-1}, \mathbf{n}_k)\}. \tag{20}$$

The Kalman gain is computed such that the mean squared error $\mathbb{E}\{|\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}|^2\}$ is minimized. This corresponds to minimizing the trace of the covariance matrix $\mathbf{P}_{k|k} = \mathbb{E}\{(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})^T\}$. Using equation (19), the covariance matrix can be written as

$$\begin{aligned}\mathbf{P}_{k|k} &= \mathbb{E}\{(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1} - \mathbf{K}_k(\mathbf{y}_k - \hat{\mathbf{y}}_k))(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1} - \mathbf{K}_k(\mathbf{y}_k - \hat{\mathbf{y}}_k))^T\} \\ &= \mathbf{P}_{k|k-1} - \mathbf{P}_{\mathbf{xy}}\mathbf{K}_k^T - \mathbf{K}_k\mathbf{P}_{\mathbf{xy}}^T + \mathbf{K}_k\mathbf{P}_{\mathbf{y}}\mathbf{K}_k^T, \end{aligned} \tag{21}$$

where $\mathbf{P}_{\mathbf{xy}} = \mathbb{E}\{(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})(\mathbf{y}_k - \hat{\mathbf{y}}_k)^T\}$ is the cross-correlation matrix and $\mathbf{P}_{\mathbf{y}} = \mathbb{E}\{(\mathbf{y}_k - \hat{\mathbf{y}}_k)(\mathbf{y}_k - \hat{\mathbf{y}}_k)^T\}$ is the observation covariance. The trace is given by

$$\mathbb{E}\{|\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}|^2\} = \mathrm{tr}(\mathbf{P}_{k|k}) = \mathrm{tr}(\mathbf{P}_{k|k-1}) - 2\mathrm{tr}(\mathbf{P}_{\mathbf{xy}}^T\mathbf{K}_k) + \mathrm{tr}(\mathbf{K}_k\mathbf{P}_{\mathbf{y}}\mathbf{K}_k). \tag{22}$$

Taking the derivative of the trace with respect to the Kalman gain $\mathbf{K}_k$ and setting it to zero gives

$$\begin{aligned}\frac{\partial}{\partial \mathbf{K}_k}\mathrm{tr}(\mathbf{P}_{k|k}) &= -2\mathbf{P}_{\mathbf{xy}} + 2\mathbf{K}_k\mathbf{P}_{\mathbf{y}} = 0 \\ \text{or} \qquad \mathbf{K}_k &= \mathbf{P}_{\mathbf{xy}}\mathbf{P}_{\mathbf{y}}^{-1}. \end{aligned} \tag{23}$$

Using Kalman gain given in equation (23), we can write the update for the error covariance in equation (21) as

$$\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1} - \mathbf{K}_k\mathbf{P}_{\mathbf{xy}}^T. \tag{24}$$

For other Kalman gains the form of equation (21) has to be used.

Original Kalman filter

The original Kalman filter published by Kálman [8] assumes that the state transition and observation can be evaluated with linear functions. The state transition function is now given by the following difference equation

$$\mathbf{x}_k = \mathbf{A}_k\mathbf{x}_{k-1} + \mathbf{B}_k\mathbf{u}_k + \mathbf{v}_k \tag{25}$$

and the observation is derived from the current state as

$$\mathbf{y}_k = \mathbf{H}_k\mathbf{x}_k + \mathbf{n}_k, \tag{26}$$

where $\mathbf{A}_k$ is the state transition matrix, $\mathbf{B}_k$ is the control matrix and $\mathbf{H}_k$ is the observation matrix.

When the dynamic model is given by equations (25) and (26), the state prediction is given by

$$\begin{aligned}\hat{\mathbf{x}}_{k|k-1} &= \mathbb{E}\{\mathbf{A}_k\mathbf{x}_{k-1} + \mathbf{B}_k\mathbf{u}_k + \mathbf{v}_k\} = \mathbf{A}_k\mathbb{E}\{\mathbf{x}_{k-1}\} + \mathbf{B}_k\mathbf{u}_k \\ &= \mathbf{A}_k\hat{\mathbf{x}}_{k-1|k-1} + \mathbf{B}_k\mathbf{u}_k. \end{aligned} \tag{27}$$

The covariance matrix of prediction error can be written as

$$\begin{aligned}\mathbf{P}_{k|k-1} &= \mathbb{E}\{(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})^T\} \\ &= \mathbb{E}\{(\mathbf{A}_k(\mathbf{x}_{k-1} - \hat{\mathbf{x}}_{k-1|k-1}) + \mathbf{v}_k)(\mathbf{A}_k(\mathbf{x}_{k-1} - \hat{\mathbf{x}}_{k-1|k-1}) + \mathbf{v}_k)^T\} \\ &= \mathbf{A}_k\mathbf{P}_{k-1|k-1}\mathbf{A}_k^T + \mathbf{Q}_k. \end{aligned} \tag{28}$$

After the observation is made available, the update of the state estimate is

$$\begin{aligned}\hat{\mathbf{x}}_{k|k} &= \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k(\mathbf{y}_k - \hat{\mathbf{y}}_k) \\ &= \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k(\mathbf{y}_k - \mathbf{H}_k\hat{\mathbf{x}}_{k|k-1}) \end{aligned} \tag{29}$$

and the estimation error covariance is

$$\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1} - \mathbf{K}\mathbf{P}_{\mathbf{xy}}^T = (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}_{k|k-1}, \tag{30}$$

where the cross-correlation matrix is given by

$$\begin{aligned}\mathbf{P}_{\mathbf{xy}} &= \mathbb{E}\{(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})(\mathbf{y}_k - \hat{\mathbf{y}}_k)^T\} \\ &= \mathbb{E}\{(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})^T\mathbf{H}_k^T\} = \mathbf{P}_{k|k-1}\mathbf{H}_k^T. \end{aligned} \tag{31}$$

The optimal Kalman gain is given by

$$\mathbf{K}_k = \mathbf{P}_{\mathbf{xy}}\mathbf{P}_{\mathbf{y}}^{-1} = \mathbf{P}_{k|k-1}\mathbf{H}_k^T(\mathbb{E}\{(\mathbf{y}_k - \hat{\mathbf{y}}_k)(\mathbf{y}_k - \hat{\mathbf{y}}_k)^T\})^{-1}$$

$$= \mathbf{P}_{k|k-1}\mathbf{H}_k^T(\mathbb{E}\{(\mathbf{H}_k(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}) + \mathbf{n}_k)(\mathbf{H}_k(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}) + \mathbf{n}_k)^T\})^{-1}$$
$$= \mathbf{P}_{k|k-1}\mathbf{H}_k^T(\mathbf{H}_k\mathbf{P}_{k|k-1}\mathbf{H}_k^T + \mathbf{R}_k)^{-1}. \quad (32)$$

The algorithm is initialized by assigning $\hat{\mathbf{x}}_{0|0} = \mathbb{E}\{\mathbf{x}_0\}$ and $\mathbf{P}_{0|0} = \mathbb{E}\{(\mathbf{x}_0 - \hat{\mathbf{x}}_{0|0})(\mathbf{x}_0 - \hat{\mathbf{x}}_{0|0})^T\}$. If the linear state transition and observation models are accurate and the noise statistics are known, the estimate $\hat{\mathbf{x}}_{k|k}$ of the state $\mathbf{x}_k$ minimizes the mean squared error [7].

Unscented Kalman filter

The original Kalman filter assumes that the system model can be accurately modeled with linear functions. The unscented Kalman filter has been developed to apply the linear estimation scheme on systems that exhibit non-linear state transition or observation models. The unscented Kalman filter uses a sampling technique called *scaled unscented transform* [9] to estimate the true mean and covariance. A minimum number of points, called *sigma points*, around the estimated mean are selected. These sigma points are fed through the non-linear system model and the weighted ensemble mean and covariance are used to estimate the true mean and covariance matrix. In this thesis the unscented Kalman filter notation follows the scheme introduced by Van Der Merwe in his dissertation [7].

The state of the stochastic process now evolves according to a non-linear function

$$\mathbf{x}_k = f(\mathbf{x}_{k-1}, \mathbf{u}_k, \mathbf{v}_k) \quad (33)$$

and the observations are given by

$$\mathbf{y}_k = h(\mathbf{x}_k, \mathbf{n}_k). \quad (34)$$

The filters state at each time step k is given by augmented versions of the state vector and covariance matrix, where the process and observation noises are incorporated. The augmented state estimate vector is given by

$$\hat{\mathbf{x}}^a_{k|k} = \begin{bmatrix} \hat{\mathbf{x}}_{k|k} \\ \mathbb{E}\{\mathbf{v}_k\} \\ \mathbb{E}\{\mathbf{n}_k\} \end{bmatrix} \quad (35)$$

and the augmented covariance matrix is given by

$$\mathbf{P}^a_{k|k} = \begin{bmatrix} \mathbf{P}_{k|k} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_k & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{R}_k \end{bmatrix}. \quad (36)$$

Let the dimension of augmented state vector be L. The sigma points around the estimated mean of state vector are selected according to

$$\begin{aligned}\mathcal{X}^a_{k-1|k-1} &= \left[\hat{\mathbf{x}}^a_{k-1|k-1},\ \hat{\mathbf{x}}^a_{k-1|k-1} + \gamma\sqrt{\mathbf{P}^a_{k-1|k-1}},\ \hat{\mathbf{x}}^a_{k-1|k-1} - \gamma\sqrt{\mathbf{P}^a_{k-1|k-1}}\right] \\ &= \begin{bmatrix}\mathcal{X}^x_{k-1|k-1} \\ \mathcal{X}^v_{k-1|k-1} \\ \mathcal{X}^n_{k-1|k-1}\end{bmatrix}, \end{aligned} \tag{37}$$

where the square root denotes the Cholesky decomposition and

$$\gamma = \sqrt{L + \lambda}. \tag{38}$$

Here λ is a scaling parameter given by $\lambda = \alpha^2(L + \kappa) - L$, where α is a parameter for how far from the mean the sigma points are selected and κ is selected to ensure positive definiteness of covariance matrix.

The weights for computing means and covariances are given by

$$W^m_0 = \frac{\lambda}{L + \lambda}, \tag{39}$$

$$W^c_0 = \frac{\lambda}{L + \lambda} + 1 - \alpha^2 + \beta \quad \text{and} \tag{40}$$

$$W^m_i = W^c_i = \frac{1}{2(L + \lambda)} \quad i = 1, \ldots, 2L. \tag{41}$$

The superscript denotes whether the weight term is used for calculating the mean or covariance. The extra term in covariance weight (40) is used to help minimize higher order errors in the estimated covariance by selecting a suitable value for parameter β. For Gaussian distributed state vector the optimal value is known to be $\beta = 2$ [9].

In addition, the unscented Kalman filter works in two phases. First, a prediction phase is carried out to give the prediction of the state. The sigma points are passed through the state transition function

$$\mathcal{X}^x_{k|k-1} = f(\mathcal{X}^x_{k-1|k-1}, \mathbf{u}_k, \mathcal{X}^v_{k-1|k-1}) \tag{42}$$

and the predicted state and prediction error covariance are estimated with weighted ensemble mean and covariance as

$$\hat{\mathbf{x}}_{k|k-1} = \sum_{i=0}^{2L} W^m_i \mathcal{X}^x_{i,k|k-1} \tag{43}$$

and

$$\mathbf{P}_{k|k-1} = \sum_{i=0}^{2L} W^c_i \left(\mathcal{X}^x_{i,k|k-1} - \hat{\mathbf{x}}_{k|k-1}\right)\left(\mathcal{X}^x_{k|k-1} - \hat{\mathbf{x}}_{k|k-1}\right)^T. \tag{44}$$

In the update phase, the transitioned sigma points are propagated through the observation model and again weighted ensemble mean and covariance are used to estimate the cross-correlation and observation covariance. The observation sigma points are given by

$$\mathcal{Y}_k = h(\mathcal{X}^x_{k|k-1}, \mathcal{X}^n_{k-1|k-1}), \tag{45}$$

from which the observation estimate is given by

$$\hat{\mathbf{y}}_k = \sum_{i=0}^{2L} W_i^m \mathcal{Y}_{i,k}, \tag{46}$$

the observation covariance

$$\mathbf{P}_{\mathbf{y}} = \sum_{i=0}^{2L} W_i^c \left(\mathcal{Y}_{i,k} - \hat{\mathbf{y}}_k\right) \left(\mathcal{Y}_{i,k} - \hat{\mathbf{y}}_k\right)^T \tag{47}$$

and the cross-correlation matrix is

$$\mathbf{P}_{\mathbf{xy}} = \sum_{i=0}^{2L} W_i^c \left(\mathcal{X}_{i,k|k-1} - \hat{\mathbf{x}}_{k|k-1}\right) \left(\mathcal{Y}_{i,k} - \hat{\mathbf{y}}_k\right)^T. \tag{48}$$

The update to the state vector and error covariance are given by the equations (19), (24) and (23) as derived earlier. Here they are given again

$$\mathbf{K}_k = \mathbf{P}_{\mathbf{xy}} \mathbf{P}_{\mathbf{y}}^{-1}, \tag{49}$$

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k(\mathbf{y}_k - \hat{\mathbf{y}}_k) \quad \text{and} \tag{50}$$

$$\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1} - \mathbf{K}_k \mathbf{P}_{\mathbf{xy}}^T. \tag{51}$$

2 Orthogonal frequency-division multiplexing system

Orthogonal frequency-division multiplexing (OFDM) is a digital modulation method where the transmitted symbol carries multiple data symbols simultaneously. The low rate data is used to modulate sub-carriers that are orthogonal in a sense that transmission over one sub-carrier does not affect the transmission over adjacent sub-carriers. Modulation and demodulation is accomplished by taking a discrete Fourier transform of symbol samples.

The advantages of using OFDM modulation is that it is robust against multipath fading and demodulation can be done without complex equalization. The spectral efficiency is good because no guard interval is used between sub-carriers in frequency and it is computationally efficient because of

discrete Fourier transform can be computed by a fast Fourier transform algorithm. The use of cyclic prefix makes the modulation scheme robust against timing synchronization errors.

The challenges of OFDM are its sensitivity to frequency synchronization errors and Doppler shift. Imperfect synchronization results in loss of orthogonality which causes phase rotation and attenuation of the signal, as well as introduces additional noise due to inter-carrier interference (ICI).

2.1 Transmitter

The transmitter is responsible for multiplexing different logical channels, modulation and radiating the signal from antennas. The OFDM symbol is a collection of orthogonal information bearing sub-carriers summed together.

Definition 2.1. *Orthogonal frequency-division multiplexing symbol* carrying a block of symbols $\mathbf{c} = [c_{-\frac{N}{2}}, c_{-\frac{N}{2}+1}, \ldots, c_{\frac{N}{2}-1}]^T$ is given by

$$s(t) = \frac{1}{\sqrt{N}} \sum_{k=-\frac{N}{2}}^{\frac{N}{2}-1} c_k e^{j2\pi f_k t}, \quad 0 \leq t < T_u, \tag{52}$$

where $e^{j2\pi f_k t} = w_{f_k}(t)$ is the sub-carrier waveform, $f_k = \frac{k}{T_u}$ is the sub-carrier frequency, T_u is the symbol duration and N is the number of sub-carriers. We assume that N is a power of two and sub-carrier index zero denotes the center sub-carrier.

The orthogonality means that even though the frequency spectrum of the sub-carriers overlap they do not contribute energy to the adjacent sub-carriers in demodulation. The orthogonality is proved by following theorem.

Theorem 2.2. *Let the effective OFDM symbol duration be T_u. Then the sub-carrier waveforms $w_f(t) = e^{j2\pi f t}$ are orthogonal if the spacing is a multiple of $\frac{1}{T_u}$.*

Proof. Two signals, $f, g : \mathbb{R} \to \mathbb{C}$, $f \neq g$, are orthogonal on the interval $[t_0, t_0 + T_u]$ if $\langle f, g \rangle_{[t_0, t_0+T_u]} = \frac{1}{T_u} \int_{t_0}^{t_0+T_u} f(t) g^*(t) dt = 0$. Let f_1 and f_2 denote the two sub-carrier frequencies such that $f_1 - f_2 = \frac{k}{T_u}$, $k \in \mathbb{Z}$. Then

$$\begin{aligned}
&\langle w_{f_1}(t), w_{f_2}(t) \rangle_{[t_0, t_0+T_u]} = \frac{1}{T_u} \int_{t_0}^{t_0+T_u} e^{j2\pi f_1 t} e^{-j2\pi f_2 t} dt \\
&= \frac{1}{T_u} \int_{t_0}^{t_0+T_u} e^{j2\pi (f_1 - f_2) t} dt = \frac{1}{T_u} \int_{t_0}^{t_0+T_u} e^{j2\pi k t / T_u} dt = \frac{e^{j2\pi k t / T_u}}{j2\pi k} \Big|_{t=t_0}^{t_0+T_u}
\end{aligned}$$

$$= \frac{e^{j2\pi kt_0/T_u}e^{j2\pi k} - e^{j2\pi kt_0/T_u}}{j2\pi k} = e^{j2\pi kt_0/T_u}e^{j\pi k}\frac{e^{j\pi k} - e^{-j\pi k}}{j2\pi k}$$
$$= e^{j\pi k(1+2t_0/T_u)}\frac{\sin(\pi k)}{\pi k} = e^{j\pi k(1+2t_0/T_u)}\mathrm{sinc}(\pi k) = \begin{cases} 1, & k = 0, \\ 0, & k \neq 0. \end{cases}$$

□

With this property the demodulation of sub-carrier k can be carried out by correlating the OFDM symbol $s(t)$ with the sub-carrier waveform conjugate as

$$\frac{\sqrt{N}}{T_u}\int_0^{T_u} s(t)w_{f_k}^*(t)dt = \frac{\sqrt{N}}{T_u}\int_0^{T_u}\frac{1}{\sqrt{N}}\sum_{k'=-\frac{N}{2}}^{\frac{N}{2}-1} c_{k'}w_{f_{k'}}(t)w_{f_k}^*(t)dt$$
$$= \sum_{k'=-\frac{N}{2}}^{\frac{N}{2}-1} c_{k'}\left\langle w_{f_{k'}}(t), w_{f_k}(t)\right\rangle = c_k. \tag{53}$$

Generating OFDM symbol using DFT

The OFDM symbols are generated by multiplexing data symbols $d_{l,k}$ and reference symbols $x_{l,k}$ in time and frequency to OFDM symbol l and sub-carrier k and forwarded to an inverse discrete Fourier transform. The resources to be mapped to sub-carrier k in OFDM symbol l are given by

$$c_{l,k} = \begin{cases} d_{l,k}, & \text{if } (l,k) \in \mathcal{D}, \\ x_{l,k}, & \text{if } (l,k) \in \mathcal{P}, \\ 0, & \text{else}, \end{cases} \tag{54}$$

where the sets $\mathcal{D}$ and $\mathcal{P}$ are the index sets containing data symbol and reference symbol locations respectively. The input block to the size N inverse discrete Fourier transforms is put into vector form as $\mathbf{c}_l = \Big[c_{l,-\frac{N}{2}}, c_{l,-\frac{N}{2}+1}, \ldots, c_{l,\frac{N}{2}-1}\Big]^T$. The output of inverse Fourier transform is the baseband OFDM symbol block

$$\mathbf{s}_l = [s_{l,0}, s_{l,1}, \ldots, s_{l,N-1}]^T = \mathbf{F}^H\mathbf{c}_l, \tag{55}$$

where

$$\mathbf{F} = \frac{1}{\sqrt{N}}\begin{bmatrix} 1 & 1 & \cdots & 1 \\ e^{-j2\pi\left(-\frac{N}{2}\right)\frac{1}{N}} & e^{-j2\pi\left(-\frac{N}{2}+1\right)\frac{1}{N}} & \ldots & e^{-j2\pi\left(\frac{N}{2}-1\right)\frac{1}{N}} \\ \vdots & \vdots & \ddots & \vdots \\ e^{-j2\pi\left(-\frac{N}{2}\right)\frac{N-1}{N}} & e^{-j2\pi\left(-\frac{N}{2}+1\right)\frac{N-1}{N}} & \ldots & e^{-j2\pi\left(\frac{N}{2}-1\right)\frac{N-1}{N}} \end{bmatrix}_{N\times N} \tag{56}$$

is the discrete Fourier transform matrix of size N. The elements of $\mathbf{s}_l$ are then given by

$$s_{l,n} = \frac{1}{\sqrt{N}} \sum_{k=-\frac{N}{2}}^{\frac{N}{2}-1} c_{l,k} e^{j2\pi \frac{kn}{N}}. \tag{57}$$

In the beginning of OFDM symbol block a cyclic prefix is added which is a copy of the last N_g samples. The operation can be carried out by multiplying OFDM symbol block with matrix

$$\mathbf{C} = \begin{bmatrix} \mathbf{0}_{N_g \times (N-N_g)} & \mathbf{I}_{N_g \times N_g} \\ & \mathbf{I}_{N \times N} \end{bmatrix} \tag{58}$$

and we get

$$\mathbf{s}_{\mathrm{cp},l} = \mathbf{C}\mathbf{s}_l = [s_{l,N-N_g}, \ldots, s_{l,N-1}, s_{l,0}, \ldots, s_{l,N-1}]^T. \tag{59}$$

Cyclic prefix is used to eliminate inter-symbol interference caused by multipath propagation, where a channel consists of multiple propagation paths that arrive at different times to the receiver. Last delayed samples of a symbol arrive at the same time as the first samples of the next symbol causing interference but the prefix acts as a buffer zone between two symbols. Additionally the cyclic prefix helps to reduce the problems due to time synchronisation errors because a cyclic shift in discrete Fourier transform input corresponds to linear phase shift in output.

After applying the prefix, the OFDM symbol samples are put to serial form and converted from digital to analogue domain, see Theorem 1.5. Let T be the sample time interval and let $T_u = NT$, $T_g = N_g T$ and $T_s = (N+N_g)T$ be the effective symbol duration, prefix duration and the duration of the complete OFDM symbol. The baseband OFDM symbol stream is given by [10]

$$s(t) = \frac{1}{\sqrt{N}} \sum_{l=-\infty}^{\infty} \sum_{k=-\frac{N}{2}}^{\frac{N}{2}-1} c_{l,k} e^{j2\pi \frac{k}{T_u}(t-T_g-lT_s)} u(t-lT_s), \tag{60}$$

where

$$u(t) = \begin{cases} 1, & 0 < t < T_s \\ 0, & \text{else} \end{cases}$$

is a rectangular windowing function. Then, the signal is divided into real and imaginary part and quadrature mixed to passband frequency and radiated through antennas. The quadrature-mixed signal is given by

$$s_c(t) = \mathrm{Re}(s(t)) \cos(2\pi f_c t) - \mathrm{Im}(s(t)) \sin(2\pi f_c t), \tag{61}$$

where f_c is the carrier frequency. It is worth mentioning that the signal $s_c(t)$ is real valued.

2.2 Channel

The signal has to pass through a radio channel. In this work, the channel is assumed a *Rayleigh fading multipath channel*. The channel consists of multiple physical propagation paths, which cause the receiver to see the superposition of multiple instances of the sent waveform that arrive at different times and are attenuated and phase shifted differently. The movement of user equipment, base station or reflecting surfaces lead to Doppler effect that causes frequency shifts to the different received instances of the signal. The transmission conditions might vary with time and frequency. The time dependent change in channel is called *time selectivity* or *fading* and frequency dependent change is called *frequency selectivity*. In addition, additional noise from surrounding environment is introduced to the signal.

The channel can be represented in time domain by its impulse response $h(\tau, t)$ which is a complex valued function of path delay τ and time t. A completely analogous representation is the response in frequency domain $H(f, t)$, which relates to the impulse response such that they are Fourier transform pairs, $H(f,t) = \int_{-\infty}^{\infty} h(\tau, t) e^{-j2\pi\tau f} d\tau$. The channel response $h(\tau, t)$ is applied to a signal $s(t)$ by convolution as $r(t) = (h(\tau, t) * s(t))(t) = \int_{-\infty}^{\infty} h(\tau, t) s(t - \tau) d\tau$. Because Fourier transform of a convolution of two signals is equal to the product of Fourier transforms of the signals, we get that $R(f) = \mathcal{F}(r(t)) = H(f, t) \cdot S(f)$, where $S(f) = \mathcal{F}(s(t))$ and $R(f) = \mathcal{F}(r(t))$. Therefore, the channel effect can be applied by convolution in time domain or by multiplication in frequency domain.

Useful correlation functions

The channel impulse response $h(\tau, t)$ is assumed to be zero mean wide-sense-stationary stochastic process in the variable t. The behavior of the channel can be characterized by the following covariance functions [11].

Definition 2.3. The channel's time domain covariance function is defined as

$$\begin{aligned} c(\tau_1, \tau_2, t_1, t_2) =& \mathbb{E}\left\{ \left(h(\tau_1, t_1) - \mathbb{E}\{h(\tau_1, t_1)\} \right)^* \left(h(\tau_2, t_2) - \mathbb{E}\{h(\tau_2, t_2)\} \right) \right\} \\ =& \mathbb{E}\{h^*(\tau_1, t_1) h(\tau_2, t_2)\} = \mathbb{E}\{h^*(\tau_1, 0) h(\tau_2, t_2 - t_1)\} \\ \equiv& c(\tau_1, \tau_2, \Delta t), \end{aligned} \tag{62}$$

where $\Delta t = t_2 - t_1$. We make the assumption that the propagation path with delay τ_1 is uncorrelated with the propagation path with delay $\tau_2 \neq \tau_1$ so

$$c(\tau_1, \tau_2, \Delta t) = c(\tau_1, \tau_2, \Delta t)\delta(\tau_2 - \tau_1) \equiv c(\tau, \Delta t), \tag{63}$$

where $\delta(t)$ is the Dirac delta function. This assumption is valid because signals arriving at different times might have travelled a completely different path to the receiver. If we set $\Delta t = 0$, the covariance function gives the average power output of the channel as a function of delay τ. That is why the function $c(\tau, 0)$ is called *the delay power spectrum* of the channel.

Definition 2.4. The frequency domain covariance function is defined as

$$\begin{aligned}
C(f_1, f_2, t_1, t_2) =& \mathbb{E}\left\{(H(f_1, t_1) - \mathbb{E}\{H(f_1, t_1)\})^* (H(f_2, t_2) - \mathbb{E}\{H(f_2, t_2)\})\right\} \\
=& \mathbb{E}\left\{H^*(f_1, t_1)H(f_2, t_2)\right\} \\
=& \mathbb{E}\left\{\int_{-\infty}^{\infty} h^*(\tau_1, t_1)e^{-j2\pi\tau_1 f_1}d\tau_1 \int_{-\infty}^{\infty} h(\tau_2, t_2)e^{j2\pi\tau_2 f_2}d\tau_2\right\} \\
=& \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{j2\pi(\tau_2 f_2 - \tau_1 f_1)}\mathbb{E}\{h^*(\tau_1, t_1)h(\tau_2, t_2)\}d\tau_1 d\tau_2 \\
=& \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{j2\pi(\tau_2 f_2 - \tau_1 f_1)}c(\tau_1, \tau_2, 0, t_2 - t_1)\delta(\tau_2 - \tau_1)d\tau_1 d\tau_2 \\
=& \int_{-\infty}^{\infty} e^{j2\pi(f_2 - f_1)\tau}c(\tau, \Delta t)d\tau \equiv C(\Delta f, \Delta t).
\end{aligned}$$

The covariance function in frequency domain is seen to be the Fourier transform of the time domain covariance function. The function $C(\Delta f, \Delta t)$ is called *spaced-time, spaced-frequency correlation function* and it gives a measure for coherence time and coherence bandwidth of the channel.

Setting $\Delta t = 0$ in $C(\Delta f, \Delta t)$ gives the spaced frequency correlation function $C(\Delta f, 0)$ which relates to the delay power spectrum $c(\tau, 0)$ such that they are Fourier transform pairs. Therefore the reciprocal of the width of the delay power spectrum T_τ gives a measure of the *coherence bandwidth* of the channel $(\Delta f)_c$. A coarse measure is given by $(\Delta f)_c \approx \frac{1}{T_\tau}$. The meaning of the coherence bandwidth is that two signals that are spaced further than the coherence bandwidth $(\Delta f)_c$ are affected differently by the channel, when signals spaced less than coherence bandwidth are affected similarly by the channel.

Taking the Fourier transform of spaced-time, spaced-frequency correlation function with respect to Δt gives the function $S(\Delta f, \lambda)$. If we set $\Delta f = 0$ we get the function $S(0, \lambda)$ which gives the signal's intensity as a function of the Doppler frequency λ. Therefore the function $S(0, \lambda)$ is called the *Doppler power spectrum* of the channel. The Doppler power spectrum relates to the

spaced-time correlation function such that they are Fourier transform pairs. The Doppler spread B_d is the range over which the Doppler power spectrum is non-zero and its reciprocal gives a measure to the *coherence time* of the channel $(\Delta t)_c$. A coarse measure is $(\Delta t)_c \approx \frac{1}{B_d}$.

Two more Fourier transform pairs can be found by defining a function $s(\tau, \lambda)$ called the *scattering function* of the channel. It is the Fourier transform pair of the covariance function $c(\tau, \Delta t)$ with respect to Δt. It is also the Fourier transform pair of the function $S(\Delta f, \lambda)$ with respect to the variable Δf. It gives the average power output of the channel as a function of the delay τ and Doppler frequency λ.

Rayleigh fading

In this work, the channel is assumed to exhibit *Rayleigh fading*. The impulse response $h(\tau, t)$ is assumed to have zero mean and phase evenly distributed between 0 and 2π. The envelope of the impulse response is then Rayleigh distributed. The probability density function of the Rayleigh distribution is

$$f(x) = \frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}}. \tag{64}$$

The scale parameter σ is given by the average power of the impulse response, $\sigma^2 = \mathbb{E}_t\{|h(\tau, t)|^2\}$.

Because the arrival of different propagation paths are grouped into clusters, the channel impulse usually is essentially non-zero at the delays τ_i, where τ_i is the delay of the cluster i. Then the channel impulse response can be approximated by $h(\tau, t) = \sum_{i=1}^{L} h_i(t)\delta(\tau_i - \tau)$, where L is the number of significant clusters. This simplification introduces modelling error, which is neglected in this work.

Example 2.5. Here we illustrate the effect of channel on the signal, when channel impulse response is given by $h(\tau, t) = \sum_{i=1}^{L} h_i(t)\delta(\tau_i - \tau)$. The channel impulse is applied to a signal by convolution as

$$\begin{aligned} h(\tau, t) * s(t) &= \int_{-\infty}^{\infty} h(\tau, t)s(t - \tau)d\tau = \int_{-\infty}^{\infty} \sum_{i=1}^{L} h_i(t)\delta(\tau - \tau_i)s(t - \tau)d\tau \\ &= \sum_{i=1}^{L} h_i(t) \int_{-\infty}^{\infty} \delta(\tau - \tau_i)s(t - \tau)d\tau = \sum_{i=1}^{L} h_i(t)s(t - \tau_i). \end{aligned}$$

2.3 Receiver

The receiver is responsible for capturing the signal, compensating for distortions, demodulation, de-multiplexing and passing received samples to de-

coder. After capturing the signal with antennas, the radio frequency signal is band pass filtered to reject unwanted frequencies and given by

$$r_c(t) = \sum_{i=0}^{L} h_i'(t) s_c(t - \tau_i) + w_c(t), \tag{65}$$

where h_i' and w_c are real valued channel response and noise. We assume that the channel impulse response is constant during one OFDM symbol and the duration $\tau_L - \tau_1$ is shorter than the cyclic prefix T_g.

Quadrature-mixing down to baseband frequency is accomplished by first multiplying the received signal with carrier frequency sinusoids that are separated by a $\frac{\pi}{2}$ phase shift. In this work, any imbalance between in-phase and quadrature components is assumed to be ideally removed. Using equation (61), the in-phase component is given by

$$\begin{aligned}
I(t) =& r_c(t)\cos(2\pi f_c t) = \sum_{i=1}^{L} h_i'(t) s_c(t-\tau_i)\cos(2\pi f_c t) + w_c(t)\cos(2\pi f_c t) \\
=& \sum_{i=1}^{L} h_i'(t)[\mathrm{Re}(s(t-\tau_i))\cos(2\pi f_c(t-\tau_i))\cos(2\pi f_c t) \\
& - \mathrm{Im}(s(t-\tau_i))\sin(2\pi f_c(t-\tau_i))\cos(2\pi f_c t)] + w_c(t)\cos(2\pi f_c t) \\
=& \sum_{i=1}^{L} \frac{h_i'(t)}{2}[\mathrm{Re}(s(t-\tau_i))(\cos(-2\pi f_c\tau_i) + \cos(2\pi f_c(2t-\tau_i))) \\
& - \mathrm{Im}(s(t-\tau_i))(\sin(-2\pi f_c\tau_i) + \sin(2\pi f_c(2t-\tau_i)))] + w_c(t)\cos(2\pi f_c t).
\end{aligned} \tag{66}$$

The signal is low-pass filtered to reject the double frequency and suitable amplification is used to get

$$\begin{aligned}
I(t) =& \sum_{i=1}^{L} h_i'(t)\,[\mathrm{Re}(s(t-\tau_i))\cos(-2\pi f_c\tau_i) \\
& - \mathrm{Im}(s(t-\tau_i))\sin(-2\pi f_c\tau_i)] + w'(t),
\end{aligned} \tag{67}$$

where $w'(t)$ is filtered noise. Similarly, for quadrature part, we get after filtering and amplification

$$\begin{aligned}
Q(t) =& \sum_{i=1}^{L} h_i'(t)\,[\mathrm{Re}(s(t-\tau_i))\sin(-2\pi f_c\tau_i) \\
& + \mathrm{Im}(s(t-\tau_i))\cos(-2\pi f_c\tau_i)] + w''(t).
\end{aligned} \tag{68}$$

The quadrature-mixer outputs, in-phase and quadrature component, are next combined to get the received baseband OFDM symbol stream. The signal is given by

$$\begin{aligned} r(t) =& I(t) + jQ(t) = \sum_{i=1}^{L} h_i'(t)[\mathrm{Re}(s(t-\tau_i))(\cos(-2\pi f_c \tau_i) + j\sin(-2\pi f_c \tau_i)) \\ &+ j\,\mathrm{Im}(s(t-\tau_i))(\cos(-2\pi f_c \tau_i) + j\sin(-2\pi f_c \tau_i))] + w'(t) + jw''(t) \\ =& \sum_{i=1}^{L} h_i'(t) e^{-j2\pi f_c \tau_i} s(t-\tau_i) + w(t) \\ =& \sum_{i=1}^{L} h_i(t) s(t-\tau_i) + w(t), \end{aligned} \tag{69}$$

where $h_i(t) = h_i'(t)e^{-j2\pi f_c \tau_i}$ is the complex valued channel impulse response and $w(t) = w'(t) + jw''(t)$ complex valued noise. The delays of the different propagation paths are unknown and the carrier frequency is usually very big, so the phase of the channel impulse response can be regarded uniformly distributed. If the real valued channel response has a Rayleigh distribution, then $h_i(t)$ exhibits Rayleigh fading. The noise $w(t)$ is white Gaussian noise because the real and imaginary components are uncorrelated with equal power. This results from the fact that the in-phase and quadrature waveforms are orthogonal.

The samples of OFDM symbol l are recovered from the received symbol stream by sampling at time indices $t_n = nT + T_g + lT_s$, $n \in \{0, 1, \ldots, N-1\}$. Here, the time variable t is advanced so that $t = 0$ denotes the time instant when the first sample of OFDM symbol 0 has reached the receiver. Then τ_i denotes the relative delay of path i compared to the first path and $\tau_1 = 0$. Because the channel impulse is assumed constant during one OFDM symbol, the index n can be omitted and $h_i(nT + T_g + lT_s) = h_i(l)$. The sampling stage incorporates the removal of cyclic prefix. The received time domain OFDM symbol block l is given by

$$\mathbf{r}_l = [r_{l,0}, r_{l,1}, \ldots, r_{l,N-1}]^T, \tag{70}$$

where

$$\begin{aligned} r_{l,n} =& \sum_{i=1}^{L} h_i(l) s(nT + T_g + lT_s - \tau_i) + w(nT + T_g + lT_s) \\ =& \sum_{i=1}^{L} h_i(l) \frac{1}{\sqrt{N}} \sum_{l'=-\infty}^{\infty} \sum_{k'=-\frac{N}{2}}^{\frac{N}{2}-1} c_{l',k'} e^{j2\pi \frac{k'}{T_u}(nT - \tau_i + (l-l')T_s)} \end{aligned}$$

$$\cdot\, u(nT + T_g - \tau_i + (l - l')T_s) + w(t_n). \tag{71}$$

The windowing function evaluates to unit only for $l' = l$ and the delayed paths of the last samples of the OFDM symbol that would fall outside sampling window are recovered from the delayed parts of the cyclic prefix. Then

$$\begin{aligned} r_{l,n} =& \frac{1}{\sqrt{N}} \sum_{k'=-\frac{N}{2}}^{\frac{N}{2}-1} c_{l,k'} e^{j2\pi k' n/N} \sum_{i=1}^{L} h_i(l) e^{-j2\pi \frac{\tau_i}{T} k'/N} + w(t_n) \\ =& \frac{1}{\sqrt{N}} \sum_{k'=-\frac{N}{2}}^{\frac{N}{2}-1} c_{l,k'} H_{l,k'} e^{j2\pi k' n/N} + w(t_n). \end{aligned} \tag{72}$$

Then the received OFDM symbol block can be written as

$$\mathbf{r}_l = \mathbf{F}^H \mathbf{H}_l \mathbf{c}_l + \mathbf{w}_l, \tag{73}$$

where $\mathbf{H}$ is the diagonal channel frequency response matrix and $\mathbf{w}_l$ is a vector of noise samples.

The demodulation of received samples is done by calculating the discrete Fourier transform. For the received symbol samples $\mathbf{y}_l$ we get

$$\mathbf{y}_l = \mathbf{F}\mathbf{r}_l = \mathbf{H}_l \mathbf{c}_l + \mathbf{F}\mathbf{w}_l = \mathbf{H}_l \mathbf{c}_l + \mathbf{W}_l. \tag{74}$$

After demodulation, each sub-carrier experiences frequency-flat channel because the channel matrix is diagonal. For example, zero-forcing or minimum mean squared error equalizer can be used to retrieve the transmitted data symbols [12]. Equalized data samples are then detected and forwarded to encoding to find the transmitted information.

2.4 Symbol timing and frequency offsets

The transmitter and receiver do not share a common frequency and timing reference, so a synchronization task has to be performed before demodulation can be carried out. The synchronization process is divided into two parts. First, an acquisition phase is carried out, where the frequency and timing are acquired without any prior knowledge with sufficient accuracy. Then, during the data reception, the frequency and timing are constantly being monitored and corrected. The estimation of the residual carrier frequency offset and symbol timing offset are addressed in this work. After initial synchronization, relatively small errors are present in the data transmission phase which need to be compensated for.

2.4.1 Symbol timing offset

Symbol timing errors are introduced in the sampling phase of receiver processing. Incorrect information of the time instant a OFDM symbol starts causes first samples of the symbol to be lost and last samples are populated with samples from the next symbol or vice versa. The samples from adjacent OFDM symbol are disturbance to the current symbol and is called *inter-symbol interference* (ISI). Losing samples causes attenuation of the signal as not all of the transmitted energy is recovered and the orthogonality between sub-carriers is lost causing *inter-carrier interference* (ICI). Cyclic prefix is used to give a buffer to mitigate the effect of too early sampling as received OFDM symbols are only circularly shifted versions of the correct symbol. This applies only in the case when the significant part of multipath channels impulse response is shorter than prefix duration.

The next examples from [10] show the effect of symbol timing offset to the received demodulated samples. We denote the symbol timing offset by $\varepsilon \in \mathbb{R}$.

Example 2.6. Assume an OFDM transmission where there is no cyclic prefix, $N_g = 0$, and perfect frequency synchronisation. Let channel introduce only additive white Gaussian noise. Assume also that the symbol timing offset is not greater than the symbol duration, i.e. $|\varepsilon| < N_s$. Because of the timing offset, the sampling of OFDM symbol l is shifted from the optimal position and done at time indices $t_n = nT + lT_u + \varepsilon T$, $n \in \{0, 1, \ldots, N-1\}$. Then, the received samples of transmitted signal are

$$\begin{aligned}
r_{l,n} =& s(nT + lT_u + \varepsilon T) + w(t_n) \\
=& \frac{1}{\sqrt{N}} \sum_{l=-\infty}^{\infty} \sum_{k'=-\frac{N}{2}}^{\frac{N}{2}-1} c_{l',k'} e^{j2\pi \frac{k'}{T_u}(nT+\varepsilon T+(l-l')T_u)} u(nT + \varepsilon T + (l-l')T_u) \\
& + w(t_n) \\
=& \frac{1}{\sqrt{N}} \sum_{l=-\infty}^{\infty} \sum_{k'=-\frac{N}{2}}^{\frac{N}{2}-1} c_{l',k'} e^{j2\pi \frac{k'n}{N}} e^{j2\pi \frac{\varepsilon k'}{N}} u(nT + \varepsilon T + (l-l')T_u) + w(t_n).
\end{aligned}$$

Demodulation of the samples with discrete Fourier transform gives

$$\begin{aligned}
y_{l,k} =& \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} r_{l,n} e^{-j2\pi kn/N} \\
=& \frac{1}{N} \sum_{n=0}^{N-1} \sum_{l'=-\infty}^{\infty} \sum_{k'=-\frac{N}{2}}^{\frac{N}{2}-1} c_{l',k'} e^{j2\pi \frac{k'n}{N}} e^{j2\pi \frac{\varepsilon k'}{N}} e^{-j2\pi \frac{kn}{N}} u(nT + \varepsilon T + (l-l')T_u)
\end{aligned}$$

$$
\begin{aligned}
&+\frac{1}{\sqrt{N}}\sum_{n=0}^{N-1} w(t_n)e^{-j2\pi kn/N}\\
=&\frac{1}{N}e^{j2\pi\varepsilon k/N}c_{l,k}\sum_{n=0}^{N-1} u(nT+\varepsilon T)\\
&+\frac{1}{N}\sum_{\substack{k'=-\frac{N}{2}\\ k'\neq k}}^{\frac{N}{2}-1} e^{j2\pi\varepsilon k'/N}c_{l,k'}\sum_{n=0}^{N-1} e^{j2\pi(k'-k)n/N}u(nT+\varepsilon T)\\
&+\frac{1}{N}\sum_{\substack{l'=-\infty\\ l'\neq 0}}^{\infty}\sum_{k'=-\frac{N}{2}}^{\frac{N}{2}-1} e^{j2\pi\varepsilon k'/N}c_{l',k'}\sum_{n=0}^{N-1} e^{j2\pi(k'-k)n/N}u(nT+\varepsilon T+(l-l')T_s)\\
&+w_{l,k}\\
=&\frac{N-|\lfloor\varepsilon\rfloor|}{N}e^{j2\pi\varepsilon k/N}c_{l,k}\\
&+\sum_{\substack{k'=-\frac{N}{2}\\ k'\neq k}}^{\frac{N}{2}-1} e^{j\pi(k'-k)\frac{N-\lfloor\varepsilon\rfloor-1}{N}}\frac{\sin\left(\pi(k'-k)\frac{N-|\lfloor\varepsilon\rfloor|}{N}\right)}{N\sin(\pi(k'-k)/N)}e^{j2\pi\varepsilon k'/N}c_{l,k'}\\
&+\sum_{k'=-\frac{N}{2}}^{\frac{N}{2}-1} e^{j\pi(k'-k)\frac{-\lfloor\varepsilon\rfloor-1}{N}}\frac{\sin\left(\pi(k'-k)\frac{|\lfloor\varepsilon\rfloor|}{N}\right)}{N\sin(\pi(k'-k)/N)}e^{j2\pi\varepsilon k'/N}c_{l+\mathrm{sgn}(\varepsilon),k'}\\
&+w_{l,k}\\
=&\frac{N-|\lfloor\varepsilon\rfloor|}{N}e^{j2\pi\varepsilon k/N}c_{l,k}+w_{l,k}^{\mathrm{ICI}}+w_{l,k}^{\mathrm{ISI}}+w_{l,k}.
\end{aligned}
$$

So, the received demodulated block can be written as

$$
\mathbf{y}_l=\frac{N-|\lfloor\varepsilon\rfloor|}{N}\mathbf{\Phi}\mathbf{c}_l+\mathbf{W}_l^{\mathrm{ICI}}+\mathbf{W}_l^{\mathrm{ISI}}+\mathbf{W}_l,
$$

where $\mathbf{\Phi}=\mathrm{diag}\{e^{j2\pi\varepsilon k/N}\}_{k=-\frac{N}{2}}^{\frac{N}{2}-1}$ is the diagonal phase shift matrix and $\mathbf{W}_l^{\mathrm{ICI}}$ and $\mathbf{W}_l^{\mathrm{ISI}}$ are the additional noise terms due to inter-carrier and inter-symbol interference. We see that the timing offset causes attenuation due to loss of signal energy and linear phase rotation. Inter-symbol and inter-carrier interference can be seen as additional noise.

Example 2.7. Consider the OFDM system from previous example. If we assign a cyclic prefix of N_g samples and we assume that the symbol timing offset is $-N_g\leq\varepsilon\leq 0$. Then, there are no samples from adjacent symbols causing inter-symbol interference and due to cyclic prefix the orthogonality

of sub-carriers is preserved. The received signal is therefore only affected by the sub-carrier specific rotation. The demodulated samples are given by

$$\mathbf{y}_l = \mathbf{\Phi}\mathbf{c}_l + \mathbf{W}_l.$$

2.4.2 Frequency offset

The frequency synchronization errors are due to *carrier frequency offset* (CFO) and *sampling frequency offset* (SFO). Carrier frequency offset is caused by the difference between transmitter and receiver oscillator frequencies and channel induced Doppler effect. When the signal is down-mixed with erroneous carrier frequency, the sub-carrier frequencies and the reference frequencies in discrete Fourier transform do not differ by a multiple of $\frac{1}{T_u}$. This means that the orthogonality between sub-carriers is lost in the receiver and inter-carrier interference is introduced. Error in sampling clock frequency causes the sampling time instant to drift gradually too early or late. If the sampling frequency offset is severe or it is not corrected, the discrete Fourier transform input window will drift away from optimal position and samples from other OFDM symbols will result in inter-symbol interference.

The carrier frequency offset in hertz is denoted by Δf. We normalise it with sub-carrier spacing in frequency and denote $\nu = \Delta f T_u$. The sampling clock frequency offset is given by Δf_s. Then the receiver sampling time interval is $T' = \frac{1}{f_s + \Delta f_s} = (1 + \xi)T$, where T is the transmitter sampling interval and $\xi = \frac{T' - T}{T} = -\frac{\Delta f_s}{f_s + \Delta f_s}$.

Example 2.8 (See [10])**.** Here we illustrate the effect of carrier and sampling clock frequency offset to the received samples $\mathbf{y}_l$. We assume perfect symbol timing, $\varepsilon = 0$. Let the transmitted signal experience additive white Gaussian noise channel. The received passband signal is given by

$$r_c(t) = s_c(t) + w_c(t).$$

The down-mixing is done with erroneous carrier frequency $f_c' = f_c + \Delta f$, where Δf is the frequency offset in hertz. Then, the filtered and amplified in-phase and quadrature outputs are given by

$$\begin{aligned} I(t) &= \mathrm{Re}(s(t))\cos(-2\pi\Delta f t) - \mathrm{Im}(s(t))\sin(-2\pi\Delta f t) + w'(t) \quad \text{and} \\ Q(t) &= \mathrm{Re}(s(t))\sin(-2\pi\Delta f t) + \mathrm{Im}(s(t))\cos(-2\pi\Delta f t) + w''(t). \end{aligned}$$

Combining the in-phase and quadrature components gives the received baseband signal as

$$r(t) = I(t) + jQ(t) = e^{-j2\pi\Delta f t} s(t) + w(t).$$

Sampling is done at time indices $t_n = (n + lN_s)(1 + \xi)T + T_g$, $n \in \{0, 1, \ldots, N-1\}$, so that $r((n + lN_s)(1 + \xi)T + T_g) = r_{l,n}$. The demodulated samples are

$$\begin{aligned}
y_{l,k} =& \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} r_{l,n} e^{-j2\pi nk/N} \\
=& \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} e^{-j2\pi \Delta f t_n} s(t_n) e^{-j2\pi nk/N} + \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} w(t_n) e^{-j2\pi nk/N} \\
=& \frac{1}{N} \sum_{l'=-\infty}^{\infty} \sum_{k'=-\frac{N}{2}}^{\frac{N}{2}-1} \sum_{n=0}^{N-1} e^{-j2\pi \Delta f[(n+lN_s)(1+\xi)T+T_g]} e^{j2\pi \frac{k'}{T_u}(n+lN_s)(1+\xi)T - l'T_s)} \\
& \cdot c_{l',k'} u((n + lN_s)(1 + \xi)T + T_g - l'T_s) e^{-j2\pi nk/N} + w_{l,k}.
\end{aligned}$$

Here we concentrate on the exponents. If we denote $\phi_{k',k} = (1+\xi)(-\nu+k') - k$ and $\phi_k = \phi_{k,k}$, we can write the above equation in the following form

$$\begin{aligned}
y_{l,k} =& \frac{1}{N} \sum_{l'=-\infty}^{\infty} \sum_{k'=-\frac{N}{2}}^{\frac{N}{2}-1} \sum_{n=0}^{N-1} e^{j2\pi \phi_{k',k} n/N} e^{-j2\pi \nu N_g/N} e^{j2\pi l \phi_{k'} N_s/N} e^{j2\pi (l-l')k' N_s/N} c_{l',k'} \\
& \cdot u((n + lN_s)(1 + \xi)T + T_g - l'T_s) + w_{l,k}.
\end{aligned}$$

The gradual drift of the sampling window causes inter-symbol interference. We assume that the sampling frequency offset ξ is so small that during one OFDM symbol the sampling window moves at most one sample forwards or backwards. This requires $|\xi N_s| < 1$ or $|\xi| < \frac{1}{N_s}$. If the sampling instances $t_n \in [-T_g + lT_s, (1+l)T_s[$ there are no samples from the previous or following OFDM symbols and inter-symbol interference doesn't occur. We concentrate only on the received OFDM symbols l for which the sampling window is overlapping with corresponding sent symbol, i.e. the timing was correct for OFDM symbol $l = 0$ and we assume sufficiently small l in the following inspections.

1. The sampling window is inside optimal interval $[-T_g + lT_s, (1 + l)T_s[$:

$$\begin{aligned}
y_{l,k} =& \frac{1}{N} \sum_{k'=-\frac{N}{2}}^{\frac{N}{2}-1} e^{-j2\pi \nu N_g/N} e^{j2\pi l \phi_{k'} N_s/N} c_{l,k'} \sum_{n=0}^{N-1} e^{j2\pi \phi_{k',k} n/N} + w_{l,k} \\
=& \sum_{k'=-\frac{N}{2}}^{\frac{N}{2}-1} e^{-j2\pi \nu N_g/N} e^{j2\pi l \phi_{k'} N_s/N} c_{l,k'} e^{j\pi \phi_{k',k} \frac{N-1}{N}} \frac{\sin(\pi \phi_{k',k})}{N \sin(\pi \phi_{k',k}/N)} + w_{l,k} \\
=& \frac{\sin(\pi \phi_k)}{N \sin(\pi \phi_k/N)} e^{-j2\pi \nu N_g/N} e^{j\pi \phi_k \frac{N-1}{N}} e^{j2\pi l \phi_k N_s/N} c_{l,k} + w_{l,k}^{\mathrm{ICI}} + w_{l,k}.
\end{aligned}$$

We see that the received symbols $c_{l,k}$ have been phase rotated and attenuated. Additionally, the orthogonality between sub-carriers has been lost and additional noise term $w_{l,k}^{\mathrm{ICI}}$ due to inter-carrier interference is introduced. For small ϕ_k the attenuation can be approximated by sinc function as $\frac{\sin(\pi\phi_k)}{N\sin(\pi\phi_k/N)} \approx \mathrm{sinc}(\pi\phi_k)$ and is close to unit. The phase shift depends on the OFDM symbol index l and sub-carrier index k.

2. Samples from adjacent OFDM symbol fall into the sampling window. Denote the offset of the sampling window outside the optimal interval by $\bar{n}$. Demodulated samples are given by

$$\begin{aligned}
y_{l,k} =& \frac{1}{N}\sum_{k'=-\frac{N}{2}}^{\frac{N}{2}-1} e^{-j2\pi\nu N_g/N} e^{j2\pi l\phi_{k'}N_s/N} c_{l,k'} \sum_{n=0}^{N-1} e^{j2\pi\phi_{k',k}n/N} \\
& \cdot u((n+lN_s)(1+\xi)T + T_g - lT_s) \\
& + \frac{1}{N}\sum_{k'=-\frac{N}{2}}^{\frac{N}{2}-1} e^{-j2\pi\nu N_g/N} e^{j2\pi l\phi_{k'}N_s/N} e^{\mathrm{sgn}(\bar{n})j2\pi k'N_s/N} c_{l+\mathrm{sgn}(\bar{n}),k'} \\
& \cdot \sum_{n=0}^{N-1} e^{j2\pi\phi_{k',k}n/N} u((n+lN_s)(1+\xi)T + T_g - (l+\mathrm{sgn}(\bar{n}))T_s) \\
& + w_{l,k} \\
=& \sum_{k'=-\frac{N}{2}}^{\frac{N}{2}-1} e^{-j2\pi\nu N_g/N} e^{j2\pi l\phi_{k'}N_s/N} c_{l,k'} e^{j\pi\phi_{k',k}\frac{N-\bar{n}-1}{N}} \frac{\sin\left(\pi\phi_{k',k}\frac{N-|\bar{n}|}{N}\right)}{N\sin(\pi\phi_{k',k}/N)} \\
& + \sum_{k=-\frac{N}{2}}^{\frac{N}{2}-1} e^{-j2\pi\nu N_g/N} e^{j2\pi l\phi_{k'}N_s/N} e^{\mathrm{sgn}(\bar{n})j2\pi k'N_s/N} c_{l+\mathrm{sgn}(\bar{n}),k'} \\
& \cdot e^{j\pi\phi_{k',k}\frac{-\bar{n}-1}{N}} \frac{\sin\left(\pi\phi_{k',k}\frac{|\bar{n}|}{N}\right)}{N\sin(\pi\phi_{k',k}/N)} + w_{l,k} \\
=& \frac{\sin\left(\pi\phi_k\frac{N-|\bar{n}|}{N}\right)}{N\sin(\pi\phi_k/N)} e^{-j2\pi\nu N_g/N} e^{j\pi\phi_k\frac{N-\bar{n}-1}{N}} e^{j2\pi l\phi_k N_s/N} c_{l,k} \\
& + w_{l,k}^{\mathrm{ICI}} + w_{l,k}^{\mathrm{ISI}} + w_{l,k}.
\end{aligned}$$

Compared to the previous case, the attenuation is more severe and additional noise term $w_{l,k}^{\mathrm{ISI}}$ due to inter-symbol interference is introduced.

2.5 Received signal model

Gathering the effects of symbol timing and frequency offsets, we arrive to the signal model that contains the effects of multipath channel, symbol timing

offset, carrier frequency offset and sampling frequency offset [10].

Theorem 2.9. *Let symbol timing offset ε be so small that no inter-symbol interference is introduced. Assume that carrier frequency offset $\nu < \frac{1}{2}$ and sampling clock frequency offset $\xi \ll 1$. Let the cyclic prefix duration be longer than the channel impulse response and assume that the channel impulse response is constant during one OFDM symbol. Then, the demodulated received samples are given by*

$$\mathbf{y}_l = \mathbf{\Phi}(\varepsilon, \nu, \xi)\mathbf{\Theta}_l(\nu, \xi)\mathbf{H}_l\mathbf{c}_l + \mathbf{W}_l^{ICI} + \mathbf{W}_l. \tag{75}$$

The elements of diagonal matrix $\mathbf{\Phi}(\varepsilon, \nu, \xi)$ are given by $\Phi_k = \frac{\sin(\pi\phi_k)}{N\sin(\pi\phi_k/N)}$ $\cdot\, e^{j\pi\left[\phi_k\frac{N-1}{N}+2\phi_{k,0}\frac{\varepsilon}{N}-2\nu\frac{N_g}{N}\right]}$, where $\phi_k = (1+\xi)(-\nu+k)-k$ and $k = -\frac{N}{2}, -\frac{N}{2}+1, \ldots, \frac{N}{2}-1$. This causes attenuation and phase shift depending on the timing ε and frequency offsets ν and ξ. The elements of the diagonal matrix $\mathbf{\Theta}_l(\nu, \xi)$ are given by $\Theta_{l,k} = e^{j2\pi l\phi_k\frac{N_s}{N}}$, $k \in \{-\frac{N}{2}, -\frac{N}{2}+1, \ldots, \frac{N}{2}-1\}$, causing frequency offset ν and ξ and OFDM symbol index l dependent phase shift. The channel matrix $\mathbf{H}_l$ is diagonal and contains the frequency response samples of the channel.

The term $\mathbf{W}_l^{ICI}$ is noise due to inter-carrier interference, given by

$$w_{l,k}^{ICI} = \sum_{k'=-\frac{N}{2};k'\neq k}^{\frac{N}{2}-1} \frac{\sin(\pi\phi_{k',k})}{N\sin(\pi\phi_{k',k}/N)} e^{j\pi\left[\phi_{k',k}\frac{N-1}{N}+2\phi_{k',0}\frac{\varepsilon}{N}-2\nu\frac{N_g}{N}\right]} \cdot e^{j2\pi l\phi_{k'}\frac{N_s}{N}} c_{l,k'}H_{l,k'}. \tag{76}$$

Term $\mathbf{W}_l$ is additive zero mean white Gaussian noise modelled as

$$w_{l,k} = \frac{1}{\sqrt{N}}\sum_{n=0}^{N-1} w((n+\varepsilon+lN_s)(1+\xi)T + T_g)e^{-j2\pi kn/N}, \tag{77}$$

with $\mathbb{E}\{|w_{l,k}|^2\} = \mathbb{E}\{|w(t)|^2\} = \sigma_w^2$.

Proof. Let $s(t)$ denote the continuous modulated signal that is transmitted. The signal is modelled as

$$s(t) = \frac{1}{\sqrt{N}}\sum_{l=-\infty}^{\infty}\sum_{k=-\frac{N}{2}}^{\frac{N}{2}-1} c_{l,k}e^{j2\pi\frac{k}{T_u}(t-T_g-lT_s)}u(t-lT_s),$$

where $u(t)$ is a windowing function given by

$$u(t) = \begin{cases} 1, & 0 < t < T_s, \\ 0, & \text{else.} \end{cases}$$

After passing through the channel, the signal is affected by channel impulse response $h(\tau,t)=\sum_{i=1}^{L}h_i(t)\delta(\tau_i-\tau)$ and additional noise $w(t)$ is introduced. Mismatch between transmitter and receiver carrier frequencies and channel induced Doppler effect cause carrier frequency offset Δf, which is seen as a phase rotation of the signal. The received signal is given as

$$r(t)=e^{-j2\pi\Delta ft}\sum_{i=1}^{L}h_i(t)s(t-\tau_i)+w(t).$$

The received signal is sampled at time indices $t_n=(n+\varepsilon+lN_s)(1+\xi)T+T_g$, $n\in\{0,1,\ldots,N-1\}$. The received samples are given by

$$\begin{aligned}
r_{l,n}=&r(t_n)=e^{-j2\pi\Delta ft_n}\sum_{i=1}^{L}h_i(t_n)s(t_n-\tau_i)+w(t_n)\\
=&e^{-j2\pi\Delta ft_n}\sum_{i=1}^{L}h_i(l)\frac{1}{\sqrt{N}}\sum_{l'=-\infty}^{\infty}\sum_{k'=-\frac{N}{2}}^{\frac{N}{2}-1}c_{l',k'}e^{j2\pi\frac{k'}{T_u}(t_n-\tau_i-T_g-l'T_s)}\\
&\cdot u(t_n-\tau_i-l'T_s)+w(t_n)\\
=&\frac{1}{\sqrt{N}}\sum_{l'=-\infty}^{\infty}\sum_{k'=-\frac{N}{2}}^{\frac{N}{2}-1}c_{l',k'}e^{j2\pi\left[-\Delta ft_n+\frac{k'}{T_u}(t_n-T_g-l'T_s)\right]}\\
&\cdot\sum_{i=1}^{L}h_i(l)e^{-j2\pi\frac{\tau_i}{T}k'/N}u(t_n-\tau_i-l'T_s)+w(t_n).
\end{aligned}$$

The term in the exponent can be written in equivalent form as

$$\begin{aligned}
&-\Delta ft_n+\frac{k'}{N}\left(\frac{t_n}{T}-N_g-l'N_s\right)\\
=&\frac{1}{N}[-\nu(n+\varepsilon+lN_s)(1+\xi)-\nu N_g+k'((n+\varepsilon+lN_s)(1+\xi)-l'N_s)]\\
=&\frac{1}{N}[(1+\xi)(-\nu+k')n+l(1+\xi)(-\nu+k')N_s-l'k'N_s\\
&+(1+\xi)(-\nu+k')\varepsilon-\nu N_g]\\
=&\phi_{k',0}\frac{n}{N}+l\phi_{k'}\frac{N_s}{N}+k'(l-l')\frac{N_s}{N}+\phi_{k',0}\frac{\varepsilon}{N}-\nu\frac{N_g}{N},
\end{aligned}$$

where we denote $\nu=\Delta fNT$, $\phi_{k',k}=(1+\xi)(-\nu+k')-k$ and $\phi_k=\phi_{k,k}$.

The received samples are demodulated by passing through a discrete Fourier transform and we get

$$y_{l,k}=\frac{1}{\sqrt{N}}\sum_{n=0}^{N-1}r_{l,n}e^{-j2\pi kn/N}$$

$$
\begin{aligned}
=&\frac{1}{N}\sum_{l'=-\infty}^{\infty}\sum_{k'=-\frac{N}{2}}^{\frac{N}{2}-1}\sum_{n=0}^{N-1} c_{l',k'}e^{j2\pi[\phi_{k',0}n/N+l\phi_{k'}N_s/N+k'(l-l')N_s/N+\phi_{k',0}\varepsilon/N-\nu N_g/N]}\\
&\cdot\sum_{i=1}^{L} h_i(l)e^{-2\pi\frac{\tau_i}{T}k'/N}u(t_n-\tau_i-l'T_s)e^{-j2\pi kn/N}\\
&+\frac{1}{\sqrt{N}}\sum_{n=0}^{N-1} w(t_n)e^{-j2\pi kn/N}\\
=&\frac{1}{N}\sum_{l'=-\infty}^{\infty}\sum_{k'=-\frac{N}{2}}^{\frac{N}{2}-1}\sum_{n=0}^{N-1} c_{l',k'}e^{j2\pi[\phi_{k',k}n/N+l\phi_{k'}N_s/N+k'(l-l')N_s/N+\phi_{k',0}\varepsilon/N-\nu N_g/N]}\\
&\cdot\sum_{i=1}^{L} h_i(l)e^{-j2\pi\frac{\tau_i}{T}k'/N}u((n+\varepsilon)(1+\xi)T+T_g+(l-l')T_s+l\xi T_s-\tau_i)\\
&+w_{l,k}\\
=&\frac{1}{N}\sum_{l'=-\infty}^{\infty}\sum_{k'=-\frac{N}{2}}^{\frac{N}{2}-1} c_{l',k'}e^{j2\pi[l\phi_{k'}S_s/N+k'(l-l')N_s/N+\phi_{k',0}\varepsilon/N-\nu N_g/N]}\\
&\cdot\sum_{n=0}^{N-1} e^{j2\pi\phi_{k',k}n/N}\sum_{i=1}^{L} h_i(l)e^{-j2\pi\frac{\tau_i}{T}k'/N}\\
&\cdot u((n+\varepsilon)(1+\xi)T+T_g+(l-l')T_s+l\xi T_s-\tau_i)+w_{l,k}.
\end{aligned}
$$

With the assumption that symbol timing and sampling clock frequency offsets are adequately small and that the cyclic prefix is longer than the channel impulse response, the windowing function $u(t)$ will evaluate unit only when $l'=l$. The equation simplifies to

$$
\begin{aligned}
y_{l,k}=&\frac{1}{N}\sum_{k'=-\frac{N}{2}}^{\frac{N}{2}-1} c_{l,k'}e^{j2\pi[l\phi_{k'}N_s/N+\phi_{k',0}\varepsilon/N-\nu N_g/N]}\\
&\cdot\sum_{n=0}^{N-1} e^{j2\pi\phi_{k',k}n/N}\sum_{i=1}^{L} h_i(l)e^{-j2\pi\frac{\tau_i}{T}k'/N}u((n+\varepsilon)(1+\xi)T+T_g+l\xi T_s-\tau_i)\\
&+w_{l,k}.
\end{aligned}
$$

Because the cyclic prefix is an exact copy of the last part of the OFDM symbol the delayed samples of the cyclic prefix compensate for the loss of delayed samples of the last part. Then

$$
y_{l,k}=\frac{1}{N}\sum_{k'=-\frac{N}{2}}^{\frac{N}{2}-1} c_{l,k'}e^{j2\pi[l\phi_{k'}N_s/N+\phi_{k',0}\varepsilon/N-\nu N_g/N]}
$$

$$
\begin{aligned}
&\cdot \sum_{n=0}^{N-1} e^{j2\pi\phi_{k',k}n/N} \sum_{i=1}^{L} h_i(l) e^{-j2\pi\frac{\tau_i}{T}k'/N} + w_{l,k} \\
=&\frac{1}{N} \sum_{k'=-\frac{N}{2}}^{\frac{N}{2}-1} c_{l,k'} e^{j2\pi[l\phi_{k'}N_s/N+\phi_{k',0}\varepsilon/N-\nu N_g/N]} e^{j\pi\phi_{k',k}\frac{N-1}{N}} \frac{\sin(\pi\phi_{k',k})}{\sin(\pi\phi_{k',k}/N)} H_{l,k'} \\
&+ w_{l,k} \\
=&\sum_{k'=-\frac{N}{2}}^{\frac{N}{2}-1} \frac{\sin(\pi\phi_{k',k})}{N\sin(\pi\phi_{k',k}/N)} e^{j2\pi\left[\phi_{k',0}\varepsilon/N-\nu N_g/N+\phi_{k',k}\frac{N-1}{2N}\right]} e^{j2\pi l\phi_{k'}N_s/N} H_{l,k'}c_{l,k'} \\
&+ w_{l,k} \\
=&\frac{\sin(\pi\phi_k)}{N\sin(\pi\phi_k/N)} e^{j\pi\left[\phi_k\frac{N-1}{N}+2\phi_{k,0}\varepsilon/N-2\nu N_g/N\right]} e^{j2\pi l\phi_k N_s/N} H_{l,k}c_{l,k} \\
&+ w_{l,k}^{\mathrm{ICI}} + w_{l,k}.
\end{aligned}
$$

□

In this work, only a small residual carrier frequency offset is assumed. The sampling clock frequency and carrier frequency are also assumed to be generated from same source. This implies that the sampling clock frequency offset is negligible and can be assumed zero.

Example 2.10. Consider an OFDM system with carrier frequency $f_c = 2.4$ GHz and sub-carrier spacing $f_d = 15$ kHz, which corresponds to a 3GPP LTE system [1]. If we have a normalized carrier frequency offset of ν, then $\Delta f = \nu f_d$ and receiver has a carrier frequency of $f_c + \nu f_d$. This means that the oscillator has a normalized error of $\frac{\nu f_d}{f_c}$ which is transferred to the sampling frequency such that $\Delta f_s = \frac{\nu f_d}{f_c} f_s$. This gives the relation between ν and ξ as $\xi = \frac{-\Delta f_s}{f_s+\Delta f_s} = -\frac{\nu f_d}{f_c+\nu f_d} \approx -\frac{f_d}{f_c}\nu$. This means that the effect of carrier frequency offset ν is $\frac{f_c}{f_d} = 160000$ times larger in magnitude than the effect of sampling clock frequency offset ξ. With carrier frequency offset in range $|\nu| < 0.1$ the sampling clock frequency offset $|\xi| < 10^{-6}$. Therefore, the sampling clock frequency offset can be considered insignificant

This simplification can be applied to the previous Theorem 2.9.

Theorem 2.11. *When the sampling frequency offset is assumed to be insignificant, we set $\xi = 0$ and received samples become*

$$
\mathbf{y}_l = \alpha(\varepsilon,\nu)\theta_l(\nu)\mathbf{\Phi}(\varepsilon)\mathbf{H}_l\mathbf{c}_l + \mathbf{W}_l^{ICI} + \mathbf{W}_l. \tag{78}
$$

The complex value $\alpha(\varepsilon,\nu) = \frac{\sin(\pi\nu)}{N\sin(\pi\nu/N)} e^{-j\pi\nu\frac{N+2N_g+2\varepsilon-1}{N}}$ *doesn't depend on the sub-carrier or the OFDM symbol indices* k *or* l*. The complex value* $\theta_l(\nu) = e^{-j2\pi l\nu\frac{N_s}{N}}$ *depends on the carrier frequency offset* ν *and OFDM symbol index* l*. The elements of diagonal matrix* $\mathbf{\Phi}(\varepsilon)$ *are given by* $\Phi_k = e^{j2\pi\frac{\varepsilon k'}{N}}$*,* $k = -\frac{N}{2}, -\frac{N}{2}+1, \ldots, \frac{N}{2}-1$*.*

The additional noise $\mathbf{W}_l^{ICI}$ *due to inter-carrier interference is given by*

$$w_{l,k}^{ICI} = \sum_{k'=-\frac{N}{2};k'\neq k}^{\frac{N}{2}-1} \frac{\sin(\pi(-\nu+k'-k)}{N\sin(\pi(-\nu+k'-k)/N)} e^{-j\pi\nu\frac{N+2N_g+2\varepsilon-1}{N}} e^{j\pi(k'-k)\frac{N-1}{N}} \cdot e^{j2\pi\frac{\varepsilon k'}{N}} e^{-j2\pi l\nu\frac{N_s}{N}} H_{l,k'} c_{l,k'}. \tag{79}$$

Proof. The result follows immediately from Theorem 2.9 by setting $\xi = 0$. □

The next result gives an approximation on the severity of increased noise due to inter-carrier interference caused by carrier frequency offset. A similar review is done in [13].

Theorem 2.12. *Assume that sent symbols* $c_{l,k}$ *are independent zero mean random samples with* $\mathbb{E}\{|c_{l,k}|^2\} = P_s$*. Assume also that channel frequency response samples* $H_{l,k}$ *are strongly correlated zero mean random samples such that* $H_{l,k}H_{l+m,k}^* = |H_{l,k}|^2$ *and* $\mathbb{E}\{|H_{l,k}|^2\} = P_H$*. Then, the inter-carrier interference term* $\mathbf{W}_l^{ICI}$ *in Theorem 2.11 can be seen as a vector of zero mean random variables* $w_{l,k}^{ICI}$ *with variance* $\mathbb{E}\{|w_{l,k}^{ICI}|^2\} \approx \frac{\pi^2}{3}P_sP_H\nu^2$*. Also, due to central limit theorem, the distribution of* $w_{l,k}^{ICI}$ *approaches to the zero mean Gaussian distribution with variance* $\sigma^2 = \frac{\pi^2}{3}P_sP_H\nu^2$ *as the number of used sub-carriers increases.*

Proof. The inter-carrier interference term $w_{l,k}^{\mathrm{ICI}}$ is given by Theorem 2.11 as

$$w_{l,k}^{\mathrm{ICI}} = \sum_{k'=-\frac{N}{2};k'\neq k}^{\frac{N}{2}-1} \frac{\sin(\pi(-\nu+k'-k)}{N\sin(\pi(-\nu+k'-k)/N)} e^{-j\pi\nu\frac{N+2N_g+2\varepsilon-1}{N}} e^{j\pi(k'-k)\frac{N-1}{N}} \cdot e^{j2\pi\frac{\varepsilon k'}{N}} e^{-j2\pi l\nu\frac{N_s}{N}} H_{l,k'} c_{l,k'}.$$

The mean $\mathbb{E}\{w_{l,k}^{\mathrm{ICI}}\} = 0$ because $H_{l,k'}$ and $c_{l,k'}$ are independent zero mean random variables. The variance

$$\begin{aligned}\sigma_{\mathrm{ICI}}^2 =& \mathbb{E}\{|w_{l,k}^{\mathrm{ICI}}|^2\} = \mathbb{E}\{w_{l,k}^{\mathrm{ICI}}(w_{l,k}^{\mathrm{ICI}})^*\} \\ =& \sum_{k'=-\frac{N}{2};k'\neq k}^{\frac{N}{2}-1} \frac{\sin^2(\pi(-\nu+k'-k))}{N^2\sin^2(\pi(-\nu+k'-k)/N)} \mathbb{E}\{|c_{l,k'}|^2\}\mathbb{E}\{|H_{l,k'}|^2\}\end{aligned}$$

$$= P_s P_H \sum_{k'=-\frac{N}{2};k'\neq k}^{\frac{N}{2}-1} \frac{\sin^2(\pi(-\nu+k'-k))}{N^2\sin^2(\pi(-\nu+k'-k)/N)},$$

where the cross product terms evaluate to zero because $c_{l,k'}$ are independent. Each term in summation is approximated with first order Taylor series with respect to $k'-k$ to get

$$\frac{\sin^2(\pi(-\nu+k'-k))}{N^2\sin^2(\pi(-\nu+k'-k)/N)} \approx \frac{\pi^2\nu^2}{N^2\sin^2(\pi(k'-k)/N)}.$$

Because the $\sin^2(x)$ function is π-periodic, the sum can be approximated as

$$\begin{aligned}\sum_{k'=-\frac{N}{2};k'\neq k}^{\frac{N}{2}-1} \frac{\sin^2(\pi(-\nu+k'-k))}{N^2\sin^2(\pi(-\nu+k'-k)/N)} &\approx 2\sum_{k'=1}^{\frac{N}{2}} \frac{\pi^2\nu^2}{N^2\sin^2(\pi k'/N)} \\ &\approx 2\nu^2 \sum_{k'=1}^{\frac{N}{2}} \frac{1}{k'^2},\end{aligned}$$

where we have approximated $\sin(x) \approx x$ with first order Taylor series. The sum approaches to the series $\sum_{k'=1}^{\infty} \frac{1}{k'^2}$, which converges, and we get

$$2\nu^2 \sum_{k'=1}^{\frac{N}{2}} \frac{1}{k'^2} < 2\nu^2 \sum_{k'=1}^{\infty} \frac{1}{k'^2} = \frac{\pi^2}{3}\nu^2.$$

Then it follows that the variance is approximated by

$$\mathbb{E}\{|w_{l,k}^{\mathrm{ICI}}|^2\} \approx \frac{\pi^2}{3} P_s P_H \nu^2.$$

This approximation holds for small ν. □

3 Timing and frequency offset estimation

The receiver needs to be synchronized with the transmitter in order to be able to detect the transmitted data symbols. Synchronization begins with an acquisition phase where the receiver does not have any prior information of the timing or frequency offsets. Special synchronization signals can be multiplexed in the transmitted signal stream that are used to make coarse estimates [1].

After acquisition phase, the accuracy of the symbol timing and carrier frequency offsets estimates need to be improved. Continuous tracking is needed

because the offsets are time-varying. Temperature changes affect the resonance frequency of the oscillator and that causes the symbol timing and frequency to gradually drift away from optimal values. In low signal to noise ratios, long averaging over time is needed to gather enough energy for good estimation of the channel and received symbols. During these long averaging periods even small offsets in oscillator frequency cause degradation to the performance of the receiver.

From our signal model in Theorem 2.11, we know that received signal samples are given by

$$y_{l,k} = \frac{\sin(\pi\nu)}{N\sin(\pi\nu/N)} e^{-j\pi\nu\frac{N+2N_g+2\varepsilon-1}{N}} e^{-j2\pi\nu\frac{N_s}{N}} e^{j2\pi\varepsilon\frac{k}{N}} H_{l,k}c_{l,k} + w_{l,k}^{\mathrm{ICI}} + w_{l,k}, \tag{80}$$

when symbol timing and carrier frequency offsets are small. The phase rotation $e^{-j\pi\nu\frac{N+2N_g+2\varepsilon-1}{N}}$ can be regarded as a constant during estimation period if the symbol timing and carrier frequency offsets ε and ν are slowly time-varying. Therefore, the phase rotation effect can be incorporated to the unknown channel frequency response. We also make the approximation $\frac{\sin(\pi\nu)}{N\sin(\pi\nu/N)} \approx 1$ and neglect the effect of inter-carrier interference. The channel is assumed to be time-invariant and also frequency-invariant during the estimation period. Then, the signal model used in the analytic derivations is given as

$$y_{l,k} = e^{j2\pi\varepsilon\frac{k}{N}} e^{-j2\pi l\nu\frac{N_s}{N}} H c_{l,k} + w_{l,k}, \tag{81}$$

where ε, ν and H are unknown constants and the noise $w_{l,k}$ is white Gaussian with zero mean and variance σ_w^2. The symbols $c_{l,k} = x_{l,k}$ are known in the positions where pilots are transmitted and we can set $|x_{l,k}|^2 = 1$. Unknown data symbols $c_{l,k} = d_{l,k}$, $(l,k) \in \mathcal{D}$, are assumed to be zero mean Gaussian random variables with variance $\mathbb{E}\{|d_{l,k}|^2\} = \sigma_d^2$.

3.1 Cramer-Rao lower bound

We derive the Cramer-Rao lower bound for the estimators of symbol timing $\hat{\varepsilon}$ and carrier frequency $\hat{\nu}$ offsets, where the estimation is done over the pilot symbols, according to Theorem 1.6. Denote the vector of incoming samples as $\mathbf{y}$ that contains N_p demodulation samples from pilot index set $\mathcal{P}$. Then $c_{l,k} = x_{l,k}$ for all $(l,k) \in \mathcal{P}$. The channel exhibits Rayleigh fading, so the real H_r and imaginary H_i parts of frequency response H are independent and can be treated as separate parameters. Denote the unknown parameter

vector by $\mathbf{\Theta} = [\varepsilon, \nu, H_r, H_i]^T$. Then, the probability density function of the received pilot samples is given by

$$f(\mathbf{y}; \mathbf{\Theta}) = \left(\frac{1}{\pi\sigma_w^2}\right)^{N_p} \exp\left(-\frac{1}{\sigma_w^2} \sum_{(l,k)\in\mathcal{P}} \left| y_{l,k} - e^{j2\pi\varepsilon\frac{k}{N}} e^{-j2\pi l\nu\frac{N_s}{N}} H x_{l,k} \right|^2\right). \quad (82)$$

The logarithm of probability density function is then given by

$$\begin{aligned}\ln f(\mathbf{y}; \mathbf{\Theta}) =& - N_p \ln(\pi\sigma_w^2) - \frac{1}{\sigma_w^2} \sum_{(l,k)\in\mathcal{P}} \left| y_{l,k} - e^{j2\pi\varepsilon\frac{k}{N}} e^{-j2\pi l\nu\frac{N_s}{N}} H x_{l,k} \right|^2 \\ =& - N_p \ln(\pi\sigma_w^2) - \frac{N_p|H|^2}{\sigma_w^2} \\ & - \frac{1}{\sigma_w^2} \sum_{(l,k)\in\mathcal{P}} \left[|y_{l,k}|^2 - 2\,\mathrm{Re}\left(y_{l,k} x_{l,k}^* e^{-j2\pi\varepsilon\frac{k}{N}} e^{-j2\pi l\nu\frac{N_s}{N}} H^* \right) \right]. \quad (83)\end{aligned}$$

Then the first partial derivatives are given by

$$\frac{\partial}{\partial\varepsilon} \ln f(\mathbf{y}; \mathbf{\Theta}) = \frac{4\pi}{N\sigma_w^2} \sum_{(l,k)\in\mathcal{P}} k\,\mathrm{Im}\left(y_{l,k} x_{l,k}^* e^{-j2\pi\varepsilon\frac{k}{N}} e^{j2\pi l\nu\frac{N_s}{N}} H^* \right), \quad (84)$$

$$\frac{\partial}{\partial\nu} \ln f(\mathbf{y}; \mathbf{\Theta}) = - \frac{4\pi N_s}{N\sigma_w^2} \sum_{(l,k)\in\mathcal{P}} l\,\mathrm{Im}\left(y_{l,k} x_{l,k}^* e^{-j2\pi\varepsilon\frac{k}{N}} e^{j2\pi l\nu\frac{N_s}{N}} H^* \right), \quad (85)$$

$$\frac{\partial}{\partial H_r} \ln f(\mathbf{x}; \mathbf{\Theta}) = - \frac{2N_p H_r}{\sigma_w^2} + \frac{2}{\sigma_w^2} \sum_{(l,k)\in\mathcal{P}} \mathrm{Re}\left(y_{l,k} x_{l,k}^* e^{-j2\pi\varepsilon\frac{k}{N}} e^{j2\pi l\nu\frac{N_s}{N}} \right) \quad \text{and} \quad (86)$$

$$\frac{\partial}{\partial H_i} \ln f(\mathbf{y}; \mathbf{\Theta}) = - \frac{2N_p H_i}{\sigma_w^2} + \frac{2}{\sigma_w^2} \sum_{(l,k)\in\mathcal{P}} \mathrm{Im}\left(y_{l,k} x_{l,k}^* e^{-j2\pi\varepsilon\frac{k}{N}} e^{j2\pi l\nu\frac{N_s}{N}} \right). \quad (87)$$

The second partial derivatives are given by

$$\frac{\partial^2}{\partial\varepsilon^2} \ln f(\mathbf{y}; \mathbf{\Theta}) = - \frac{8\pi^2}{N^2\sigma_w^2} \sum_{(l,k)\in\mathcal{P}} k^2\,\mathrm{Re}\left(y_{l,k} x_{l,k}^* e^{-j2\pi\varepsilon\frac{k}{N}} e^{j2\pi l\nu\frac{N_s}{N}} H^* \right), \quad (88)$$

$$\frac{\partial^2}{\partial\nu^2} \ln f(\mathbf{y}; \mathbf{\Theta}) = - \frac{8\pi^2 N_s^2}{N^2\sigma_w^2} \sum_{(l,k)\in\mathcal{P}} l^2\,\mathrm{Re}\left(y_{l,k} x_{l,k}^* e^{-j2\pi\varepsilon\frac{k}{N}} e^{j2\pi l\nu\frac{N_s}{N}} H^* \right), \quad (89)$$

$$\frac{\partial^2}{\partial H_r^2} \ln f(\mathbf{y}; \mathbf{\Theta}) = \frac{\partial^2}{\partial H_i^2} \ln f(\mathbf{y}; \mathbf{\Theta}) = -\frac{2N_p}{\sigma_w^2}, \quad (90)$$

$$\frac{\partial^2}{\partial\nu\partial\varepsilon} \ln f(\mathbf{y}; \mathbf{\Theta}) = \frac{8\pi^2 N_s}{N^2\sigma_w^2} \sum_{(l,k)\in\mathcal{P}} lk\,\mathrm{Re}\left(y_{l,k} x_{l,k}^* e^{-j2\pi\varepsilon\frac{k}{N}} e^{j2\pi l\nu\frac{N_s}{N}} H^* \right), \quad (91)$$

$$\frac{\partial^2}{\partial\varepsilon\partial H_r}\ln f(\mathbf{y};\mathbf{\Theta}) = \frac{4\pi}{N\sigma_w^2}\sum_{(l,k)\in\mathcal{P}} k\,\mathrm{Im}\left(y_{l,k}x_{l,k}^* e^{-j2\pi\varepsilon\frac{k}{N}}e^{j2\pi l\nu\frac{N_s}{N}}\right), \tag{92}$$

$$\frac{\partial^2}{\partial\varepsilon\partial H_i}\ln f(\mathbf{y}:\mathbf{\Theta}) = -\frac{4\pi}{N\sigma_w^2}\sum_{(l,k)\in\mathcal{P}} k\,\mathrm{Re}\left(y_{l,k}x_{l,k}^* e^{-j2\pi\varepsilon\frac{k}{N}}e^{j2\pi l\nu\frac{N_s}{N}}\right), \tag{93}$$

$$\frac{\partial^2}{\partial\nu\partial H_r}\ln f(\mathbf{y};\mathbf{\Theta}) = -\frac{4\pi N_s}{N\sigma_w^2}\sum_{(l,k)\in\mathcal{P}} l\,\mathrm{Im}\left(y_{l,k}x_{l,k}^* e^{-j2\pi\varepsilon\frac{k}{N}}e^{j2\pi l\nu\frac{N_s}{N}}\right), \tag{94}$$

$$\frac{\partial^2}{\partial\nu\partial H_i}\ln f(\mathbf{y};\mathbf{\Theta}) = \frac{4\pi N_s}{N\sigma_w^2}\sum_{(l,k)\in\mathcal{P}} l\,\mathrm{Re}\left(y_{l,k}x_{l,k}^* e^{-2\pi\varepsilon\frac{k}{N}}e^{j2\pi l\nu\frac{N_s}{N}}\right) \quad \text{and} \tag{95}$$

$$\frac{\partial^2}{\partial H_r\partial H_i}\ln f(\mathbf{y};\mathbf{\Theta}) = 0. \tag{96}$$

Then the Fisher information matrix can be formed as

$$\mathbf{F}(\mathbf{\Theta}) = \begin{bmatrix}\mathbf{F}_{11} & \mathbf{F}_{12}\\ \mathbf{F}_{12}^T & \mathbf{F}_{22}\end{bmatrix}, \tag{97}$$

where

$$\mathbf{F}_{11} = \frac{8\pi^2|H|^2}{N^2\sigma_w^2}\begin{bmatrix}\sum\limits_{(l,k)\in\mathcal{P}} k^2 & -N_s\sum\limits_{(l,k)\in\mathcal{P}} lk\\ -N_s\sum\limits_{(l,k)\in\mathcal{P}} lk & N_s^2\sum\limits_{(l,k)\in\mathcal{P}} l^2\end{bmatrix}, \tag{98}$$

$$\mathbf{F}_{12} = \frac{4\pi}{N\sigma_w^2}\begin{bmatrix}-H_i\sum\limits_{(l,k)\in\mathcal{P}} k & H_r\sum\limits_{(l,k)\in\mathcal{P}} k\\ N_sH_i\sum\limits_{(l,k)\in\mathcal{P}} l & -N_sH_r\sum\limits_{(l,k)\in\mathcal{P}} l\end{bmatrix} \tag{99}$$

and

$$\mathbf{F}_{22} = \frac{2N_p}{\sigma_w^2}\mathbf{I}_{2\times 2}. \tag{100}$$

Inverting $\mathbf{F}(\mathbf{\Theta})$ using block-wise inversion gives

$$\mathbf{F}^{-1}(\mathbf{\Theta}) = \begin{bmatrix}(\mathbf{F}_{11}-\mathbf{F}_{12}\mathbf{F}_{22}^{-1}\mathbf{F}_{12}^T)^{-1} & -(\mathbf{F}_{11}-\mathbf{F}_{12}\mathbf{F}_{22}^{-1}\mathbf{F}_{12}^T)^{-1}\mathbf{F}_{12}\mathbf{F}_{22}^{-1}\\ -\mathbf{F}_{22}^{-1}\mathbf{F}_{12}^T(\mathbf{F}_{11}-\mathbf{F}_{12}\mathbf{F}_{22}^{-1}\mathbf{F}_{12}^H)^{-1} & (\mathbf{F}_{22}-\mathbf{F}_{12}^T\mathbf{F}_{11}^{-1}\mathbf{F}_{12})^{-1}\end{bmatrix}. \tag{101}$$

We are interested in the upper left part which gives

$$(\mathbf{F}_{11}-\mathbf{F}_{12}\mathbf{F}_{22}^{-1}\mathbf{F}_{12}^T)^{-1} = \left(\frac{8\pi^2|H|^2}{N^2\sigma_w^2}\begin{bmatrix}J(1) & J(2)\\ J(2) & J(3)\end{bmatrix}\right)^{-1}$$

$$= \frac{N^2\sigma_w^2}{8\pi^2|H|^2\left(J(1)J(3)-J(2)^2\right)} \begin{bmatrix} J(3) & -J(2) \\ -J(2) & J(1) \end{bmatrix}, \tag{102}$$

where

$$J(1) = \sum_{(l,k)\in\mathcal{P}} k^2 - \frac{1}{N_p}\left(\sum_{(l,k)\in\mathcal{P}} k\right)^2, \tag{103}$$

$$J(2) = -N_s\left(\sum_{(l,k)\in\mathcal{P}} lk - \frac{1}{N_p}\left(\sum_{(l,k)\in\mathcal{P}} l\right)\left(\sum_{(l,k)\in\mathcal{P}} k\right)\right) \quad \text{and} \tag{104}$$

$$J(3) = N_s^2\left(\sum_{(l,k)\in\mathcal{P}} l^2 - \frac{1}{N_p}\left(\sum_{(l,k)\in\mathcal{P}} l\right)^2\right). \tag{105}$$

The lower bound for the variance of unbiased estimators for symbol timing and carrier frequency offsets ε and ν is then given by

$$\mathbb{E}\{|\varepsilon-\hat{\varepsilon}|^2\} \geq \frac{N^2\sigma_w^2}{8\pi^2|H|^2}\frac{J(3)}{J(1)J(3)-J(2)^2} \quad \text{and} \tag{106}$$

$$\mathbb{E}\{|\nu-\hat{\nu}|^2\} \geq \frac{N^2\sigma_w^2}{8\pi^2|H|^2}\frac{J(1)}{J(1)J(3)-J(2)^2}. \tag{107}$$

Result for the carrier frequency offset is the same as derived in [14], where they did not consider the estimation of timing offset.

3.2 Maximum likelihood estimate

The maximum likelihood estimate is given by the global maximum of the log-likelihood function. The estimation is done over known reference symbols found in positions $y_{l,k}$, where l and k pairs are from the set $\mathcal{P}$. We assume that the parameters to be estimated are unknown constants and that the noise $w_{l,k}$ has a Gaussian distribution. The same maximum likelihood estimate is also derived in [15] where only frequency offsets were considered.

The log-likelihood function of the unknown parameters is given by

$$\begin{aligned}\Lambda(\tilde{\varepsilon},\tilde{\nu},\tilde{H}) &= \ln(p(\mathbf{y}_l|\tilde{\varepsilon},\tilde{\nu},\tilde{H})) \\ &= -N_p\ln(\pi\sigma_w^2) - \frac{1}{\sigma_w^2}\sum_{(l,k)\in\mathcal{P}}\left|y_{l,k} - e^{j2\pi\tilde{\varepsilon}\frac{k}{N}}e^{-j2\pi l\tilde{\nu}\frac{N_s}{N}}\tilde{H}x_{l,k}\right|^2. \end{aligned} \tag{108}$$

To estimate only symbol timing and carrier frequency offsets, we need to first maximize the log-likelihood function with respect to the unknown channel.

This is achieved by finding the root of the partial derivative with respect to $\tilde{H}$. We take the derivative

$$\begin{aligned}\frac{\partial}{\partial \tilde{H}}\Lambda(\tilde{\varepsilon},\tilde{\nu},\tilde{H}) =& -\frac{1}{\sigma_w^2}\sum_{(l,k)\in\mathcal{P}}\frac{\partial}{\partial \tilde{H}}\left|y_{l,k}-e^{j2\pi\tilde{\varepsilon}\frac{k}{N}}e^{-j2\pi l\tilde{\nu}\frac{N_s}{N}}\tilde{H}x_{l,k}\right|^2\\ =& -\frac{1}{\sigma_w^2}\sum_{(l,k)\in\mathcal{P}}\frac{\partial}{\partial \tilde{H}}\Big(|y_{l,k}|^2-y_{l,k}e^{-j2\pi\tilde{\varepsilon}\frac{k}{N}}e^{j2\pi l\tilde{\nu}\frac{N_s}{N}}\tilde{H}^*x_{l,k}^*\\ & -y_{l,k}^*e^{j2\pi\tilde{\varepsilon}\frac{k}{N}}e^{-j2\pi l\tilde{\nu}\frac{N_s}{N}}\tilde{H}x_{l,k}+\tilde{H}\tilde{H}^*\Big)\\ =& -\frac{1}{\sigma_w^2}\sum_{(l,k)\in\mathcal{P}}\Big(-y_{l,k}^*e^{j2\pi\tilde{\varepsilon}\frac{k}{N}}e^{-j2\pi l\tilde{\nu}\frac{N_s}{N}}x_{l,k}+\tilde{H}^*\Big)\\ =& -\frac{N_p}{\sigma_w^2}\tilde{H}^*+\frac{1}{\sigma_w^2}\sum_{(l,k)\in\mathcal{P}}y_{l,k}^*x_{l,k}e^{j2\pi\tilde{\varepsilon}\frac{k}{N}}e^{-j2\pi l\tilde{\nu}\frac{N_s}{N}}. \end{aligned}\tag{109}$$

Setting the derivative to zero and solving for $\tilde{H}$ gives the maximum likelihood estimate of the channel with dependence on the parameters $\tilde{\varepsilon}$ and $\tilde{\nu}$. The estimate is given by

$$\hat{H}(\tilde{\varepsilon},\tilde{\nu})=\frac{1}{N_p}\sum_{(l,k)\in\mathcal{P}}y_{l,k}x_{l,k}^*e^{-j2\pi\tilde{\varepsilon}\frac{k}{N}}e^{j2\pi l\tilde{\nu}\frac{N_s}{N}}.\tag{110}$$

Placing the result in the log-likelihood function gives the log-likelihood function only for the timing and frequency offset parameters as

$$\begin{aligned}\Lambda(\tilde{\varepsilon},\tilde{\nu}) =&\Lambda(\tilde{\varepsilon},\tilde{\nu},\hat{H}(\tilde{\varepsilon},\tilde{\nu}))\\ =& -N_p\ln(\pi\sigma_w^2)-\frac{1}{\sigma_w^2}\sum_{(l,k)\in\mathcal{P}}\left|y_{l,k}-e^{j2\pi\tilde{\varepsilon}\frac{k}{N}}e^{-j2\pi l\tilde{\nu}\frac{N_s}{N}}\hat{H}(\tilde{\varepsilon},\tilde{\nu})x_{l,k}\right|^2\\ =& -N_p\ln(\pi\sigma_w^2)-\frac{1}{\sigma_w^2}\sum_{(l,k)\in\mathcal{P}}\Big(|y_{l,k}|^2-y_{l,k}x_{l,k}^*e^{-j2\pi\tilde{\varepsilon}\frac{k}{N}}e^{j2\pi l\tilde{\nu}\frac{N_s}{N}}\hat{H}^*(\tilde{\varepsilon},\tilde{\nu})\\ & -y_{l,k}^*x_{l,k}e^{j2\pi\tilde{\varepsilon}\frac{k}{N}}e^{-j2\pi l\tilde{\nu}\frac{N_s}{N}}\hat{H}(\tilde{\varepsilon},\tilde{\nu})+|\hat{H}(\tilde{\varepsilon},\tilde{\nu})|^2\Big).\end{aligned}\tag{111}$$

We want to maximize the log-likelihood with respect to $\tilde{\varepsilon}$ and $\tilde{\nu}$, so we can ignore the terms independent of them and the scaling factor $\frac{1}{\sigma_w^2}$. Then

$$\begin{aligned}\Lambda(\tilde{\varepsilon},\tilde{\nu}) =& \hat{H}^*(\tilde{\varepsilon},\tilde{\nu})\sum_{(l,k)\in\mathcal{P}}y_{l,k}x_{l,k}^*e^{-j2\pi\tilde{\varepsilon}\frac{k}{N}}e^{j2\pi l\tilde{\nu}\frac{N_s}{N}}\\ &+\hat{H}(\tilde{\varepsilon},\tilde{\nu})\sum_{(l,k)\in\mathcal{P}}y_{l,k}^*x_{l,k}e^{j2\pi\tilde{\varepsilon}\frac{k}{N}}e^{-j2\pi l\tilde{\nu}\frac{N_s}{N}}-N_p|\hat{H}(\tilde{\varepsilon},\tilde{\nu})|^2\\ =&N_p|\hat{H}(\tilde{\varepsilon},\tilde{\nu})|^2=\frac{1}{N_p}\left|\sum_{(l,k)\in\mathcal{P}}y_{l,k}x_{l,k}^*e^{-j2\pi\tilde{\varepsilon}\frac{k}{N}}e^{j2\pi l\tilde{\nu}\frac{N_s}{N}}\right|^2.\end{aligned}\tag{112}$$

The maximum likelihood estimates are given by

$$[\hat{\varepsilon}_{\mathrm{ML}}, \hat{\nu}_{\mathrm{ML}}] = \arg\max_{\tilde{\varepsilon},\tilde{\nu}} \Lambda(\tilde{\varepsilon}, \tilde{\nu}). \tag{113}$$

No closed form solution for the estimates is developed. The maximum needs to be solved by doing a two dimensional grid search over all possible pairs $(\tilde{\varepsilon}, \tilde{\nu})$. Therefore, the maximum likelihood estimation is not feasible in practical implementations. Because of the periodicity of the complex exponential function, the estimation range for the carrier frequency offset is

$$|\hat{\nu}| < \frac{N}{2mN_s} \tag{114}$$

and range for the symbol timing offset estimate is

$$|\hat{\varepsilon}| < \frac{N}{2m}, \tag{115}$$

where m is the minimum number of symbols or sub-carriers that the known reference symbols are separated by.

For tracking the slowly time-variant change in frequency offset with high noise power, the estimator is modified to make use of information from the previous time steps. A forgetting factor $0 < \lambda \leq 1$ is used to incorporate reference symbols from longer period. The estimator is now given as

$$[\hat{\varepsilon}_t, \hat{\nu}_t] = \arg\max_{\tilde{\varepsilon},\tilde{\nu}} |\mathbf{R}_t|, \tag{116}$$

where

$$(\mathbf{R}_t)_{ij} = \lambda \sum_{(l,k)\in\mathcal{P}} y_{l,k} x^*_{l,k} e^{-j2\pi\tilde{\varepsilon}_j \frac{k}{N}} e^{j2\pi l\tilde{\nu}_i \frac{N_s}{N}} + (1-\lambda)(\mathbf{R}_{t-1})_{ij}. \tag{117}$$

The maximum search is done over predefined values for $\tilde{\varepsilon}$ and $\tilde{\nu}$. The selection of forgetting factor λ affects to how accurately the time-variant change can be tracked. Small forgetting factor leads to better noise rejection but estimator does not respond to rapid changes in estimation parameters. Different strategies for selecting suitable value for the forgetting factor exist. One approach would be to derive the so called steady state Kalman gain discussed in [16]. Another possibility is to find suitable forgetting factor from simulations. The latter approach is used in the simulations.

When the estimate is used to correct the receiver oscillator, the maximum value in $\mathbf{R}_{t-1}$ points to the error before the correction. Using this value in the filtering causes a delay for the estimation because if the adjustment was correct the true offset is now 0. Therefore, if the estimate is used in a feedback loop, the matrix $\mathbf{R}_{t-1}$ needs to be permuted such that the peak is situated to denote zero carrier frequency offset.

3.3 Frequency domain correlation based estimate

As shown by Theorem 2.11, the carrier frequency offset ν is seen as a common phase rotation between demodulation samples in consecutive OFDM symbols. This phase rotation can be measured from known symbols or from repetitive structure in consecutive OFDM symbols. The OFDM symbols contain scattered pilot symbols in predefined places in time-frequency grid. The pilot symbol phase can then be removed and common phase rotation can be estimated. Measuring the common phase difference from the transmitted data symbols is complicated by the unknown phase of the data symbol. If there is a repetitive structure in consecutive received OFDM symbols, then correlating a data symbol with its repetition removes the symbol phase and the common phase rotation is left. The channel is assumed constant for the two consecutive samples to be correlated and all rotation between two samples is assumed to be due to carrier frequency offset, not from the time varying nature of the channel. This type of estimator is derived in [17], [18] and [19].

This estimation can be extended to estimate the symbol timing offset by measuring the phase rotation between sub-carriers in the same OFDM symbol. If the channel's frequency response is constant for the consecutive samples in frequency domain, the phase rotation is relative to the symbol timing offset ε.

Using scattered pilot symbols

The received OFDM symbols contain known pilots in symbols l and sub-carriers k, where $(l,k) \in \mathcal{P}$. The pilot symbol phase can be removed by multiplying the received sample $y_{l,k}$ with the conjugate of the transmitted pilot $x^*_{l,k}$. We get

$$y_{l,k}x^*_{l,k} = e^{j2\pi\varepsilon\frac{k}{N}}e^{-2\pi l\nu\frac{N_s}{N}}H + w_{l,k}x^*_{l,k}. \tag{118}$$

Because the phase of the noise term is assumed to be uniformly distributed on the interval $[0, 2\pi[$, the term $w_{l,k}x^*_{l,k}$ is also circularly-symmetric complex normal random variable and the multiplication with $x^*_{l,k}$ does not change the distribution of the noise term.

Multiplying these phase de-rotated samples with conjugates, that are m OFDM symbols apart, gives

$$y_{l,k}x^*_{l,k}(y_{l+m,k}x^*_{l+m,k})^* = e^{j2\pi m\nu\frac{N_s}{N}}|H|^2 + w', \tag{119}$$

where the noise term is $w' = e^{j2\pi\varepsilon\frac{k}{N}}e^{-j2\pi l\nu\frac{N_s}{N}}Hw^*_{l+m,k}x_{l+m,k}$
$+ e^{-j2\pi\varepsilon\frac{k}{N}}e^{j2\pi(l+m)\nu\frac{N_s}{N}}H^*w_{l,k}x^*_{l,k} + w_{l,k}x^*_{l,k}w^*_{l+m,k}x_{l+m,k}$. It has zero mean and

expected power

$$\mathbb{E}\{|w'|^2\} = 2|H|^2\sigma_w^2 + \sigma_w^4. \tag{120}$$

This increase in noise power is significant because even small powered noise get amplified by the doubled channel gain in the first term. If the noise power is higher than the signal power the squaring in the second term increases the noise contribution drastically.

Next, the products are summed together and the principal phase angle is measured to give

$$\text{angle}\left(\sum_{(l,k)\in\mathcal{P}'} y_{l,k}x_{l,k}^* y_{l+m,k}^* x_{l+m,k}\right) = 2\pi m\nu\frac{N_s}{N} + p, \tag{121}$$

where p is the phase noise term and $\mathcal{P}'$ is the pilot index set such that last indexes are omitted.

The carrier frequency estimate is given by

$$\hat{\nu} = \frac{N}{2\pi m N_s}\text{angle}\left(\sum_{(l,k)\in\mathcal{P}'} y_{l,k}x_{l,k}^* y_{l+m,k}^* x_{l+m,k}\right). \tag{122}$$

The estimation range for $\hat{\nu}$ becomes

$$|\hat{\nu}| = \frac{N}{2\pi m N_s}\left|\text{angle}\left(\sum_{(l,k)\in\mathcal{P}'} y_{l,k}x_{l,k}^* y_{l+m,k}^* x_{l+m,k}\right)\right| \leq \frac{N}{2mN_s}. \tag{123}$$

For symbol timing offset estimation, we consider the multiplication of pilots adjacent in frequency domain. The multiplication of phase de-rotated samples separated by m sub-carriers is given by

$$y_{l,k+m}x_{l,k+m}^* y_{l,k}^* x_{l,k} = e^{j2\pi\varepsilon\frac{m}{N}}|H|^2 + w''. \tag{124}$$

The noise w'' is also zero mean and has the same expected power as in equation (120). The symbol timing offset estimate is then given by

$$\hat{\varepsilon} = \frac{N}{2\pi m}\text{angle}\left(\sum_{(l,k)\in\mathcal{P}''} y_{l,k+m}x_{l,k+m}^* y_{l,k}^* x_{l,k}\right). \tag{125}$$

Here $\mathcal{P}''$ is the pilot index set with last sub-carrier indices removed. The estimation range for $\hat{\varepsilon}$ is

$$|\hat{\varepsilon}| \leq \frac{N}{2m}. \tag{126}$$

To track time-variant change in carrier frequency offset with high noise power, a forgetting factor $0 < \lambda \leq 1$ is introduced as in previous section. Due to noise statistics not remaining approximately Gaussian after measuring the angle, the averaging should be done in advance. The recursive estimator is given by

$$\hat{\nu}_t = \frac{N}{2\pi m N_s} \text{angle}\left(R_t\right), \tag{127}$$

where

$$R_t = \lambda \sum_{(l,k)\in\mathcal{P}'} y_{l,k} x^*_{l,k} y^*_{l+m,k} x_{l+m,k} + (1-\lambda) R_{t-1}. \tag{128}$$

The forgetting factor is again found from simulations.

For the case, when the estimate is used to continuously correct the oscillator, the value R_{t-1} needs to be phase rotated according to the correction made. If the estimate is used to correct the offset in one step then R_{t-1} needs to be rotated such that it lies on the positive real axis.

3.4 Kalman filter based estimate

The carrier frequency offset and symbol timing offset estimation reduces to estimating a frequency of a complex sinusoid when inspecting the received pilot samples. The unscented Kalman filtering has been used to estimate frequency of a complex sinusoid [20]. In carrier frequency estimation it is discussed in [21].

The received reference symbol samples can be seen as an autoregressive process, given that the channel is almost stationary. The reference symbol sample, with the known symbol phase removed, at time index $l+m$ is given by

$$y_{l+m,k} = e^{-j2\pi(l+m)\nu\frac{N_s}{N}} e^{j2\pi\varepsilon\frac{k}{N}} H + n_{l+m,k} = e^{-j2\pi m\nu\frac{N_s}{N}} y_{l,k} + w_{l+m,k}, \tag{129}$$

where $\mathbb{E}\{|w_{l,k}|^2\} = \sigma_w^2$.

We select the phase $\phi = -2\pi m\nu\frac{N_s}{N}$ as the state parameter we want to estimate. The carrier frequency offset estimate is given by the relation $\hat{\nu} = -\frac{\hat{\phi}N}{2\pi m N_s}$. The state transition model is

$$\phi_l = \phi_{l-m} + v_l, \tag{130}$$

where v_l is zero mean white Gaussian noise with variance Q_l. The variance is also a design parameter to incorporate uncertainty of time dependent change

in ν. With larger variance, the filter reacts faster to the changes in underlying system and smaller variance gives more accurate estimates.

The unscented Kalman filter is initialized by setting $\hat{\phi}_{0|0} = 0$ and $P_{0|0} = P_0$. If we assume that the carrier frequency offset is uniformly distributed to the interval $[-\pi, \pi]$, we can set $P_0 = \frac{\pi^2}{3}$. The state transition model is linear so the original Kalman filter prediction step can be used. The prediction of the state is given by

$$\hat{\phi}_{l|l-m} = \hat{\phi}_{l-m|l-m} \tag{131}$$

and the prediction error variance is

$$P_{l|l-m} = P_{l-m|l-m} + Q_l. \tag{132}$$

For the observation model, the unscented Kalman filter update step is used. The sigma points are calculated as

$$\mathcal{X}_l = [\hat{\phi}_{l|l-m},\ \hat{\phi}_{l|l-m} + \gamma\sqrt{P_{l|l-m}},\ \hat{\phi}_{l|l-m} - \gamma\sqrt{P_{l|l-m}}]. \tag{133}$$

The scaling parameter is given by equation (38). We set $\alpha = 0.5$, $\beta = 2$ and $\kappa = 6$ to keep covariance matrices positive semi-definite.

The predicted observation points are given by passing each sigma point through the mapping

$$\mathcal{Y}_{i,l} = h(\mathcal{X}_{i,l}, y_{l-m,k}) = \begin{bmatrix} \mathrm{Re}(e^{j\mathcal{X}_{i,l}} y_{l-m,k}) \\ \mathrm{Im}(e^{j\mathcal{X}_{i,l}} y_{l-m,k}) \end{bmatrix}. \tag{134}$$

Here the symbol samples are rotated with the phases in sigma point vector, to predict the reference symbol samples in the next time step.

Observation estimate is calculated as

$$\hat{\mathbf{y}}_l = \sum_{i=0}^{2} W_i^m \mathcal{Y}_{i,l}, \tag{135}$$

observation covariance matrix is given by

$$\mathbf{P_y} = \sum_{i=0}^{2} W_i^c \left(\mathcal{Y}_{i,l} - \hat{\mathbf{y}}_l\right)\left(\mathcal{Y}_{i,l} - \hat{\mathbf{y}}_l\right)^T + \mathbf{R}_l \tag{136}$$

and the cross correlation vector

$$\mathbf{P_{xy}} = \sum_{i=0}^{2} W_i^c \left(\mathcal{X}_{i,l} - \hat{\phi}_{l|l-m}\right)\left(\mathcal{Y}_{i,l} - \hat{\mathbf{y}}_l\right)^T. \tag{137}$$

The noise covariance matrix is given by $\mathbf{R}_l = \frac{\sigma_w^2}{2}\mathbf{I}$.

Then, the phase estimate is updated according to

$$\hat{\phi}_{l|l} = \hat{\phi}_{l|l-m} + \mathbf{K}_l(y_{l,k} - \hat{\mathbf{y}}_l) \tag{138}$$

and the variance is updated as

$$P_{l|l} = P_{l|l-m} - \mathbf{K}_l \mathbf{P}_{\mathbf{xy}}^T, \tag{139}$$

where the Kalman gain is given as $\mathbf{K}_l = \mathbf{P}_{\mathbf{xy}}\mathbf{P}_{\mathbf{y}}^{-1}$. The prediction and update steps are carried out for each reference symbol pair that are separated by m OFDM symbols and are on the same sub-carrier.

The symbol timing offset is estimated with a similar scheme. Now, we estimate the phase shift between reference symbol samples in same OFDM symbol but different sub-carriers. The samples are given by

$$y_{l,k+m} = e^{-j2\pi l\nu\frac{N_s}{N}} e^{j2\pi\varepsilon\frac{k+m}{N}} H + n_{l,k+m} = e^{j2\pi\varepsilon\frac{m}{N}} y_{l,k} + n_{l,k+m}. \tag{140}$$

The system state is the phase rotation $\phi = 2\pi\varepsilon\frac{m}{N}$ and the timing offset estimate is given by $\hat{\varepsilon} = \frac{N}{2\pi m}\hat{\phi}$. The algorithm is the same as for carrier frequency offset estimation, except that the observation function h uses reference symbol samples from adjacent sub-carrier instead of adjacent OFDM symbol as

$$\mathcal{Y}_{i,k} = h(\mathcal{X}_{i,k}, y_{l,k-m}) = \begin{bmatrix} \mathrm{Re}(e^{j\mathcal{X}_{i,k}} y_{l,k-m}) \\ \mathrm{Im}(e^{j\mathcal{X}_{i,k}} y_{l,k-m}) \end{bmatrix}. \tag{141}$$

4 Numerical simulations

A simulation model has been developed to evaluate the results numerically. The model is implemented with Anaconda distribution of the Python programming language [22]. It implements the OFDM transmitter and receiver as well as channel as discussed in Section 2.

System parameters

The OFDM symbol transmissions are divided into frames, each containing 14 OFDM symbols and each symbol has 12 sub-carriers for data. The modulation is done using a 128 point fast Fourier transform algorithm, where the unused sub-carriers are filled with zeros. The sub-carrier spacing is selected as 15 kHz and the cyclic prefix is 9 samples long. Reference symbols are mapped according to Figure 1 and selected from the set of quadrature phase shift keying symbols $\{e^{j\frac{\pi}{4}+k\frac{\pi}{2}}\}_{k=0}^{3}$. The symbols inside a frame are

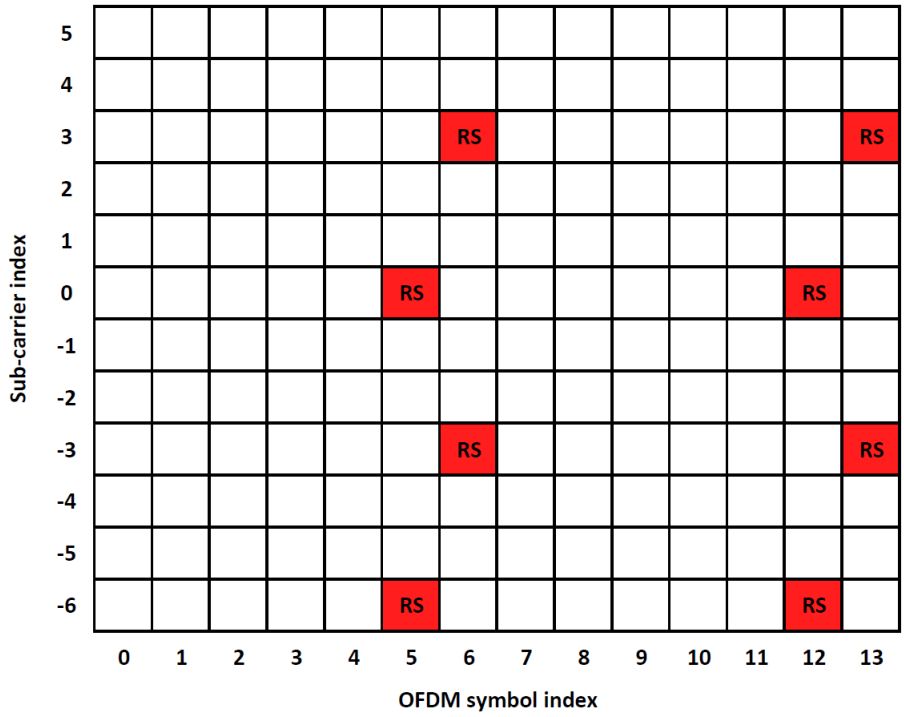

Figure 1: Reference symbol locations in one frame.

transmitted together and are experiencing the same symbol timing and carrier frequency offsets. The sampling frequency is selected as 1.92 MHz. The parameters are selected to reflect the 3GPP narrow-band internet of things (NB-IoT) specification [1].

4.1 Simulations

Estimation mean

To study the mean of the estimators presented in Section 3, a constant carrier frequency offset and constant symbol timing offset is injected to the received signal. The offsets are estimated from 20 frames for maximum likelihood estimates and correlation estimates. The estimator means are measured from 500 simulation runs with signal to noise power ratios of 0 dB and -10 dB. The results are presented in Figures 2 and 3.

We can see that the maximum likelihood and correlation estimators are unbiased in the center of the estimation range and suffer from underestimation near the edges. The increase in noise power causes the area where

Frame	14 symbols × 12 sub-carriers
FFT size	128
Sub-carrier spacing	15 kHz
Sampling frequency	1.92 MHz
Cyclic prefix	9 samples
Symbol duration	71 ns
Frame duration	1 ms

Table 1: Simulation parameters.

underestimation occurs to enlarge. This behavior is explained by the 2π periodicity of the correlation based estimate. The maximum likelihood CFO estimate uses search grid from -1500 Hz to 1500 Hz with 1001 equally spaced points. When the carrier frequency offset is less than -500 Hz or greater than 500 Hz there is a significant peak in cost function 2000 Hz from the actual value which causes the bias in Figure 3. Reducing the estimation grid to the interval $[-1000, 1000]$ removes this bias. The Kalman based estimate is not able to accurately estimate the carrier frequency offset in the signal to noise ratio case of -10 dB. In the case of 0 dB, the estimate has the largest error of the three. This occurs because the state of the estimate is initialized to zero and it does not reach the true value in the estimation period.

For the symbol timing offset, the same conclusion can be made. The maximum likelihood estimate does not show any bias due to the search grid being narrower than the distance between two consecutive peaks in the cost function, see Figures 4 and 5.

Mean squared error

The estimation accuracy is evaluated by calculating the mean squared error. Now the number of available frames for estimation is 20 for every estimator. The lower bounds for the errors are given by the Cramer-Rao lower bounds given by equations (106) and (107). The results are presented in Figures 6 and 7.

It can be seen that the maximum-likelihood estimates reach the Cramer-Rao lower bound when the signal to noise ratio is greater than -8 dB. With higher relative noise powers the mean squared error of the estimators increases rapidly and it does not follow Cramer-Rao lower bound. This is caused by the non-linear threshold effect in the phase noise statistics when the high-powered noise causes abrupt changes to the phase of received samples.

The correlation-based estimate shows the penalty of the noise enhance-

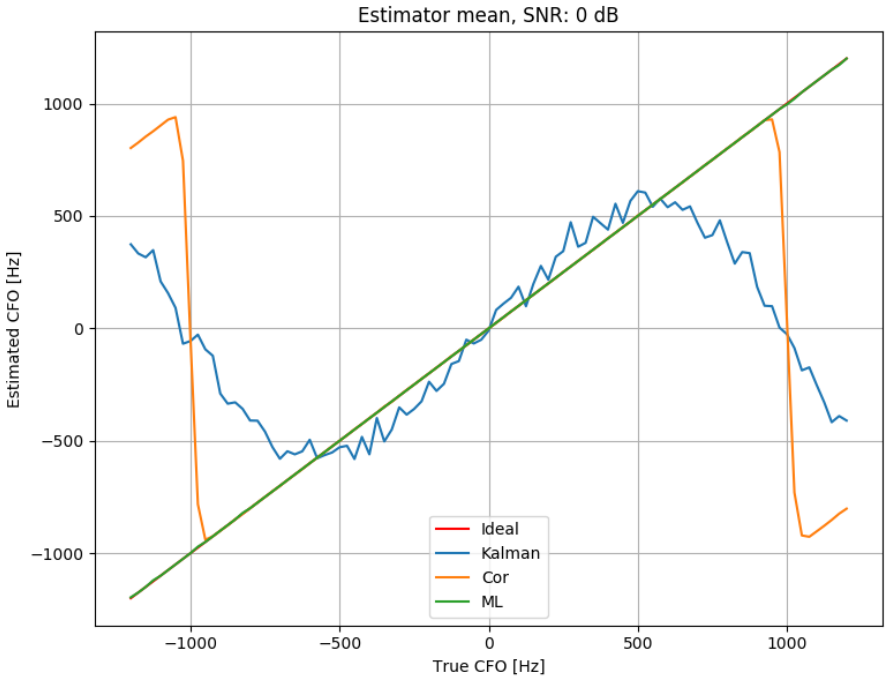

Figure 2: The mean of the carrier frequency offset estimation with signal to noise ratio of 0 dB.

ment due to multiplication of two noisy samples. As a function of signal to noise power ratio, the mean squared error of symbol timing offset estimate is behind the Cramer-Rao bound by the increase in noise power given in equation (120). The mean squared error of carrier frequency offset estimate shows larger signal to noise ratio loss than symbol timing offset estimate.

The Kalman filter based estimate shows poor performance and is not suitable for estimating the residual carrier frequency or symbol timing offsets from short samples. The filter needs sufficiently long time to settle to the correct value of the underlying parameters, which is not reached within 20 frames.

Tracking time-variant change in CFO

The time variant change in the carrier frequency is due to the temperature changes of the crystal oscillator. We assume that the oscillator has a worst-case temperature drift of 0.6 ppm/Hz and the temperature changes at a rate of 0.5 °C/s due to heat dissipating loads, such as the power amplifiers. This

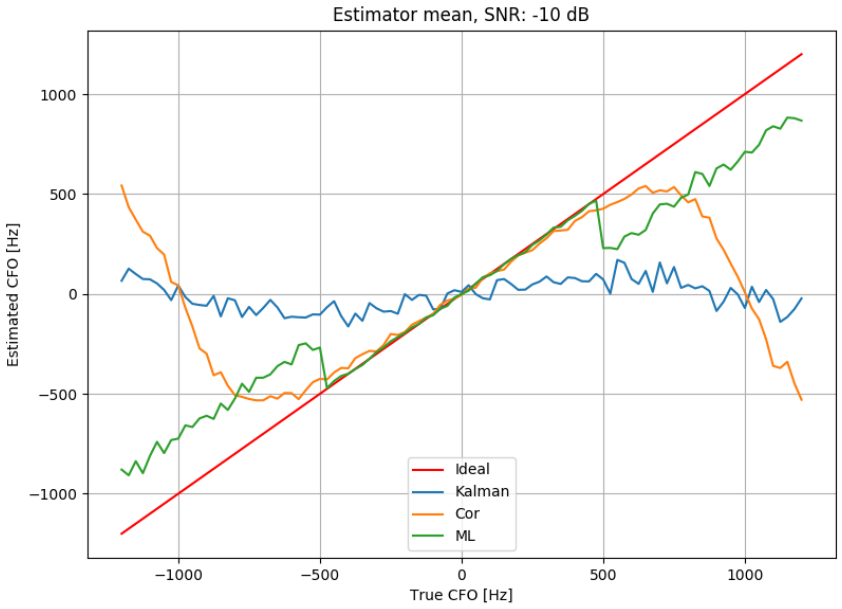

Figure 3: The mean of the carrier frequency offset estimation with signal to noise ratio of -10 dB.

results in the time-variation of generated carrier frequency of 0.6 ppm/s or 750 Hz for carrier frequency of 2.5 GHz. The time-variant behavior is modeled by letting the residual carrier frequency offset change according to a linear slope and the value is updated for each transmitted frame. The target is to have the estimate within 0.1 ppm or in this case 250 Hz at all times. Two cases are studied, one where the carrier frequency offset is only estimated and no corrections are made and one where the estimate is used to correct the receivers oscillator.

In the first case, the forgetting factor is set to 0.01 for maximum likelihood and correlation estimators. The estimation is done from four frames before applying the forgetting factor filtering. The state noise variance for Kalman estimator is set to 0.00001 and estimate is returned after four frames. A time series of the tracking is presented in Figure 8 and the mean squared error is shown in Figure 9. It can be seen that the mean squared error of maximum likelihood estimate and correlation-based estimates show similar relative performance than in the previous case of constant frequency estimation. The Kalman filter shows ability to track the carrier frequency offset at times, but it might lock on to the CFO that corresponds to one full rotation of the reference symbols, i.e. to CFO of ± 2002 Hz with the set parameters. For

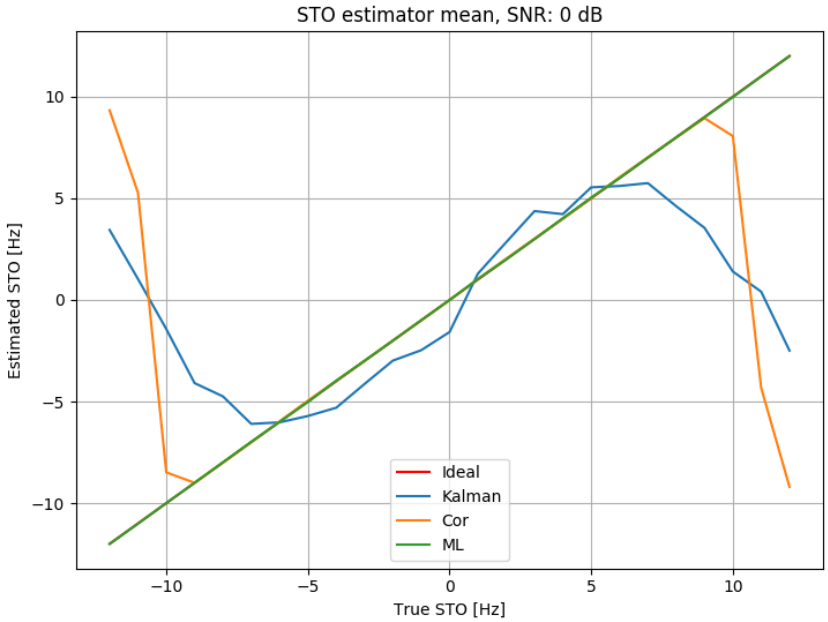

Figure 4: The mean of the symbol timing offset estimation with signal to noise ratio of 0 dB.

this reason, the Kalman based estimate cannot be regarded as very robust estimator.

In the second case, the carrier frequency offset estimate is used to adjust the receiver frequency in order to have the effective carrier frequency offset less than 250 Hz. The parameters for estimators are the same except, when a correction is made, the receiver oscillator and estimators are adjusted accordingly. For maximum likelihood and correlation estimators, the knowledge from the previous time step is adjusted to show zero offset and for Kalman estimator the state is corrected to zero. An initial carrier frequency offset of 600 Hz, which changes at a rate of 750 Hz/s, is induced to the received signal. A time series of how the carrier frequency offset behaves, when the corrections are applied, is presented in Figure 10 and the mean squared error is presented in Figure 11. The maximum likelihood estimate is shown to be the most accurate in the signal to noise power ratio larger than -8 dB. An error floor is reached that is relative to the used forgetting factor. The mean squared error shows more erratic behavior below -6 dB compared to the previous case. This is due to the possibility of estimator correcting the carrier frequency offset to the value of ± 2002 Hz due to periodicity of complex

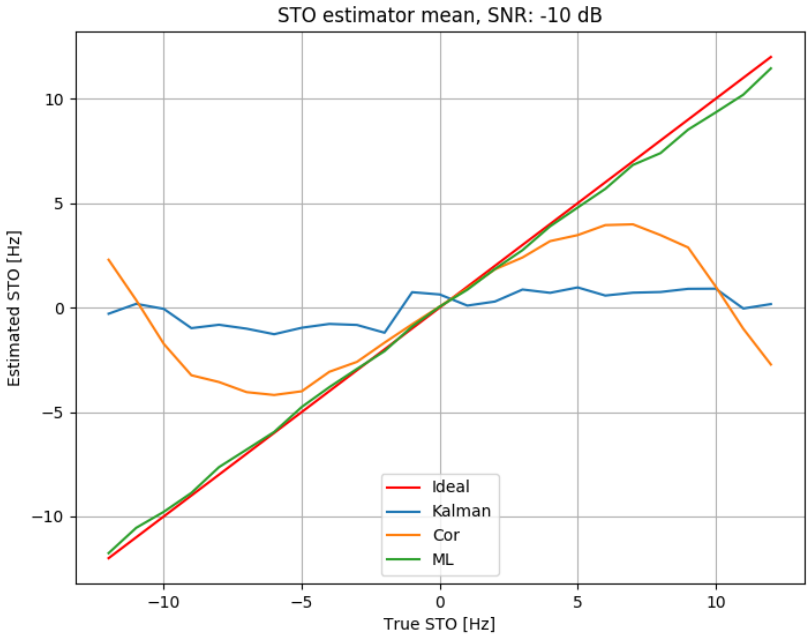

Figure 5: The mean of the symbol timing offset estimation with signal to noise ratio of -10 dB.

exponential function.

4.2 Conclusion

The simulations show that the computationally intensive maximum likelihood estimate is the most accurate estimator of the three. However, it is unsuitable for practical implementations because of the two-dimensional search for peak finding. The correlation-based estimate suffers from the noise enhancement after multiplication. Without any noise reduction filtering, the accuracy is not good enough to keep the carrier frequency error adequately small. The tracking case simulations show that utilizing the estimate from previous time step in continuous tracking increases the performance, such that the required accuracy is reached in a short estimation window. For further performance improvement, an adaptive scheme for selecting the forgetting factor should be studied. With a fixed forgetting factor the performance is optimized for certain signal to noise ratio range and frequency error slope.

The Kalman filter based estimate is not suitable for carrier frequency offset and symbol timing offset estimation according to this study. The number

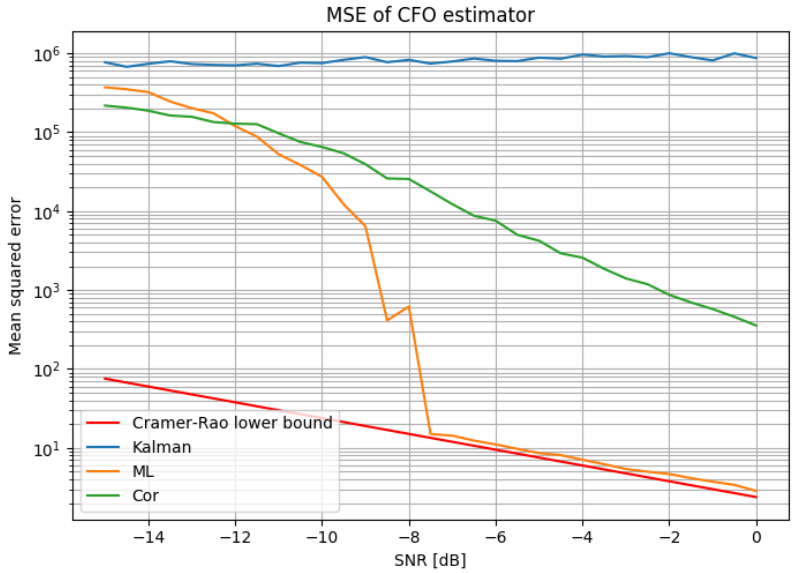

Figure 6: Mean square error of the carrier frequency offset estimator.

of samples required to have the filter settle and the possibility of locking to the wrong value, due to the periodicity of exponent function, makes the estimator unstable. Behavior that is more robust is needed for carrier frequency tracking and the estimator should not worsen the receiver performance.

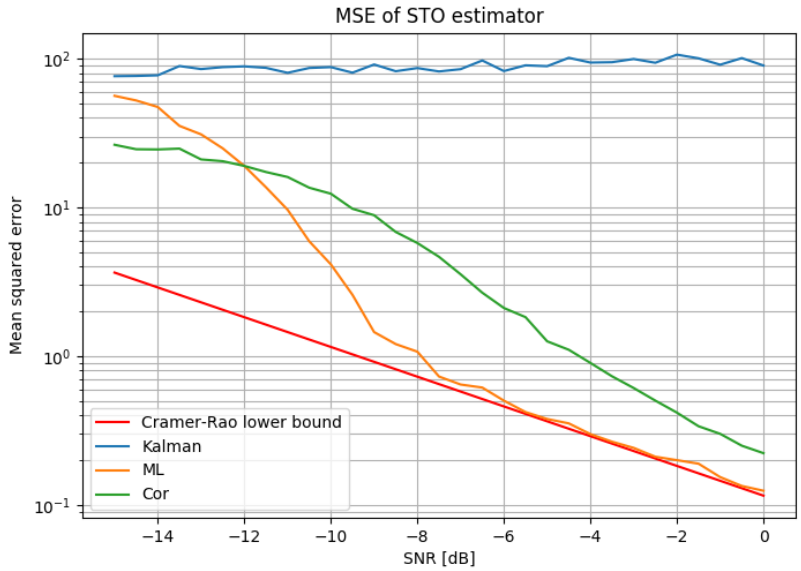

Figure 7: Mean square error of the symbol timing offset estimator.

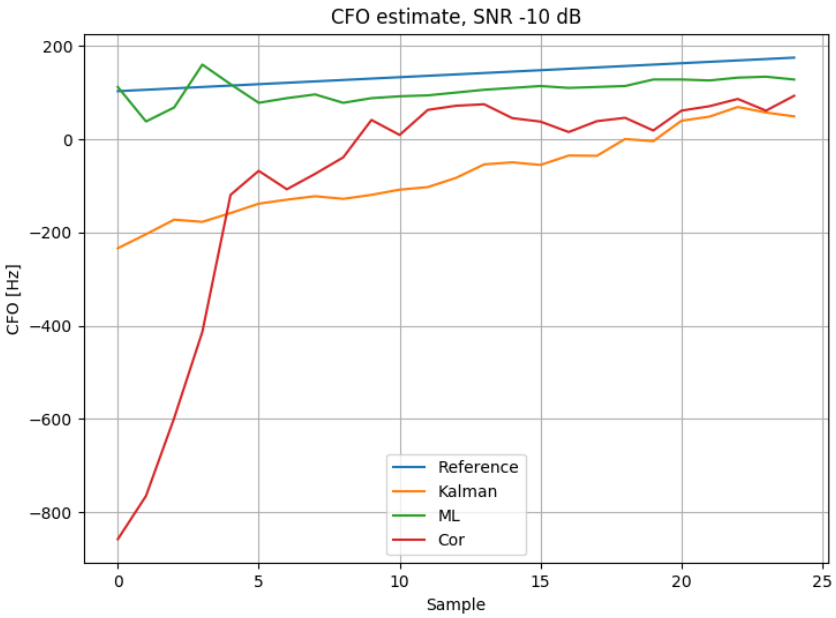

Figure 8: Time series of tracking capability of the estimators.

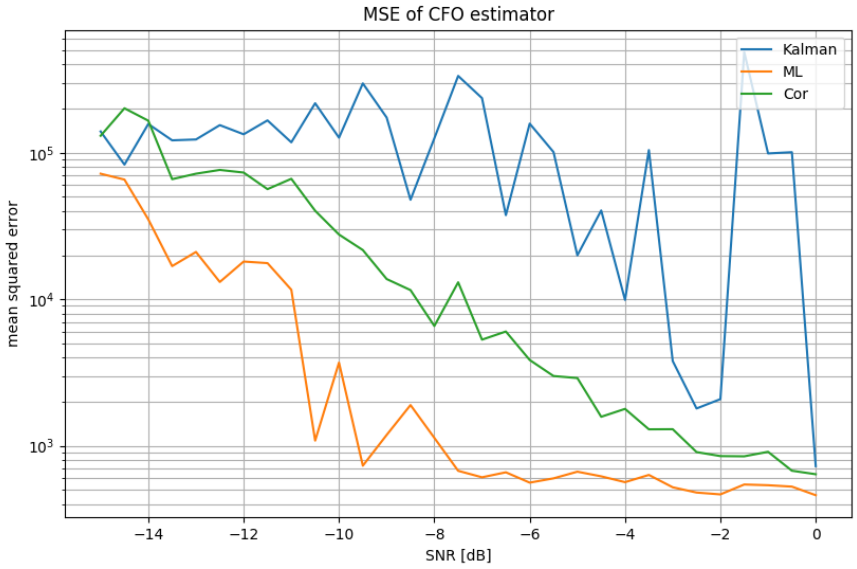

Figure 9: Mean squared error of the carrier frequency offset estimate with induced frequency error of 750 Hz/s.

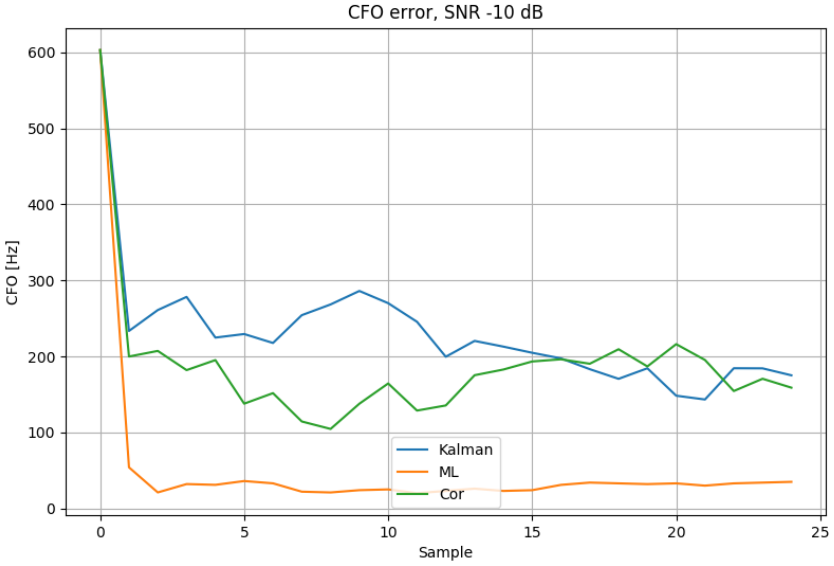

Figure 10: Time series of carrier frequency offset with induced frequency error slope of 750 Hz/s.

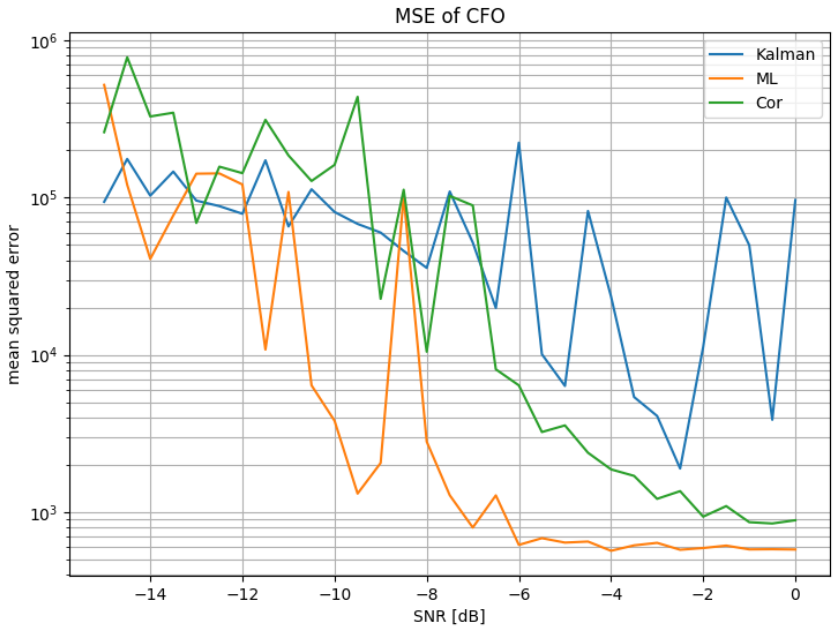

Figure 11: MSE of the carrier frequency offset with induced frequency error slope of 750 Hz/s.

References

[1] *Evolved Universal Terrestrial Radio Access (E-UTRA); Physical channels and modulation*, 3GPP.

[2] I. L. S. Committee *et al.*, *IEEE 802.11-Wireless LAN medium access control (MAC) and physical layer (PHY) specifications.*

[3] E. ETSI, *302 755 V1.4.1*, 2015.

[4] A. C. King, J. Billingham, and S. R. Otto, *Differential equations: linear, nonlinear, ordinary, partial.* Cambridge University Press, 2003.

[5] C. E. Shannon, "Communication in the presence of noise," *Proceedings of the IRE*, vol. 37, no. 1, pp. 10–21, 1949.

[6] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume 1: Estimation Theory.* Pearson Education, 2009. [Online]. Available: https://books.google.no/books?id=pDnV5qf1f6IC

[7] R. Van Der Merwe, "Sigma-point Kalman filters for probabilistic inference in dynamic state-space models," Ph.D. dissertation, Oregon Health & Science University, 2004.

[8] R. E. Kalman *et al.*, "A new approach to linear filtering and prediction problems," *Journal of basic Engineering*, vol. 82, no. 1, pp. 35–45, 1960.

[9] S. J. Julier, "The scaled unscented transformation," in *American Control Conference, 2002. Proceedings of the 2002*, vol. 6. IEEE, 2002, pp. 4555–4559.

[10] M. Speth, S. A. Fechtel, G. Fock, and H. Meyr, "Optimum receiver design for wireless broad-band systems using OFDM. I," *IEEE Transactions on communications*, vol. 47, no. 11, pp. 1668–1677, 1999.

[11] J. Proakis, *Digital Communications*, ser. Electrical engineering series. McGraw-Hill, 2001. [Online]. Available: https://books.google.no/books?id=sbr8QwAACAAJ

[12] J.-J. Van De Beek, O. Edfors, M. Sandell, S. K. Wilson, and P. O. Borjesson, "On channel estimation in OFDM systems," in *Vehicular Technology Conference, 1995 IEEE 45th*, vol. 2. IEEE, 1995, pp. 815–819.

[13] T. Pollet, M. Van Bladel, and M. Moeneclaey, "BER sensitivity of OFDM systems to carrier frequency offset and Wiener phase noise," *IEEE Transactions on communications*, vol. 43, no. 234, pp. 191–193, 1995.

[14] M. Morelli and M. Moretti, "Fine carrier and sampling frequency synchronization in OFDM systems," *IEEE Transactions on Wireless Communications*, vol. 9, no. 4, pp. 1514–1524, 2010.

[15] C. Oberli, "ML-based tracking algorithms for MIMO-OFDM," *IEEE Transactions on Wireless Communications*, vol. 6, no. 7, 2007.

[16] M. I. Ribeiro, "Kalman and extended Kalman filters: Concept, derivation and properties," *Institute for Systems and Robotics*, vol. 43, 2004.

[17] M. Speth, S. Fechtel, G. Fock, and H. Meyr, "Optimum receiver design for OFDM-based broadband transmission. II. a case study," *IEEE Transactions on communications*, vol. 49, no. 4, pp. 571–578, 2001.

[18] K. Guo, W. Xu, and G. Zhou, "Differential carrier frequency offset and sampling frequency offset estimation for 3GPP LTE," in *Vehicular Technology Conference (VTC Spring), 2011 IEEE 73rd.* IEEE, 2011, pp. 1–5.

[19] M. Luise and R. Reggiannini, "Carrier frequency recovery in all-digital modems for burst-mode transmissions," *IEEE Transactions on Communications*, vol. 43, no. 2, 3, 4, pp. 1169–1178, 1995.

[20] P. Dash, S. Hasan, and B. Panigrahi, "Adaptive complex unscented Kalman filter for frequency estimation of time-varying signals," *IET science, measurement & technology*, vol. 4, no. 2, pp. 93–103, 2010.

[21] T. Kang and R. A. Iltis, "Iterative decoding, offset and channel estimation for OFDM using the unscented Kalman filter," in *Signals, Systems and Computers, 2007. ACSSC 2007. Conference Record of the Forty-First Asilomar Conference on.* IEEE, 2007, pp. 1728–1732.

[22] Anaconda software distribution. ContinuumAnalytics. [Online]. Available: https://continuum.io