

# **Statistical Modeling of Real Inductor Impedances**

*Jianhang Duan*

A dissertation submitted in partial fulfillment  
of the requirements for the degree of  
**Undergraduate of Statistics**  
of  
**University College London.**

Department of Statistical Science  
University College London

April 12, 2023

I, Jianhang Duan, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the work.

# Abstract

Inductors are electronic components whose impedance as a function of frequency is challenging to model. This study compares a sequence of six models of increasing complexity. **Starting** from the RLC model which adds a resistor (R) and a capacitor (C) to an ideal inductor to approximate an imperfect inductor. Bayesian optimisation is used for maximum likelihood parameter estimation. While the model fits qualitative trends well, **their** residuals are clearly structured, and fit is poor near the self-resonance frequency. In addition, **I** will utilise statistical hypothesis testing to evaluate the influence of the coupling effect between series-connected inductors.

# Acknowledgements

Acknowledge all the things!

# Contents

<b>1</b>	<b>Introduction</b>	<b>7</b>
<b>2</b>	<b>Introduction to Electronics for Statisticians</b>	<b>9</b>
2.1	Alternating Current & Direct Current . . . . .	9
2.2	Electric Components . . . . .	9
2.3	Ohm's Law . . . . .	10
2.4	Insulator & Conductor . . . . .	11
2.5	Electric field . . . . .	12
2.6	Capacitor . . . . .	13
2.6.1	Capacitive reactance . . . . .	14
2.6.2	Q-factor of capacitors . . . . .	14
2.7	Magnetic Field . . . . .	15
2.8	Inductor . . . . .	15
2.8.1	Quality-factor of Inductor . . . . .	20
2.9	Reactance & Impedance . . . . .	20
2.10	Parallel Circuit & Series Circuit . . . . .	21
2.11	Resonance . . . . .	23
<b>3</b>	<b>Statistical Methodologies</b>	<b>24</b>
3.1	Non-linear Models . . . . .	24
3.2	Optimisation . . . . .	25
3.2.1	Newton method . . . . .	26
3.3	Broyden–Fletcher–Goldfarb–Shanno Method . . . . .	27



3.4	Bayesian optimization . . . . .	30
3.5	Evaluation Methods . . . . .	33
3.5.1	Log-likelihood function . . . . .	33
3.5.2	Akaike Information Criterion . . . . .	34
<b>4</b>	<b>Data &amp; Research Questions</b>	<b>37</b>
4.1	Research Questions . . . . .	38
<b>5</b>	<b>Results</b>	<b>40</b>
5.1	Simulation of 4 Microhenry Toroidal Inductor . . . . .	40
5.2	Simulation of 0.5 Microhenry to 2 Microhenry Cylindrical Inductor	55
5.3	Modeling the Low Frequency Range . . . . .	57
<b>6</b>	<b>Inductance for 2 inductors connected in series</b>	<b>59</b>
<b>7</b>	<b>Conclusion</b>	<b>62</b>
7.1	Future work . . . . .	63
	<b>Appendices</b>	<b>64</b>
<b>A</b>	<b>Appendix</b>	<b>64</b>
	<b>Bibliography</b>	<b>67</b>

## Chapter 1

# Introduction

Inductors serve multiple purposes and are one of the most prevalent and widely used electronic components in daily life. Given the special characteristic of resonance of the inductor, it can be used as a sensor to take a specific signal frequency when combined with a capacitor. Apart from that, When multiple inductors are combined in the electric systems, transformer that altering the power of transmission can be formed due to magnetic coupling effect. Almost all electronic circuits contain inductors, whose primary function is to store energy in the form of magnetic field. Therefore, modeling an inductor is important and necessary when employing it. The behaviour of the inductor and whole electric circuit under various inductance becomes predictable, thereby preventing situations where they overheat or fail to function.



The objective of this project is to find a appropriate model that can simulate the inductor's impedance properly. In this study, I will begin with a rudimentary conceivable model that taken into account different types of inductor and frequency, and then applying statistical optimization (Quassi-Newton and Bayesian Optimisation method) to find the maximum log-likelihood estimate. Subsequently, I will continuously update the current model based on residual images from previous models. After attempting several non-linear models, the optimal outcome is determined based on the information criterion and significance results.

Traditionally, it was possible to create an actual circuit in the physical world and simulate the behaviour of the intended electronic component, such as the in-



ductance or the impedance. Simulating the behaviour of electronic components can help us learn how the electronic component respond at different frequencies. However, inductor simulation is notoriously difficult in physics due to inductor resonance and magnetic saturation. Kolster [27] began modelling inductors in the early 1910s, but the predictive power of the model he created was limited. In order to model the impedance of inductors with varying degrees of inductance, numerous models must be tested. Using conceivable non-linear models to imitate the actual electronic components is of great research interest in order to save a substantial amount of circuit construction time.

At the end of the result section of this report, I will use the log-likelihood ratio test to determine whether series-connected inductor coupling effect exists in our measured data. Which simply tests whether series-connected inductors, such as  $1\ \mu H$  and  $4\ \mu H$ , produce a theoretical inductance of  $5\ \mu H$  in an electric circuit. If the existence of the coupling effect can be verified, what is the extent of the coupling effect, and if multiple inductors are modeled, can their electromagnetic field interactions be ignored?

The report has the following format: Chapter 2 provides an introduction to the fundamental concepts of electronics and physics. The focus of Chapter 3 was on the statistical optimization techniques and their implementation, which will be used to enhance the non-linear models by finding the maximum log-likelihood estimate. The fourth chapter is an explanation of the data sets and the rationale for our research topics. Chapter 5 covered the reason and process of getting the current 6 candidate models, and numerous statistical results, including residual plots and kolmogorov-smirnov test are employed to help me evaluate the model's quality. Chapter 6 focuses mainly on an investigation into whether the impedance of two series-connected inductors would behave as anticipated. chapter 7 summarised the previous statistical observations, and drew the final conclusion, and provided a prognosis for future work.



## Chapter 2

# Introduction to Electronics for Statisticians

This part intends to offer a basic introduction to the physics theory of circuitry, which is required for the subsequent statistical analysis. It is intended to be a summary with the objective of providing an introduction that statisticians can understand. The construction of the circuit and its constituent electronic components are described in this chapter.

## 2.1 Alternating Current& Direct Current

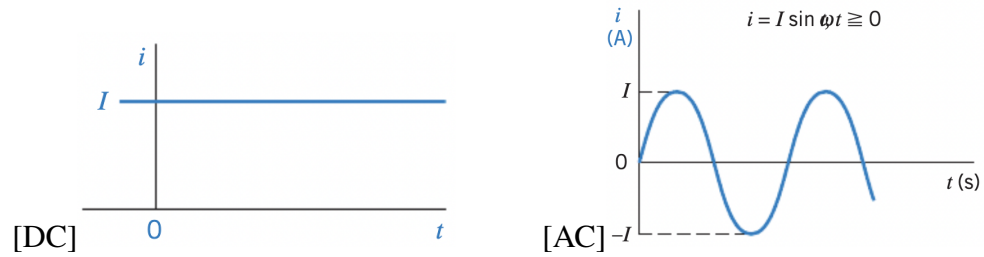
Direct current (DC) is a current of constant magnitude[45]. The direction of the direct current in the circuit will not change over time. Typical direct current sources include batteries, cells, solar cells, etc.

Alternating current(AC) is the exact opposite of direct current, which implies that the direction of the current changes regularly throughout time. Figure 2.1 represents the alternating current as a sinusoidal curve that fluctuates regularly along the time axis[40].

## 2.2 Electric Components

We often categorise electronic components into two primary types. The first category consists of active electronic components, which can be thought of as power transmitters. In contrast to active **electronics**, passive **electronics** do not convey





**Figure 2.1:** Electric current against Time [40]

power to our circuits [41].

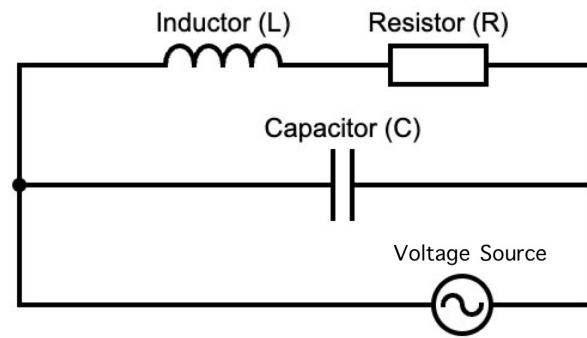
- Active components need a source of energy, such as an alternating current (AC) or direct current (DC) power supply, in order to function and deliver current or other forms of energy to our circuits. A generator is a common active component that supplies electricity to our circuits.
- The passive component in our circuit does not provide any power. For instance, there are resistors, inductors, and capacitors.

## 2.3 Ohm's Law

Only passive components will be considered for this project. Typically, these components fall into one of three categories: resistors, capacitors, and inductors, each of which will be discussed in succession. Ohm's law, which is one of the basic rules of electricity, was found by Georg Simon Ohm (1789-1854), a German physicist, and plays a significant role in the practical electrical applications and it will help us comprehend resistors[25].

Ohm's law states that if we have a conductor and precisely measure the voltage (V) and current (A) at both ends, we will discover that they are proportional to one another[19]. The ampere (A) is the standard unit of current, and its symbol is I. Resistance (R) and Voltage (V) are measured in Ohms and Volts, represented respectively by the Greek letter Omega and the letter V.

$$V(V) = I(A)R(\Omega) \quad (2.1)$$



**Figure 2.2:** Simple Electric Circuit (RLC)

More generally,

$$R = V/I. \quad (2.2)$$

Figure 2.2 depicts a circuit that can be viewed as a route through which an electrical current flows, with connected electronic components. This specific combination of electronic components is referred to as an RLC circuit, which consists of a resistor, an inductor, and a capacitor.

As shown in Figure 2.2, the resistance symbol in the circuit is a small rectangle, whose main function is to provide resistance to the circuit. We can conceive of resistance in a circuit as an impediment to the flow of electric current. The greater the resistance, the greater the degree to which the current is prevented from passing.

## 2.4 Insulator & Conductor

By categorising a material's resistance, we typically arrive at two broad categories: conductor and insulator[23].

- A conductor is a substance with low resistance that readily transmits electricity. A significant number of charge carriers, which are charged particles that move freely, exist in the conductor. Typically, metals are excellent conductors, with silver being the finest.
- A insulator is a substance with extremely high resistance. The tremendous resistance makes it difficult for electricity to get through. Common insulators include rubber, wood, air, ceramic, etc.

The resistance of the resistor in the electric circuit can be easily calculated by applying Ohm's law. We can use the voltmeter and ammeter to measure the voltage and current of the terminals of the resistor, and put our measured data into the formula of Ohm's law.

## 2.5 Electric field

Everything is composed of atoms, which consist of electrons and positively charged nuclei. Letter  $q$  represent the electric charge of both the electron and the proton;  $q = 1.60217663410 \times 10^{19}C$ , where C is the unit of charge known as the coulomb. In contrast to the electron, which has a relative charge of -1, the proton has a relative charge of 1.

According to a basic aspect of physics, unlike charges attract and like charges repel. Positively charged particles are surrounded by their own electric field. If two positively charged particles are brought together, their electric fields will interact and cause displacement. Therefore, the electric field can be regarded as the interaction force of charged particles.

In physics, force is a vector, thus we may comprehend that electric fields are physical fields with direction and magnitude. This force, represented by  $\mathcal{F}$ , is defined by [43]

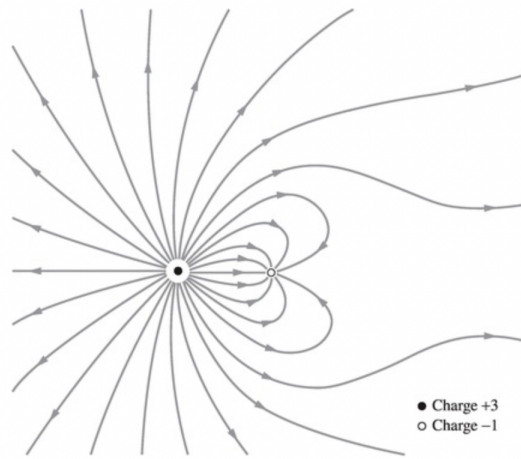
$$\mathcal{F} = qF, \quad (2.3)$$

where F is the electric field strength and defined by

$$F = \frac{Q}{4\pi\epsilon_0 r^2}. \quad (2.4)$$

Here,  $\epsilon_0$  represent the permittivity of the environment (often air with a permittivity of 1.00051) and  $r$  is the distance between positively charged particles and the location of another negatively charged particle[24].

Figure 2.3 demonstrates the electric field around two charged particles, and it is clear that the positively charged particle's electric field is directed toward the negatively charged particle; as the two charged particles carry opposite charges,



**Figure 2.3:** Electric field[42]

they are attracted to one another.



## 2.6 Capacitor

The components of a capacitor are **dielectric material** and two metal plates. The medium between the two metal plates is **called dielectric material**. Since dielectric material has an extremely low electrical conductivity, it is also an insulator.

The electrical sign for capacitor in RLC circuit (Figure 2.2) consists of two parallel vertical lines, which also symbolise the capacitor's constituent metal plates. When a capacitor is connected to a power source, current flows into it, charging its plates. One of the metal plates is positively charged while the other is negatively charged. Therefore, when our plates are charged, an electric field is produced between the two plates and the charge is stored. Because the electric field between the metal plates does not dissipate when the power supply is disconnected, our capacitor continues to output energy.

The relationship between capacitance (C), current (I) and voltage (V) can be represented as [20]

$$I = C \frac{dV}{dt}. \quad (2.5)$$

### 2.6.1 Capacitive reactance

Reactance (X) is the obstructive effect of an inductor and capacitor on the alternating current in a circuit [46]. Examples of reactive components are inductor and capacitor. Later we shall explore inductive reactance, but for now we will simply consider capacitive reactance ( $X_C$ ). The property of reactance is identical to that of resistance, whereas resistance is a real number, reactance is a complex number. We can derive the formula of capacitive reactance using Ohm's law (Equation 2.2), which we mentioned earlier.



All voltages and currents in a linear circuit have the same angular frequency  $\omega$  as the power source. Such voltage (V) can be expressed in Euler's formula

$$V = e^{j\omega t} = \cos(\omega t) + j\sin(\omega t), \quad (2.6)$$

where  $j$  represents the imaginary unit in complex numbers and angular frequency  $\omega$  is defined as

$$\omega = 2\pi f. \quad (2.7)$$

Letter  $f$  appeared in this equation is denoted as the frequency of the alternating current.

When we substitute the voltage in Euler's form into Equation 2.5, we obtain

$$I = C \frac{de^{j\omega t}}{dt} = C(j\omega e^{j\omega t}) = Cj\omega V. \quad (2.8)$$

According to Ohm's law (Equation 2.2), Equation 2.8 can be rearranged to give the capacitive reactance ( $X_C$ ).

$$\frac{V}{I} = \frac{1}{Cj\omega} = X_C \quad (2.9)$$

### 2.6.2 Q-factor of capacitors

Quality factor is abbreviated as Q-factor which indicates the efficiency of capacitors. A capacitor is similar to a battery in that it stores electrical energy. The energy intake and output of a capacitor with a high level of efficiency should be relatively similar. Nevertheless, if the difference between the two values is too large, which means the

majority of the energy stored in the capacitor will be lost as heat.

When considering a capacitor in series with a resistor, the following formula is used to compute the Q-factor of the capacitor. By substituting the definition of capacitive reactance in Equation 2.9 we get

$$Q = \frac{|X_c|}{R_s} = \frac{1}{2\pi f C R_s}. \quad (2.10)$$

$X_L$  denotes capacitive reactance, whereas  $R_s$  represents series resistance.

## 2.7 Magnetic Field

Any magnetic object has two poles, denoted by N for the North Pole and S for the South Pole. These two poles are located at either end of the magnetic item and are the most magnetic locations on the magnet or magnetic object as a whole. **When a magnetic item is divided in half, two new magnets with their own north and south poles are created at the moment of the split** [10].



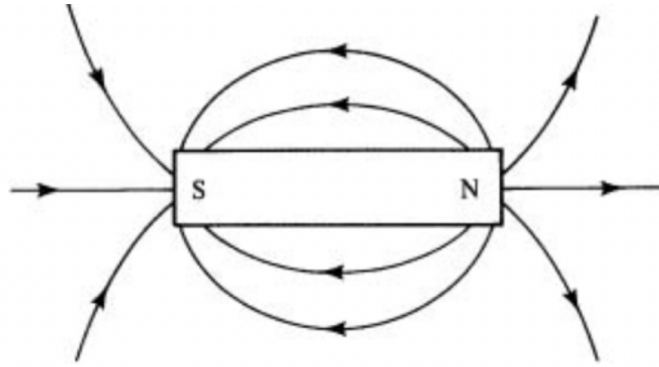
By spinning or orbiting an atomic nucleus, charged particles that are in motion generate a magnetic field. Electrons in non-magnetic objects are also in motion, but their motion is primarily random, so the magnetic fields between the electrons cancel out and the item is not magnetic. The reason that our bar magnet has a magnetic field is because its electrons are moving in unison.

Figure 2.4 depicts a bar magnet, where the curve represents the magnet's magnetic field and the arrow represents the magnetic field's direction. Each magnetic field emanated from the North pole of the bar magnet to the South pole, and the closer the magnetic field was to the magnetic pole of the magnet, the stronger it was.

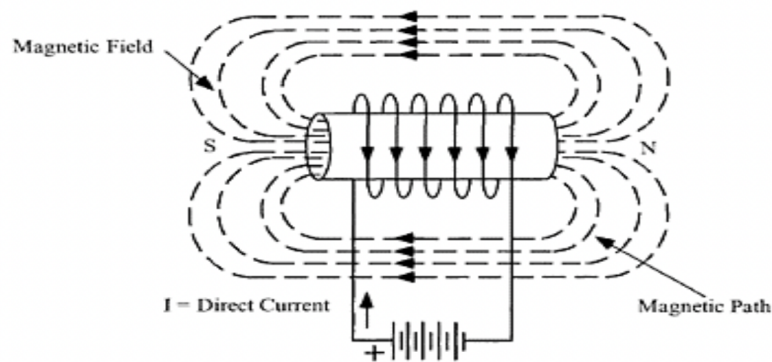
## 2.8 Inductor

The inductor is the final and most crucial electronic component that needs to be modelled for this research.

Figure 2.5 depicts a circuit with only an inductor, and from this figure we can see that the inductor is made up of wire that has been wrapped around the core



**Figure 2.4:** Bar Magnet[10]



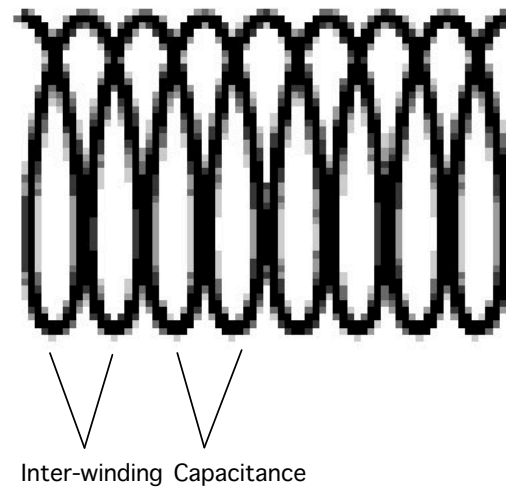
**Figure 2.5:** Iron Core Inductor[33]

in a cylindrical shape. When current  $I$  flows through our coil, a magnetic field is produced (Dotted line surrounding the coil in Figure 2.5), and the inductor stores energy within the magnetic field. This phenomenon is referred to as inductance, and the unit of inductance is Henry (H). Inductors are also known as coils. The electric symbol of an inductor is given in Figure 2.2, it is just like a string of bent wire.

The capacitor utilised in this study differs slightly from the one introduced previously. The capacitor utilised in this instance is known as an Inter-winding Capacitor. The metal plate is replaced with a single coil without the core. Inter-winding capacitance is created between neighboring turns of the wire (Shown in Figure 2.6).

The core of an inductor is composed of a variety of materials, including iron, air, copper, etc. However, the inductance of various materials varies. The vast majority of substances are poor conductors of magnetic flux; their permeability is very





**Figure 2.6:** Inter-winding Capacitance [26]



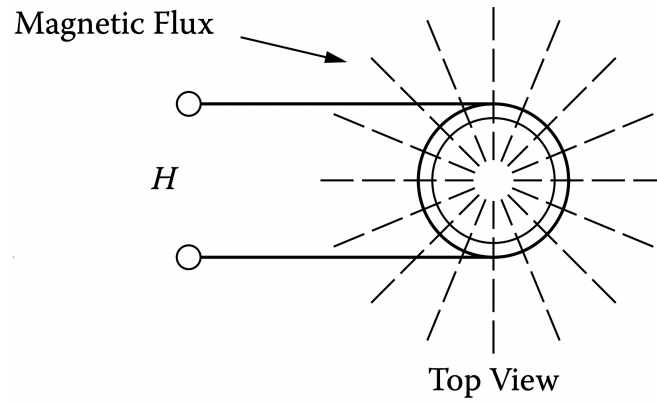
**Figure 2.7:** Comparison of Cylindrical inductor and Toroidal inductor

low. Air, paper, and copper, which are all nonmagnetic, have the same permeability as vacuum. Each has an extremely low permeability of 1.0. Nevertheless, there are ferromagnetic materials, such as iron and cobalt, with a high permeability of up to tens of millions. In this research, iron powder is the primary material utilised for the toroidal core which is built of coils encircled by circular toroid (Figure 2.72).



Air-coil (Cylindrical inductor without the core) inductor shown in Figure 2.71 are not wrapped in any ferromagnetic material; they are supported mechanically by the wires themselves. Unlike cylindrical inductor, the coils on the toroid are firmly connected and form a circle, there is no gap between the coils and the magnetic field is concentrated on the toroid. In general, toroidal inductors are more efficient than cylindrical inductors because their Q-factor is greater.

Magnetic flux is a vector that represents the magnetic field's direction and strength. When a magnetic object is placed in a magnetic field, the force on the



**Figure 2.8:** Magnetic flux from top vies.[34]

object can be determined by studying the magnetic flux. Figure 2.8 illustrates the magnetic flux lines from top view of the inductor in Figure 2.5.

Magnetic flux ( $\Phi$ ) of the electromagnet is defined by

$$\Phi = \frac{NI}{\frac{l_{mag}}{\mu_0 + \mu_r S} + \frac{d}{\mu_0 S}}. \quad (2.11)$$



$S, N, l_{mag}, d, \mu_0, \mu_r$  represent, respectively, the cross-sectional area of the coil, the number of turns of the coil, the length of the magnetic core, the length of the air gap, the permeability of air (Typically assumed to be 1), and the permeability of the magnetic core [49].

There are numerous ways to express inductance, but the expression related to magnetic flux is

$$L = \frac{N\Phi}{I} = \frac{N(BS)}{I}. \quad (2.12)$$

As shown by Equation 2.11, when ferromagnetic materials such as iron is utilized as magnetic cores, the magnetic flux will increase, thereby increasing the inductance of the inductor. This is because the magnetic flux is positively proportional to the inductance in the expression of Equation 2.12. In contrast, if we use materials with low permeability, both the magnetic flux and the inductance will decrease.

In addition to the core material, various other variables can influence the magnetic flux, such as the coil's cross-sectional area and the number of turns, both of which are proportional to the magnetic flux. In other words, the greater the cross-

sectional area of the coil and the greater the number of turns of coils, the greater the strength of our magnetic flux.

Electromotive force (EMF) and Voltage both represent the generated potential difference, so we assume they are equivalent. The relationship between output voltage and the time derivative of the total magnetic flux can be explained by Faraday's law which is one of the fundamental law of inductors. We can calculate the output voltage from inductor using Faraday's law which is defined as

$$\varepsilon = -N \frac{d\Phi_b}{dt}, \quad (2.13)$$

where  $\varepsilon$  represent the output voltage or electromotive force,  $N$  is the number of coils,  $\Phi_b$  is the magnetic flux, and  $t$  represent the time. According to Faraday's law of induction, When the current through the coil changes, the resulting magnetic field changes over time, inducing an electromagnetic force, which then creates the voltage in the conductor.

The magnetic flux, which was previously defined in Equation 2.11, also has a shorter form, which is  $\Phi_b = BS$ . In this expression,  $B$  represent the Magnetic Flux Density, and  $S$  represent the cross sectional area of the coil [29]. During our research we consider these two parameters to be constant. Thus, we can simplify the formula further.

$$\varepsilon = -NS \frac{dB}{dt} \quad (2.14)$$

Inductive reactance can be found by the formula

$$X_L = j\omega L. \quad (2.15)$$

Inductance is represented by the letter  $L$  in this formula, while  $\omega$  and  $j$  have already been defined in subsection 2.61.

### 2.8.1 Quality-factor of Inductor

Q-factor of in inductor is the ratio between inductive reactance ( $X_L$ ) and series resistance ( $R_s$ ). When using alternating current with a high frequency, the performance of inductors with a high Q-factor is closer to our ideal circumstance and their performance is more stable.

$$Q = \frac{\omega L}{R_s} = \frac{2\pi f L}{R_s} \quad (2.16)$$

When we use ferromagnetic material as the inductor's core or increase the cross-sectional area of the coil, the inductance will increase, thereby increasing the Q-factor, and the inductor's quality can be enhanced.

In reality, there is both inductance and resistance of an inductor. Resistance is the primary cause of energy loss in inductors. Due to winding resistance, the passing speed of electrons will be hindered when they attempt to pass through the wire. As electrons collide, kinetic energy is transformed into heat, resulting in a loss of energy. The iron core's heat loss is more than that of the air core. The changing magnetic field created by alternating current is the primary cause of the core loss of Iron power core, which ultimately results in greater heat loss [13].

## 2.9 Reactance & Impedance

When inductor and capacitor are connected in a series circuit, the total electrical reactance of the entire circuit  $X$  can be calculated and expressed as

$$X = X_L + X_c = j\omega L + -\frac{j}{\omega C}. \quad (2.17)$$

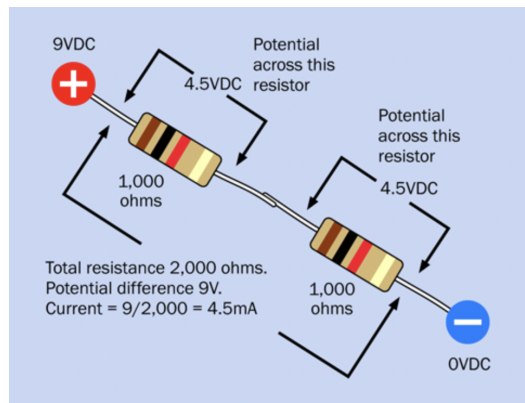
Impedance is the opposition to the alternating current in a circuit that results from the interaction of resistance and reactance [1].

The Impedance itself is a complex number that consists of both real and imaginary parts. Impedance is defined as

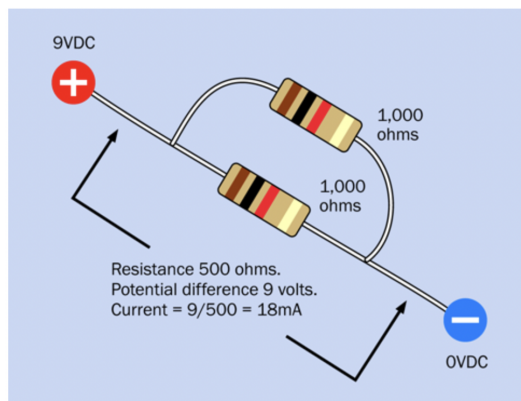
$$Z = R + jX. \quad (2.18)$$

The real part of the impedance is the resistance, the imaginary part represent the total reactance(X). In general, a component whose impedance contains a non-zero imaginary part is referred to be "reactive".

## 2.10 Parallel Circuit& Series Circuit



**Figure 2.9:** Electric Circuit in series[39]



**Figure 2.10:** Electric Circuit in series[39]

The most fundamental two circuit types, series circuit and parallel circuit, are used in applications ranging from simple circuits to complicated circuits.

- Electric components in series circuit are connected one follow the other. It is shown in Figure2.9
- Electric components in parallel circuit are connected side by side. It is shown in Figure2.10

In a series circuit, the current remains constant throughout, and the voltage of the circuit equals the sum of the branch voltages; hence, the voltage will be divided by the series circuit. In contrast, in a parallel circuit, the voltage is the same everywhere, and the current of the circuit is the sum of the branch currents.

Considering series circuit. Assume number  $N$  resistors connected in series, then we can find the total resistance and voltage of the circuit.

$$R_{total} = (R_1 + R_2 + R_3 + \dots + R_N) \quad (2.19)$$

The voltage of an electric circuit is the sum of the voltages measured at each component's two terminals ( $V_1, V_2, \dots, V_N$ ).

$$V = (V_1 + V_1 + V_3 + \dots + V_N) \quad (2.20)$$

Current in the circuit can be calculated using Ohm's Law (Equation 2.2).

$$I = \frac{V}{(R_1 + R_2 + R_3 + \dots + R_n)} \quad (2.21)$$

Figure 2.10 shows a parallel circuit with two resistors connected side by side. Current in the electric circuit shown by Figure 2.10 has two routes to run. The combined current is equal to the sum of the currents in each branch.

$$I_{total} = (I_1 + I_2) \quad (2.22)$$

Each branch's voltage is the same as the supply voltage.

$$V_{supply} = V_{branch1} = V_{branch2} \quad (2.23)$$

The total resistance for the parallel circuit is

$$\frac{1}{R_{total}} = \left( \frac{1}{R_1} + \frac{1}{R_2} \right). \quad (2.24)$$

Using the same strategy, the reactance of a circuit can be computed by substituting

the resistances with the impedance of inductors and capacitors in formulas 2.24 and 2.19.

## 2.11 Resonance

In RLC circuits Figure 2.2, for instance, electrical resonance happened when the imaginary part of the impedance between the inductor and capacitor cancel out, and the reactance ( $X$ ) of the circuit is equal to zero. The impedance( $Z$ ) of the RLC circuit (Figure 2.2) can be expressed using the formula below.

$$Z = R + (2\pi fL - \frac{1}{2\pi fC})j \quad (2.25)$$

Resonant frequency is the frequency when

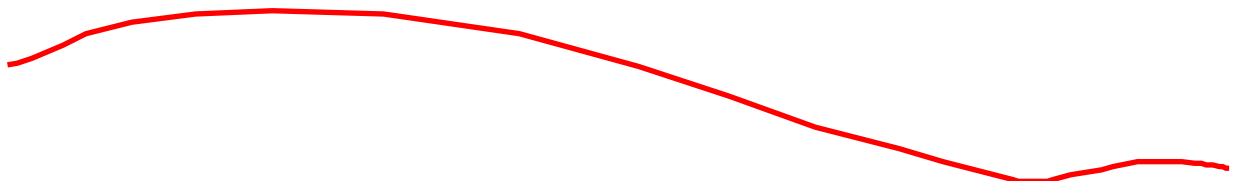
$$|2\pi fL| = |-\frac{1}{2\pi fC}|. \quad (2.26)$$

Resonant frequency can be calculated by solving the Equation 2.26.

$$f_{resonant} = \frac{1}{2\pi\sqrt{LC}} \quad (2.27)$$

Impedance of the resonant circuit now become

$$Z = R + (0)j. \quad (2.28)$$



## Chapter 3

# Statistical Methodologies

In this section, we will discuss statistical methods required for later analysis, including optimization methods to obtain the optimal value we want and evaluation methods to select the most appropriate model balancing goodness of fit and number of parameters.

### 3.1 Non-linear Models

We utilise the nonlinear model to predict the current observation. General relationship between  $Y$  and observation  $X$  in non-linear model can be expressed as a non-linear function of  $X$ ,  $f : R \rightarrow R$ , and error term  $\varepsilon$  [22]. The dependent variable  $Y$  is the output of our model, which is the measured impedance, whereas  $X$  is the measured frequency of the alternating current. Non-linear model is trained using observed frequencies and corresponding impedance. To ensure that the output value of our model is not significantly different from the measured impedance data at a given frequency, we want to minimize the model's sum of squares.

$$Y = f(X; \theta) + \varepsilon_i \quad (3.1)$$

white noise  $\varepsilon_i \in N$  is a collection of independent, identically distributed(iid) random variables  $\varepsilon_1, \varepsilon_2, \varepsilon_3, \dots, \varepsilon_N$ , in which  $N$  represent the number of observations. And  $\varepsilon_i \sim N(0, \sigma^2)$ , where  $\sigma^2 \in R^+$ .  $\theta \in \Theta$  is a  $k$  dimensional parameter (i.e. a  $k \times 1$  vector) with parameter space  $\Theta$ , such that  $\theta = (\theta_1, \dots, \theta_k)^T$ . Letter  $T$  represents the transpose of the vector.





In this research, the three primary model parameters are inductance, resistance, and capacitance. The circuit can let the existence of numerous electronic components to imitate our existing data; therefore, the dimension of our parameter is dependent on the number of electronic components in the modeling circuit. The effect of the parameter on our model is crucial; therefore, optimization is performed to identify the model parameter that best fits.



The most common non-linear model in medical statistics is the cox proportional hazard model, which is used to estimate the probability of future events occurring at a specific time. For instance, it may be used to forecast the instantaneous risk five years later of a sample, such as a 65-year-old Asian male with advanced lung cancer. In this scenario, the relationship between the observed time and the probability of the elderly person passing away is non-linear and can be represented using a non-linear model [6].

## 3.2 Optimisation

Before we look at statistical optimization methodologies we need to know what is optimization and why we should use optimization.

Optimization is the task of finding  $\hat{\theta}$  such that the objective function  $f$  : in terms of k-dimensional parameter  $\theta$  is as large or as small as it can be. The process of locating the maximum of an objective function  $f(\theta)$  does not differ from that of locating its minimum, since

$$\theta^* = \arg \max_{\theta \in \Theta} f(\theta) = \arg \min_{\theta \in \Theta} -f(\theta). \quad (3.2)$$



Numerous facets of academic fields can be improved through optimization. In the financial industry, for example, if I want to develop different products, X and Y, and I know my cost function, I may use optimization to determine the best quantity of X and Y that minimises my cost. Clearly, this is the most elementary example. In the real world, we must additionally consider the available resources, the profitability of the products, etc.

### 3.2.1 Newton method

Newton method is a prevalent optimization technique utilised in numerous fields. Utilizing Taylor expansion and the iterative method to approximate the ideal value of the entire function is its central concept. Newton's approach has several advantages, including the ability to utilise an iterative solution when it is difficult to obtain the root, its speed, and its ability to precisely locate the optimal value.

When we consider one dimension, and we are trying to optimise the twice differentiable function  $f : R \rightarrow R$ . The first step is to carry out the quadratic Taylor polynomial of the function  $f : X$  and supposing they are approximating [11].

$$f(\theta) \approx f(\theta_n) + f'(\theta_n) + \frac{1}{2}f''(\theta_n)^2 \quad (3.3)$$

Consider that  $\theta_n$  in this equation represents a sequence of numbers,  $\{\theta_0, \theta_1, \dots, \theta_n\}$ . If we take the derivative of the Equation 3.3 and suppose that it is equal to zero, we will have the following formula, which is our iteration formula.

$$\theta_{n+1} = \theta_n - \frac{f'(\theta_n)}{f''(\theta_n)^2} \quad (3.4)$$

$\theta_0 \in R$  is the initial guess of the function root. Then, we use the Equation 3.4 as the iteration formula to fill in  $\theta_0$ , and we repeat the iteration process to get  $\theta_1, \theta_2, \theta_3, \dots$  that is closer to optimal value according to degree of convergence.

When applying Newton's method in higher dimensions, our iterative formula can be represented as

$$\Theta_{n+1} = \Theta_n - [Hf(\Theta_n)]^{-1} \nabla f(\Theta_n), \quad (3.5)$$

in which  $\Theta_n = (\theta_1^n, \theta_2^n, \dots, \theta_k^n)^T$ ,  $k$  is the number of dimensions.  $Hf(\Theta_n)$  represent the Hessian matrix of the function  $f(\Theta_n)$ . The formula for iteration can alternatively be stated compactly in matrix form.

$$\begin{bmatrix} \theta_1^{n+1} \\ \theta_2^{n+1} \\ \vdots \\ \theta_k^{n+1} \end{bmatrix} = \begin{bmatrix} \theta_1^n \\ \theta_2^n \\ \vdots \\ \theta_k^n \end{bmatrix} - \begin{bmatrix} \frac{\partial^2 f(\Theta_n)}{\partial \theta_1^2} & \frac{\partial^2 f(\Theta_n)}{\partial \theta_1 \theta_2} & \dots & \frac{\partial^2 f(\Theta_n)}{\partial \theta_1 \theta_k} \\ \frac{\partial^2 f(\Theta_n)}{\partial \theta_2 \theta_1} & \frac{\partial^2 f(\Theta_n)}{\partial \theta_2^2} & \dots & \frac{\partial^2 f(\Theta_n)}{\partial \theta_2 \theta_k} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f(\Theta_n)}{\partial \theta_k \theta_1} & \frac{\partial^2 f(\Theta_n)}{\partial \theta_k \theta_2} & \dots & \frac{\partial^2 f(\Theta_n)}{\partial \theta_k^2} \end{bmatrix}^{-1} \begin{bmatrix} \frac{\partial f(\Theta_n)}{\partial \theta_1} \\ \frac{\partial f(\Theta_n)}{\partial \theta_2} \\ \vdots \\ \frac{\partial f(\Theta_n)}{\partial \theta_k} \end{bmatrix} \quad (3.6)$$

In this instance, the Hessian matrix is a  $k$ -by- $k$  square matrix. When  $k$  is large, it will be costly to calculate the hessian matrix. Due to the complexity caused by the Hessian matrix, the application of the Newton method in high dimensions is no longer appropriate. Quasi-Newton method, on the other hand, is an alternative optimization approach based on Newton's method that solves the problem by not requiring the actual computation of the Hessian matrix. Rather, the gradient vector is utilised to continuously update an approximation of the Hessian matrix.

### 3.3 Broyden–Fletcher–Goldfarb–Shanno Method

The Broyden–Fletcher–Goldfarb–Shanno (BFGS) method **is the most used Quasi-Newton algorithm**; it is named after its discoverers Broyden, Fletcher, Goldfarb, and Shanno [37]. It is an iterative method for solving unconstrained nonlinear optimization problems in numerical optimization [14].



Suppose we are attempting to optimise differentiable function  $f(\Theta)$  in dimension  $k$ , where  $\Theta$  at  $n$  iterations can be expressed as  $\Theta_n = (\theta_1^n, \theta_2^n, \dots, \theta_k^n)^T$ . In later formulations, the derivative of the function at  $\Theta_n$  will be represented by  $\nabla f(\Theta_n)$  and  $f'(\Theta_n)$ .

In the BFGS approach, the aim of our iterative formula is to continuously update  $B_{\Theta_n}$  which is a  $k \times k$  symmetric matrix that approximates the Hessian matrix. Before taking the iteration, we need to confirm the search direction  $p_{\Theta}$ .

$$p_{\Theta_n} = -B_{\Theta_n}^{-1} \nabla f(\Theta_n), \quad (3.7)$$

When we define the direction of our search, we can find a point along that direction,

so our iteration formula can be defined as

$$\theta_{n+1} = \theta_n + \alpha_n(-H_n f'(n)). \quad (3.8)$$

In the iteration formula,  $\alpha_n$  is referred to as step length, and we seek a  $\alpha_n$  that

$$\min_{\alpha} (f(\theta_n + \alpha p_{\Theta_n})) \quad (3.9)$$

Wolfe conditions will be utilised as a standard to determine a suitable alpha value and its definition contains three inequalities which are [31]

$$f(\theta_n + \alpha_n p_n) \leq f(\theta_n) + \tilde{\beta} \alpha_n \nabla f(\Theta_n) p_n^T, \quad (3.10)$$

$$\nabla f(\theta_n + \alpha_n p_n) p_n^T \geq \beta \alpha_n \nabla f(\Theta_n) p_n^T, \quad (3.11)$$

$$0 \leq \tilde{\beta} \leq \beta \leq 1. \quad (3.12)$$

Using the same procedure as Newton method, assume the second-order Taylor polynomial obtained by Taylor expansion is approximating to our objective function and we will get

$$f(\theta) \approx f(\theta_{n+1}) + f'(\theta_{n+1})(\theta - \theta_{n+1}) + \frac{1}{2} f''(\theta_{n+1})(\theta - \theta_{n+1}). \quad (3.13)$$



When we differentiate this expression, we get

$$f'(\theta) = \nabla f(\theta_{n+1}) + H_{n+1}(\theta - \theta_{n+1}), \quad (3.14)$$

where  $H_{n+1}$  represents the Hessian matrix, and since we assume that the previously defined matrix  $B$  approximates the genuine Hessian matrix  $H$ , therefore  $H_{n+1} = B_{n+1}$ . And we can rewrite the Equation 3.14 and obtain

$$B_{n+1} s_n = y_n. \quad (3.15)$$

with

$$s_n = (\theta_{n+1} - \theta_n) \quad (3.16)$$

$$y_n = (\nabla f(n+1) - \nabla f(n)) \quad (3.17)$$

According to the addition rule for matrices, we suppose that  $B_{n+1}$  matrix is equal to  $B_n$  plus an unknown  $k \times k$  matrix at  $n$  iterations which is defined as  $\varpi_n = \alpha z_n z_n^T + \beta q_n q_n^T$ . And  $z_n, q_n$  are  $k \times 1$  matrix. (I won't go into great explanation as to why  $\varpi_n = \alpha z_n z_n^T + \beta q_n q_n^T$ , instead relying on the result found by mathematicians Broyden, Fletcher, Goldfarb, and Shanno. Their contribution to the non-linear optimisation is invaluable and pioneering.)

$$B_{n+1} = B_n + \varpi_n \quad (3.18)$$

When  $z_n = y_n$  and  $q_n = s_n$  we get the formula [15]

$$(B_n + \alpha y_n y_n^T + \beta s_n s_n^T) s_n^T = y_n \quad (3.19)$$

After solving the equation we get the expression for  $\alpha$  and  $\beta$ .

$$\begin{aligned} \alpha &= \frac{1}{y_n^T s_n} \\ \beta &= -\frac{1}{s_n^T y_n s_n} \end{aligned} \quad (3.20)$$

Finally, we can get the iteration formula

$$B_{n+1} = B_n - \frac{B_n s_n s_n^T B_n}{s_n^T B_n s_n} + \frac{y_n y_n^T}{y_n^T s_n} \quad (3.21)$$

We use the result obtained here for  $k+1$  as the input in Equation 3.8 and thus continue iterations until  $X_n$  is converging to a specific number.

Finite difference approximation is utilised to approximate the derivative of the objective function at point  $\theta$  and it is defined as

$$f(\theta) \approx \frac{f(\theta - \nabla \theta) - f(\theta)}{\nabla \theta} + \varepsilon. \quad (3.22)$$

The approximation of the objective function's derivative will improve for smaller  $\nabla\theta$ . However, this is not always the case, and when  $\nabla\theta$  becomes extremely small, we must consider the floating point ~~of the error~~. When  $\nabla\theta$  is equal to  $1 \times 10^{-20}$ , which is a very small value, we would expect to obtain a highly accurate approximation of the derivative. However, when we calculate using computer programming software, the error of the result is greater since  $\theta$  surpasses the upper limit of significant numbers that can be recognised by software, so it is calculated as 0 in the process of calculation. Therefore, it generates an abrupt increase in errors.

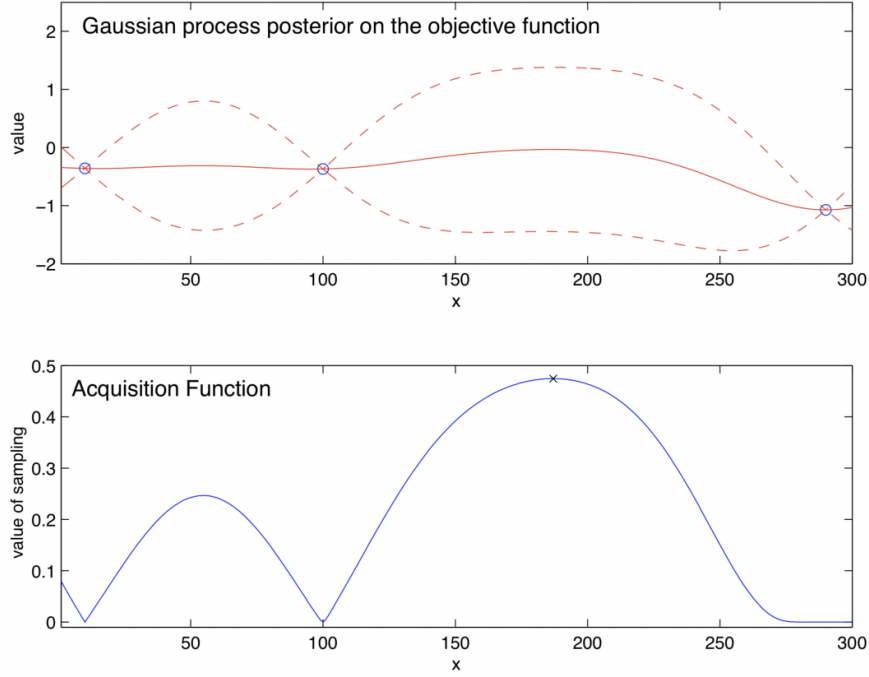
Equation 3.22 represent the finite difference approximation with an additional error term  $\varepsilon$ , due to the presence of the error term, we may anticipate that the objective function  $f(x)$  will not only be twice differentiable, but also higher-order differentiable. BFGS can be utilised for large dimensions, but the objective has always been to locate the local maximum after setting the step length and search direction; this is the restriction of BFGS method. Consequently, the BFGS approach can only handle the converging objective function. Since the purpose of optimisation is to discover the global maximum of the objective function, it is inappropriate and costly to continue using the BFGS approach when the objective function has many local maximum. The subsequent optimization technique is known as Bayesian Optimisation, which is not dependent on Newton's method and seeks global optimal values.

### 3.4 Bayesian optimization

Bayesian optimization is a technique for optimising lengthy-to-evaluate objective functions. Bayesian optimization functions in higher dimensions, tolerates stochastic noise in function evaluations, and is most effective for optimization over continuous domains with fewer than 20 dimensions [17].

Assume we are now attempting to optimise  $K$  dimensional  $f : \Theta \rightarrow R$ ,  $\Theta \in (\theta_1, \theta_2, \dots, \theta_k)$ , via Bayesian optimisation. Red line shown in Figure 3.1 reflects our objective function. Initially, we do not know what our objective image would look like, thus we substitute the surrogate function to represent the behaviour of target

function. The surrogate function achieves this by collecting  $n$  limited points at random and then fitting through them [35]. we can combine all the object<sup>ion</sup> function's output  $f(\Theta)$  into a vector form, for each limited random point  $\Theta = (\theta_1, \theta_2, \dots, \theta_n)$



**Figure 3.1:** Bayesian Optimisation[17]

Because of its versatility, the Gaussian process prior is selected as the surrogate function and function's prior. Therefore, the distribution of one object<sup>ion</sup> function's value  $f(\theta_n)$  is expressed as

$$f(\theta_n) \sim N(\mu_0(\theta_n), \sigma^2(\theta_n)). \quad (3.23)$$

Then, we must identify an Acquisition function, whose objective is to assist us based on our previous function distribution(Gaussian process) and our current observation to obtain posterior distribution according to Bayesian posterior updating process [44].

Gaussian process regression is used to produce Bayesian credible intervals which measures the uncertainty around the estimated parameter values. It is shown by the dotted red line in Figure 3.1. Because we simulate our goal function using

a Gaussian process regression, our Bayesian credible intervals may be expressed as the mean and standard deviation of the Gaussian process regression. The Upper Confidence Bound (UCB) shows the magnitude of our forecast as well as the confidence interval. UCB can be expressed as

$$UCB = \mu(\theta) \pm k\sigma(\theta), \quad (3.24)$$

in which  $k$  represent the  $Z_{0.975} = 1.96$ .

Bayesian optimization is searching global optimum, hence the acquisition function cannot solely evaluate the present lowest or maximum value, but must also investigate unexplored regions where global optimum may exist. These characteristics is typically known as Exploitation and Exploration.

Acquisition function is denoted by

$$u : \Theta \rightarrow R^+ \quad (3.25)$$

$$u(\Theta) = u(\Theta | \{\theta_n, y_n\}) \quad (3.26)$$

$$y_n \sim N(f(\theta_n), \sigma^2) \quad (3.27)$$

$\sigma^2$  represent the variance of the noise.  $\theta_n, y_n$  represent the previous observation.

If the findings of the previous iterations were too distinct from the value of objective function, then this step should be skipped and not investigated. Alternatively, if a **significant** value is noticed based on historical data and it may be the global optimum, we would deploy additional points around this point **to mimic the most real objective function.**

By analyzing the previous data, we can make new judgments about optimal value. The posterior distribution after given the observation now follow the normal distribution with mean  $\mu_n(\theta)$  and variance  $\sigma_n^2(\theta)$  [44].

$$f(\theta) | f(\theta_{1:n}) \sim N(\mu_n(x), \sigma_n^2(\theta)) \quad (3.28)$$



According to the Gaussian process regression, the distribution of our function value is now known, allowing us to forecast the predicted function output in any unknown location.

The most common and most practical acquisition function is EI (Expected improvement) and it can be evaluated by the following equation.

$$EI_{\theta}(\theta) = \sigma_n(\theta) \phi\left(\frac{\mu_n(\theta) - f(\theta^*)}{\sigma_n(\theta)}\right) - (\mu_n(\theta) - f(\theta^*) - \beta) \Phi\left(\frac{\mu_n(\theta) - f(\theta^*)}{\sigma_n(\theta)}\right) \quad (3.29)$$

$f(\theta^*)$  is formerly observed greatest value and  $\theta^*$  is the corresponds x-point. It can be expressed as  $\theta^* = \arg \max f(\theta)$ , where n is the number of times the function has been evaluated.  $\Phi(x)$  is the standard normal p.d.f., and  $\phi(x)$  is the cumulative distribution function of the standard Gaussian process.  $\beta$  represent the learning-rate of the function which is used to balance between Exploration and Exploitation [48].

The expected improvement algorithm then find the point with the greatest expected improvement [17].



$$X_{NextEvaluation} = \arg \max EI_{\theta}(\theta) \quad (3.30)$$

The Bayesian Optimisation approach is employed before the BFGS. Bayesian Optimization can therefore assist in locating the global maximum for the entire likelihood function. Then, BFGS methods are employed to more precisely find our maximum Likelihood estimates based on the results of Bayesian optimisation. This sequence makes up for the fact that our BFGS algorithm finds the local maximum as opposed to the global maximum.

In this research, I utilised the R package [50] in performing Bayesian Optimisation of my objective function [50].

## 3.5 Evaluation Methods

### 3.5.1 Log-likelihood function

Suppose we have a random sample with n observations  $X = (x_1, x_2, x_3, \dots, x_n)$  with probability density(mass) function  $f_{X_i}(x_i; \theta)$  with parameter(s)  $\theta$ . The likelihood

function for  $\theta$  is defined as [32]

$$L(\theta; x) = \prod_{i=1}^n f_{X_i}(x_i; \theta). \quad (3.31)$$

The likelihood function is just the joint probability density(mass) function of  $X$  and the the log-likelihood function for  $\theta$  is defined as [2]

$$\ell(\theta; x) = \log L(\theta; x). \quad (3.32)$$

Maximum likelihood estimate (MLE) was first proposed in 1920 by R.A. Fisher. In the previous definition, we assumed that log-likelihood is actually a probability density (mass) function, therefore when we attempt to make inferences about parameter theta, it is natural for us to select the data's highest probability of occurrence, hence the concept of maximum likelihood estimation [36].

Maximum likelihood estimate  $\hat{\theta}$  refers to the parameter that maximise the likelihood function. The calculation process is same as the optimisation problem shown in Equation 3.2 after substituting the log-likelihood function (L) to objective function  $f(\theta)$ .

### 3.5.2 Akaike Information Criterion

Considering that we have some data, our goal is to determine the optimal model for predicting these variables in order to reduce information loss. The information criteria used to quantify the performance of the candidate model in this project is Akaike information criterion, abbreviated as AIC, which is typically the initial model selection criterion that should be applied in practice[16]. AIC provide a surprisingly straightforward estimate of the typical out-of-sample deviation [32]. Out-of-sample deviation refers to past data that was not utilised for training the model. As a general rule, 75% of the data should be used for fitting and the remaining 25% for validation. When I am entirely satisfied with the fitted model, I use it to forecast the remainder of the data and assess its performance. Therefore, AIC evaluates the prediction quality of our model based on historical data.

AIC will seek the lowest negative likelihood of the models and choose the

candidate model with the best performance[47], taking into account the number of parameters. The relationship between them can be presented as [2]

$$AIC = -2\text{Log}(L) + 2p, \quad (3.33)$$

Where L is the Likelihood of the model and p is the number of the parameters in the model.

AIC was used in this project to evaluate the predictive impedance performance of the model, which was then utilised for model selection. When comparing two nonlinear models, the model with the smaller AIC value is always preferred.

The Bayesian information criterion (BIC) is a model selection method that differs from the Akaike information criterion (AIC) by the second term.

$$BIC = -2\text{Log}(L) + p\text{Log}(n) \quad (3.34)$$

letter n in the above formula represented the number of data sets.

We may discover that the Bayesian information criterion (BIC) quite similar to the Akaike information criterion (AIC), as both are based on the likelihood function. Hence, the quantity of AIC and BIC has close to each other and have little influence on the model selection. I ultimately chose AIC for two reasons. The first reason is that AIC and BIC have distinct implications. When using AIC, we believe that the candidate is not the genuine model. But I'd like to utilise the best "wrong" model for predicting the inductor impedance. However, the BIC technique suggests that the true model is contained within the candidate model, and I'll eventually find it after collecting more samples[3]. AIC is preferable due to the complexities of modelling impedance that caused by non-ideal behaviour of the inductor. Another reason is that the AIC has a straightforward estimation of the out-of-sample deviation. When just one data set is available, for instance, we can randomly divide the raw data set into training data and test data before fitting the model to the training data. Compare model predictions with observed values using test data. Repeat many times. Using the average prediction error, models are then compared. Due to Monte



Carlo error, comparisons conducted only a few times are unreliable[18]. We cannot afford several training data fits. Hence, it is more effective to use out-of-sample bias; therefore, I chose AIC as the model selection approach.

2778

## Chapter 4

# Data&Research Questions

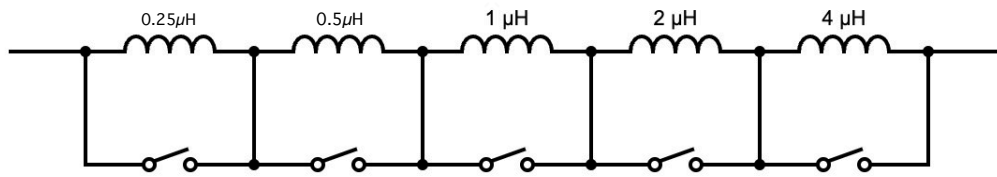
The data consist of measurements of the impedance of 31 different coils across a range of frequencies from 1.8 MHz to 30 MHz. The measured impedance are expressed as reflection coefficients relative to a system impedance of 50 Ohms. The link between reflection coefficients and impedance is given via (4.1).

$$\Gamma = \frac{Z_L - Z_0}{Z_L + Z_0} \quad (4.1)$$

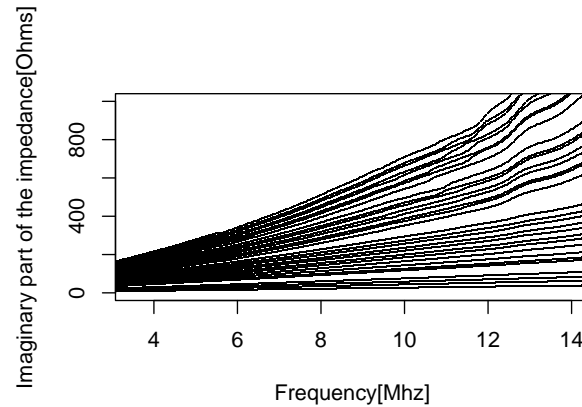
In this formula,  $Z_L$  represents the complex load impedance, whereas  $Z_0$  represents the characteristic impedance, which defaults to being ~~nearly~~ purely real (50 ohm) for subsequent operations in order to obtain the load impedance in the electric circuit.

The 31 inductances were actually obtained as serial circuits of five fixed inductors as displayed in Figure 4.1. Since each inductor is connected in parallel to a switch, the switch can control each inductor, allowing us to create an electric circuit with different inductors. Moreover, the impedance varies for each inductor in the circuit, ranging from low to high from  $0.25\mu H$  to  $4\mu H$  from the first inductor to the fifth inductor. Inductors with inductances ranging from  $0.25\mu H$  to  $2\mu H$  are cylindrical air-cored inductor(Shown in 2.71). However, they are distinct from the fifth inductor with inductance of  $4\mu H$ , toroidal inductor(Shown in 2.72) with iron powder core.

Observing the differences between the imaginary part of the impedance and frequency of 31 sets of data (Figure4.2) reveals that the impedance of 31 sets of data



**Figure 4.1:** This graph illustrates the circuit diagram of five inductors connected in series.



**Figure 4.2:** Imaginary part of the Impedance against Frequency graph

increases as the AC frequency rises. Because the inductive reactance is proportional to the frequency of alternating current (Shown in Equation 2.15). Considering that only inductors are connected in series, the total impedance of the circuit is equal to the sum of the impedance of each inductor. When using a circuit with a relatively higher impedance, the rate of increase of impedance is greater than a circuit with a lower impedance.

## 4.1 Research Questions

Our objective is to develop the most precise models to simulate impedance( $X$ ) at various frequencies( $\nu$ ). The object we modelled is the inductors shown in Figure 4.1.

This statistical model resembles an unidentified electrical circuit. The purpose of statistical analysis is to compare the model's impedance to the actual measured value. The comparison results would aid me in optimising my initial model and determining the optimal model by comparing various models using statistical method-

ologies.

387

## Chapter 5

# Results

Figure 2.2 depicts the simple electric circuit (RLC circuit) that served as the starting point of our modelling.

Chapter 2.9 introduces the formula and content that express the impedance of inductor and capacitor respectively. Combining this with the series and parallel calculation methods described in Chapter 2.10 allows for the expression of the impedance of the RLC model to be derived. The Expression below is my initial nonlinear-model consists of inductance, capacitance, and resistance.

Expression for the impedance of the RLC circuit is described below:

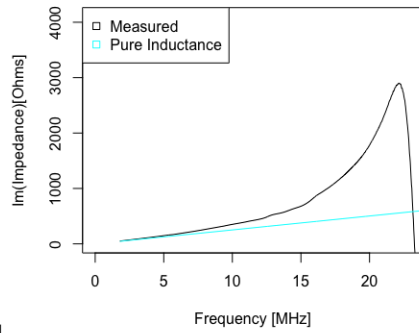
$$Z = (\frac{1}{R + 2\pi f L j} + 2\pi f C j)^{-1}. \quad (5.1)$$

I will start in simulating the corresponding data set for the fifth inductor( $4\mu\text{H}$ ) demonstrated in Figure 4.1. In order to evaluate the predictive performance of the candidate model, I will compare the model's prediction with the measured impedance as well as its information criterion.

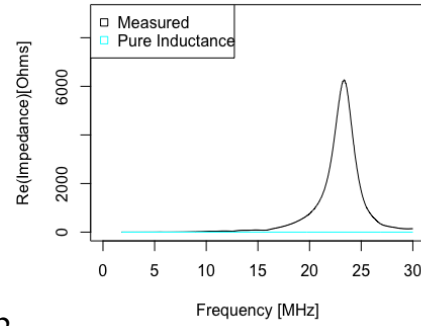
### 5.1 Simulation of 4 Microhenry Toroidal Inductor

Figure 5.1 depicts the imaginary and real part of the Impedance versus Frequency graph. The black curve can be seen represents the impedance measured in practise. I conduct exploratory data analysis, as we initially do not know which parameters for our initial model, the RLC model, are appropriate. The blue line represents the simplest conceivable model, i.e. The RLC model that consist of ideal  $4\mu\text{H}$  inductor





5.11



5.12

**Figure 5.1:** The measured impedance of  $4\mu\text{H}$  datasets is compared to the impedance predicted by conceivable model which has only a ideal  $4\mu\text{H}$  inductor

only. Other parameters such as capacitance and resistance are therefore null. There will be a sequence of models of increasing complexity and better fit to the data. The model with inductance only is at the extremely simple end of this sequence.

Observing the blue line in the image of imaginary part of the impedance against frequency graph in this instance reveals that our model fits well when the frequency is low (Around 2Mhz). Nonetheless, as frequency increases, the discrepancy between model's prediction and actual measurement results grows. It is obvious that this model is inappropriate since it is not simulating the data set at all at higher frequency.

However, the blue line depicts in the Figure 5.11 is the tangent to the black curve (measured impedance) at about 4 Mhz, indicating that selecting the inductance of  $4\mu\text{H}$  of the inductor to simulate the  $4\mu\text{H}$  data set is an good decision.

One possible explanation for the inaccuracy of model predictions is that, by default, only inductance is available for our circuit, so the predicting power for capacitor and resistor are lost. Figure 5.12 depicts the real part of impedance versus frequency. We can see that current model is incapable of predicting the actual resistance (Real part of the impedance). The RLC model with only the inductance and other parameters set to zero, such as capacitance and resistance, does not appear to be ideal for model fitting. This is an example of erroneous parameters, which will diminish our model's capacity for precise prediction. In order to find the optimal parameter for our model, we endeavour to optimise these three parameters which

can be determined by finding the maximum log-likelihood estimate.

Loss represents the difference between our model's output and actual measurement findings. If the predictive capacity of model is very accurate, then loss will be reduced and close to 0, however if the model's fit is not ideal, then the quantity of loss will be large.

Given the current model, the-log likelihood function in the normalized model can be interpreted as the probability that the output of our model closely resembles the measured data. To achieve a reliable model, we must minimise the loss, which is the same concept as minimising the negative log likelihood function. Thus, we assume that the loss function is equivalent to the negative log likelihood function up to a multiplicative parameter that is inverse the error variance. The expression for loss is shown below:



$$\sum_{\nu} (X_{mes}(\nu) - X(\nu, \theta))^2 = loss = -\ell, \quad (5.2)$$

in which predicted impedance  $X(\nu, \theta)$  can be viewed as a function with respect to frequency parameter  $\nu$  and vector  $\theta \in [0, \infty]$ . And  $X_{mes}(\nu)$  represent the measured Impedance at give frequency  $\nu$ .

Considering the Equation 5.1, the predicted impedance of RLC model can be expressed in terms of frequency  $\nu$  and vector  $\theta = [\theta_1, \theta_2, \theta_3]$ , in which  $\theta_1, \theta_2, \theta_3$  represent Inductance, Capacitance, and Resistance respectively.  $X_{RLC}(\nu, \theta)$  can therefore defined as follows:

$$X_{RLC}(\nu, \theta) = \left( \frac{1}{\theta_3 + 2\pi\nu\theta_1j} + 2\pi\nu\theta_2j \right)^{-1} \quad (5.3)$$

The maximum log-likelihood estimate under normalised model is

$$\hat{\theta} = \arg \min_{\theta} (-\ell(\theta)) = \arg \max_{\theta} \ell(\theta). \quad (5.4)$$

Model error ( $\epsilon$ ) is the stochastic error that model the difference between pre-

dicted impedance and the actual measured value.

$$X_{mes}(\mathbf{v}) = X(\mathbf{v}, \boldsymbol{\theta}) + \boldsymbol{\varepsilon} \quad (5.5)$$

Impedance is a complex number consisting of both a real and an imaginary component. When modelling the residual, it is natural for us to consider residual as the random vector and the imaginary and real parts as two dependent normal random variables. Therefore, a bivariate normal distribution might be a appropriate choice to model the probability distribution of the error. And it can be defined as:

$$\begin{pmatrix} \varepsilon_{Re} \\ \varepsilon_{Im} \end{pmatrix} = \boldsymbol{\varepsilon} \sim N(0, \Sigma), \Sigma = \begin{bmatrix} \sigma_r^2 & \sigma_r \sigma_i \\ \sigma_r \sigma_i & \sigma_i^2 \end{bmatrix} \quad (5.6)$$

We can now compute the joint probability density function of  $\varepsilon_{Re}$  and  $\varepsilon_{Im}$ . It can be written as

$$f_X(x; \boldsymbol{\theta}) = \prod_{i=1}^N \frac{1}{(\sqrt{\det(2\pi\Sigma)})} \exp\left\{-\frac{1}{2}(X_{mes}(\mathbf{v}) - X(\mathbf{v}, \boldsymbol{\theta}))^T \Sigma^{-1} (X_{mes}(\mathbf{v}) - X(\mathbf{v}, \boldsymbol{\theta}))\right\}. \quad (5.7)$$

The quantity in  $\{\cdot\}$  is a quadratic form. The measured impedance and modeled impedance are in the vector form written as:

$$X_{mes} = \begin{bmatrix} Re(X_{mes}) \\ Im(X_{mes}) \end{bmatrix}, X(\mathbf{v}; \boldsymbol{\theta}) = \begin{bmatrix} Re(X(\mathbf{v}; \boldsymbol{\theta})) \\ Im(X(\mathbf{v}; \boldsymbol{\theta})) \end{bmatrix}. \quad (5.8)$$

According to the definition of the log-likelihood function expressed in Equation 3.32, we can then take a logarithm of the joint probability distribution and compute the log-likelihood function in terms of the covariance matrix.

$$\ell(x; \boldsymbol{\theta}) = -\frac{1}{2} \sum_1^n [(X_{mes}(\mathbf{v}) - X(\mathbf{v}, \boldsymbol{\theta}))^T \Sigma^{-1} (X_{mes}(\mathbf{v}) - X(\mathbf{v}, \boldsymbol{\theta})) + \log(2\pi\Sigma)] \quad (5.9)$$

In order to determine the maximum log-likelihood estimates, we will now employ the optimization methodologies introduced in Chapter 2 to the log-likelihood function of RLC model. As an illustration, the Best estimate obtained by using



different optimisation methods is shown below:

Simple Inductance Model	Inductance[H]	Capacitance[F]	Resistance[Ω]	Log-likelihood
Initial Parameters	4	0	0	-17522.83
Parameter after BaysOpt	4.0157	11.5308	55.9391	-11796.56
Parameter after BaysOpt&BFGS	3.9801	11.6554	53.7760	-11709.51

**Table 5.1:** A comparison of the quantity of the log-likelihood resulting from NO statistical optimization, the BFGS method alone, and combination of BFGS and Bayesian Optimization together.

Observing the above table visually reveals the trend of improvement of the predictive performance of the RLC model that using different sequence of optimisation method. Initial log-likelihood has been significantly improved from -17522 to -11709. Model optimised by the BFGS method after Bayesian optimisation is preferable to models with no optimisation or Bayesian optimisation alone. Because its log-likelihood is less negative than others.

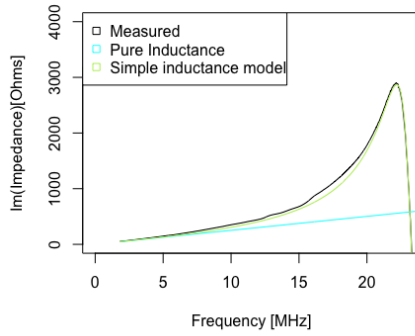
Table demonstrated below represent the iterations and the updating procedure accompanied with the quantity of log-likelihood of Bayesian Optimisation.

Iterations of BaysOpt	Inductance	Capacitance	Resistance	Log-likelihood
1st sampling point	4.551552	28.75003	38.48987	-17733.31
2nd sampling point	5.065576	17.28041	99.26881	-17704.78
3rd sampling point	3.186615	29.44952	43.4751	-17610.38
4th sampling point	3.877136	24.72166	29.4267	-17819.73
5th sampling point	4.760303	35.87559	73.30402	-17587.10
1st iteration	4.162245	10.93917	64.24394	-13736.15
2nd iteration	4.073188	10.47663	61.77328	-16459.45
3rd iteration	4.696899	11.28179	64.37451	-17111.47
4th iteration	4.155704	11.11235	56.73725	-12582.85
5th iteration	4.203839	11.3686	57.3363	-13591.75
⋮	⋮	⋮	⋮	⋮

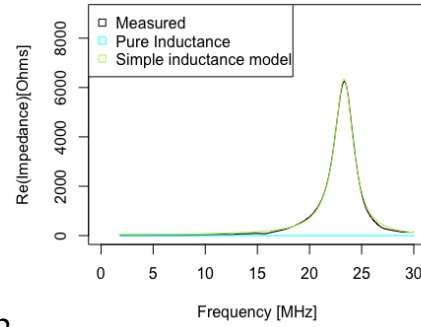
**Table 5.2:** Quantity of log-likelihoods and parameters for each iterations and sampling points of Bayesian optimisation

The first five rows of the table indicate five points on the log-likelihood function that were randomly selected by the surrogate function. The remaining rows of the table indicate the subsequent iteration and updating process obtained by learning the prior randomly sampled points. I have only displayed the first five iterations

of Bayesian Optimisation; BayesOpt would require at least 50 iterations to get a reliable estimate. The iteration with the largest value of log-likelihood is already displayed in Table 5.1.



6.11



6.12

**Figure 5.2:** The measured impedance of 4  $\mu\text{H}$  datasets is compared to the impedance predicted by model 1(RLC Model) after statistical optimization.

The green lines in Figure 6.1 indicate the model's forecast after the optimisation methods was applied. It was evident that the simulation conditions resulting from the optimisation procedure for both the imaginary and real parts of the impedance were vastly superior to those of the original model (Inductance model shown in blue). Our model's forecast nearly matched the measured impedance in the real electric circuit. However, there are still modest deviations from the prediction for frequencies 15–20 MHz and 25–30 MHz for the imaginary and real parts of the impedance, respectively. Y-axis has a large scale, so even a small deviation may cause significant issues.

Looking at Figure 6.12, we can see that at the low frequency range, the real part of the predicted impedance resemble a straight line. This is because the inductor acts as a short circuit in the RLC model (Figure 2.5) , and the capacitor acts as an open circuit (Since an ideal capacitor has infinite resistance, so electricity will not pass through.), leaving only the resistor in the circuit.

In order to have a deeper understanding of the performance of our model, we must consider its residuals. Residuals are utilised in numerous processes aiming to detect various forms of model-data discrepancies. In conventional usage, we also use the term "error". Chesher and Irish's work from the early 1980s forms the

foundation for a significant proportion of the prevalent residual analysis techniques [9].

The Affine relationship between the index and the frequency  $f_i$  can be expressed as:

$$f_i = i \times \Delta f + f_0. \quad (5.10)$$

$f_0 = 1.8$  MHz is the frequency at which the measured data begins.  $\Delta f$  is a constant value that equal to 0.028 MHz. Index  $i \in \{1, 2, 3, \dots, n\}$ , in which  $n$  denotes the number of measurements available for each inductor.

Residuals are the difference between the predicted and measured data and are defined by the following formula, where  $Y_f$  and  $y_f$  are  $n \times 1$  vectors representing the fitted value and measured data at frequency  $f_i$ , respectively [5].

$$Y_f = \begin{bmatrix} Y_{f_0} \\ Y_{f_1} \\ \vdots \\ Y_{f_n} \end{bmatrix}, y_f = \begin{bmatrix} y_{f_0} \\ y_{f_1} \\ \vdots \\ y_{f_n} \end{bmatrix} \quad (5.11)$$

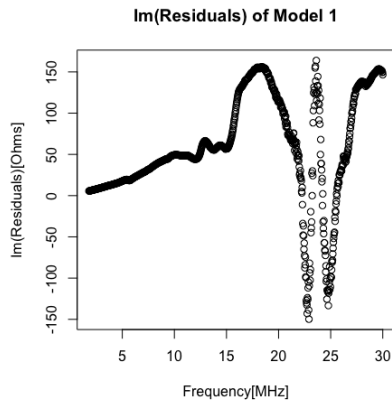
Therefore, the residuals can represented as:

$$Z_f = y_f - Y_f. \quad (5.12)$$

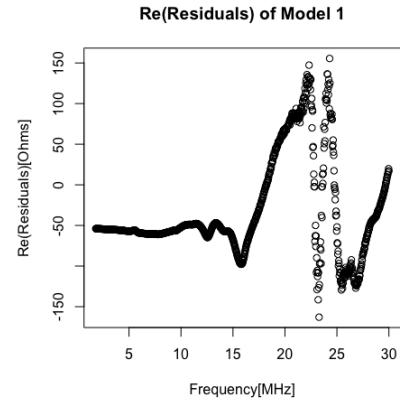
We can also express this more compactly using matrix notation:

$$\begin{bmatrix} Z_{f_0} \\ Z_{f_1} \\ \vdots \\ Z_{f_n} \end{bmatrix} = \begin{bmatrix} y_{f_0} \\ y_{f_1} \\ \vdots \\ y_{f_n} \end{bmatrix} - \begin{bmatrix} Y_{f_0} \\ Y_{f_1} \\ \vdots \\ Y_{f_n} \end{bmatrix} \quad (5.13)$$

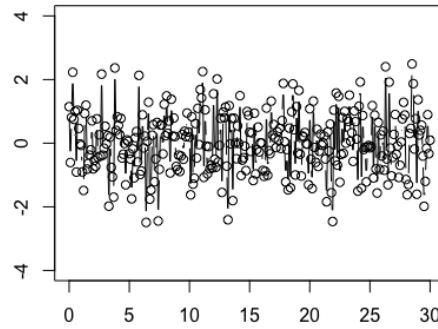
On the basis of an idealised theory, we believe that our model error follows a normal distribution with mean zero and constant variance, the Residuals against Frequency plots should vary around 0 and not exhibit a regular pattern, akin to random noise shown in Figure 5.5.



**Figure 5.3:** Imaginary part of the residuals of Model 1 testing the  $4\mu\text{H}$  data sets



**Figure 5.4:** Real part of the residuals of Model 1 testing the  $4\mu\text{H}$  data sets



**Figure 5.5:** Expected plot of the Residual against Frequency

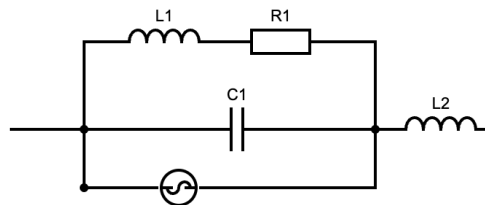
The two pictures of Figure 5.3 and 5.4 depict, respectively, the imaginary and real parts of residuals against with frequency. The preceding images Figure 6.1 led us to conclude that the simulation power of our model after statistical optimization was favourable. However, the result depicted in Figure 5.3 and 5.4 is not **optimistic**. Both imaginary and real part of the residuals are structured and depicts a clear increasing tendency. A significant peak can be observed at the frequency of 17 Mhz and 24 Mhz.

Moreover, the variance of the residual is not constant over time, and variance occurs in clusters, i.e., there are periods of extreme volatility, such as at the reso-

nance frequency of approximately 23 MHz. There are also periods of low volatility between 5Mhz and 15Mhz. Moreover, Figure 5.3 differs significantly from the expected Residual distribution (Figure 5.3 ). The study of the evident upward trend suggests that the impedance of our current model is significantly lower than the measured impedance, and therefore the current model's predictive performance is insufficient.

Likewise, the real part of the Residual against Frequency plot ( Figure 5.4) is structured and therefore **pessimistic**. However, by observing the imperfect residual image, I can attempt to improve my current model by incorporating additional electronic components. When frequency is around 1 MHz, the residual is roughly -50 ohms, and it remains at this value until frequency is increased to 15 MHz. Such an image characteristic indicates that the total resistance in simple electric circuit is greater than the measured data, causing the curve to oscillate at -50 ohms. Upon inspection of the residual plots, it is evident that the residual does not follow the normal distribution.

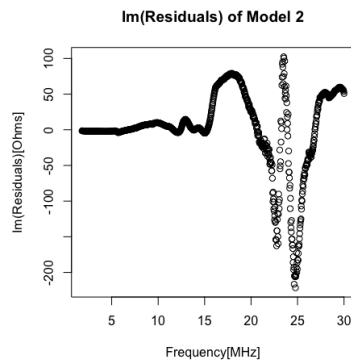
After examining the imaginary part of the residuals of the first model, in order to combat the rising trend shown in Figure 5.3, we can attempt to increase the impedance of the electric circuit by adding a series-connected inductor. The modifications made to model 1 result in the formation of a new models, model 2 which is given in Figure 5.8.



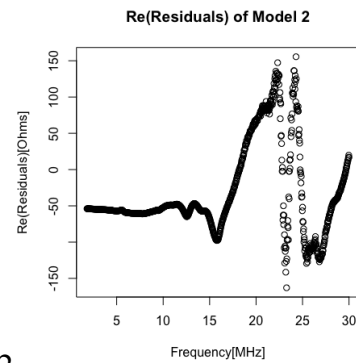
**Figure 5.6:** Model 2

On the basis of model 1(Figure 2.2), model 2 adds a series-connected inductor, which only affects the imaginary part of the impedance. Hence, we anticipate no change in the real part of the residual plot of model 2. Furthermore, as demonstrated in Figure 5.7, Model 2 does resolve the rising trend in Model 1's image, indicating





5.71



5.72

**Figure 5.7:** Residual against Frequency graph of Model 2 testing the  $4\mu\text{H}$  data sets

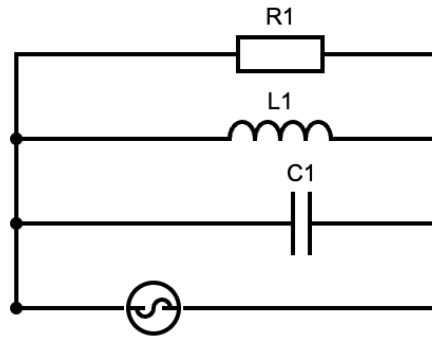
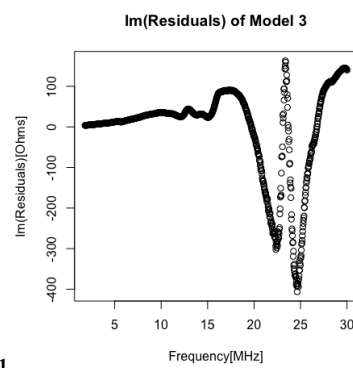
that our model fitting has been improved. Yet, the general appearance of the residual plot of model 2 is still significantly different from what we anticipated (Figure 5.5), and the residual distribution is definitely not normal.

Magnetic saturation, which is the collapse of the magnetic field, thermal phenomena, and inductor resonance are three explanations for **inductor's** poor predictive ability. We are using iron power toroidal inductor with low signal which is unlikely to reach magnetic saturation. Blanchard [7] and Detka [12] explain the underlying physics behind the difficulty of simulating a resonant inductor, and why thermal phenomena must be taken into consideration with toroidal inductors in particular. These demand a deeper understanding of physics, and fitting models like this goes slightly beyond the scope of the research.

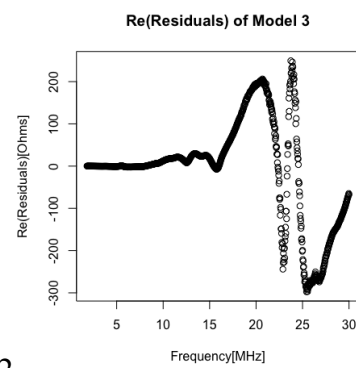


Model 1 and Model 2, their imaginary part of the residual plot are both not satisfying, since they are oscillating around -50 Ohms and we expect they are oscillating around 0. On the basis of model 1, model 3 (shown in Figure 5.8) is presented, which, by replacing the series connection between the resistor and inductor with a parallel connection, may possibly alleviate the problem of the current model's excessive resistance.

Compare the Residual plot of model 3 (Figure 5.9) with model 1 (Figure 5.3 and Figure 5.4). Consider initially the imaginary part of residuals. It is worth noting that, Model 1 has a maximum absolute residual of approximately 100 at its resonant frequency, whereas model 3 reaches a shocking 450. This indicates that model 3's ability to forecast at resonant frequency is inferior to model 1's. In addition, When

**Figure 5.8:** Model 3

5.91



5.92

**Figure 5.9:** Residual against Frequency graph of Model 3 testing the  $4\mu\text{H}$  data sets

we consider the real part of residual. Model 3's maximum absolute residual at the resonant frequency is 300 ohms, which is roughly double that of Model 1 (150 Ohms). Consequently, in this example based on the image, model 3 is inferior to model 1.

After optimising the parameters of each model and computing the log-likelihood, the AIC of each model can be derived using the Chapter 2 definition of AIC (Equation 3.33).

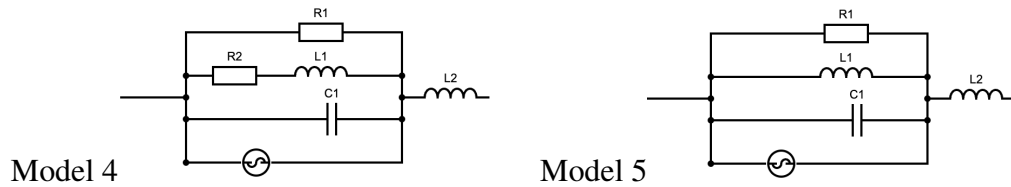
	Model 1	Model 2	Model 3
Akaike Information Criterion	23561.17	22213.40	24695.37

**Table 5.3**

In our context, the number of Model's parameters is not excessively high which is approximately around 5. Hence, the error of the model, as indicated by the model's likelihood function, would be the primary factor that influence AIC. So,

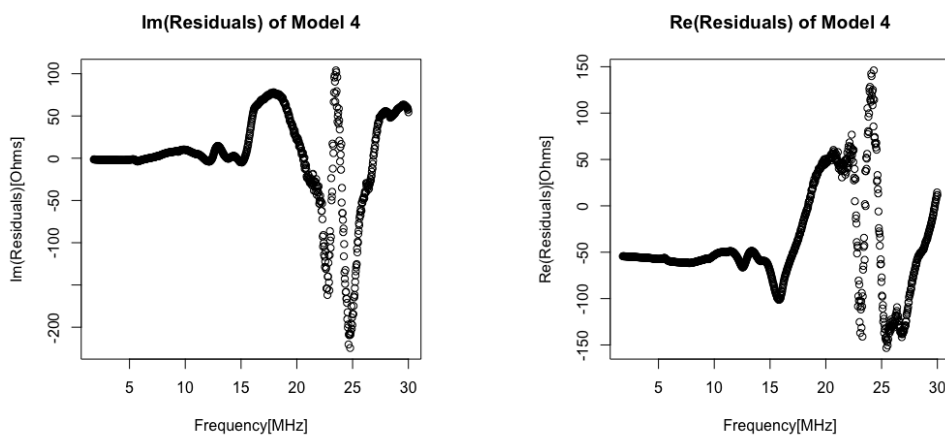
we may intuitively assess our model's predictive performance based on the quantity of AIC. Table 5.3 data reveals that Model 2 has the smallest AIC, hence we may consider model 2 to be the best model at present. Predictive performance has been greatly enhanced by the addition of an inductor to Model 1. Additionally, the AIC of Model 3 is significantly higher than that of the previous model, which confirms our preliminary judgment based on the residual plot of Model 3.

Model 2 is the superior model. Yet, although model 3's predictive accuracy is quite weak, but its real part of the residuals no longer fluctuate around -50 Ohms. Hence, we attempt to integrate them in two distinct ways to create two new models, model 4 and model 5, depicted in Figure 5.10.



**Figure 5.10:** Two new candidate models, Models 4 and 5, will be evaluated for their predictive performance.

The residual plot for Model 5 is shown in Figure 5.15, which reveals that the residual plot of Model 4 is analogous to that of Model 2. Both fluctuation and kurtosis are essentially unchanged.



**Figure 5.11:** Residual against Frequency graph of Model 4 testing the  $4\mu\text{H}$  data sets

Table 5.4 shows the AIC of Model 2 and Model 4. Model 4's AIC is only

	Model 2	Model 4
AIC	22213.40	22246.96

**Table 5.4**

slightly greater than that of model 2. Consequently, we can infer that the addition of a parallel resistor does not appear to greatly enhance our model's simulation performance.

Table 5.5 represents the log-likelihood estimate ( $\hat{\theta}$ ) of Model 5 that was calculated using the Bayesian Optimisation and BFGS methods. 'Inductance(2)' appeared in the Table 5.5 corresponds to the newly added inductor labeled with 'L2' shown in Figure 5.10. The extremely low value of Inductance(2) indicates that the L2 inductor contributes almost nothing towards the prediction of the impedance. Hence, the predictive ability of model 5 is essentially identical to that of its nested model, model 3.

Model 5	Inductance[H]	Capacitance[F]	Resistance[ $\Omega$ ]	Inductance(2)[H]
Parameters	4.1065	11.4026	6350.8349	$2.2 \times 10^{-9}$

**Table 5.5:** Values of Model 4's parameters after applying the BFGS and Bayesian optimisation.

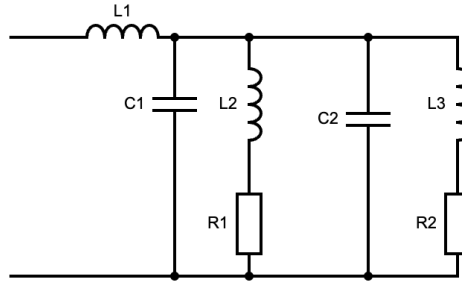
Table 5.6 demonstrates that the Akaike information criteria for Models 3 and 5 are similar, confirming our hypothesis that Model 5 is comparable to Model 3.

	Model 3	Model 5
AIC	24695.37	24701.43

**Table 5.6**

Initial attempted model, Model 1 and Model 2, have performed exceptionally well in model predictive performance to  $4\mu H$  data set. Based on the AIC values of models 4 and 5, we determined that adding or repositioning electronic components had a negligible effect on the simulation of  $4\mu H$  data. It is worth exploring the scenario in which connecting Model 1 and Model 2 **in series** and then to determine whether combining the two models would considerably enhance the model's predicted accuracy. I anticipated Figure 5.12 to depict a generic circuit illustration.

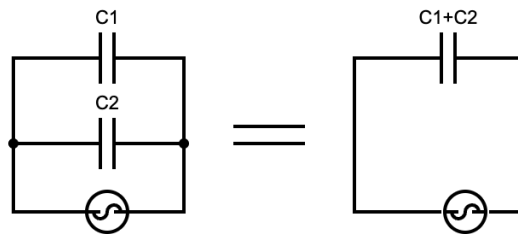




**Figure 5.12:** The speculated model that combine Model 1 and Model 2.

The combination of electric components of  $L1$ ,  $C1$ ,  $L2$ , and  $R1$  in Figure 5.12 is Model 2, and  $C2$ ,  $L3$ , and  $R2$ , which represent Model 1. Nevertheless, circuit diagram shown in Figure 5.12 cannot be utilised straightly as model 6. Because the two capacitors ( $C1$  and  $C2$ ) are non-identifiable in the electric circuit since they are observationally equivalent [21]. There are distinct capacitance values of capacitors that result in the same data distribution.

As illustrated in Figure 5.13, two parallel capacitors each with capacitance  $C1$  and  $C2$  respectively, is identical to the capacitance ( $C1+C2$ ) of a single parallel capacitor. Similarly, the  $L2$  and  $R1$  branch in Figure 5.12 are not identifiable from the



**Figure 5.13**

$L3$  and  $R2$  branch. Nonetheless, for our new model, retaining a set of unidentifiable branches and eliminating the unnecessary capacitor is acceptable. The ultimate version of Model 6 is illustrated in Figure 5.14.

When comparing the residual plot of model 6 to the residual plot of the image of model 2, the best model at present (Figure 5.7), the difference between them is negligible, and their images are nearly identical.

Table 5.7 and displays the AIC of the all candidate models.

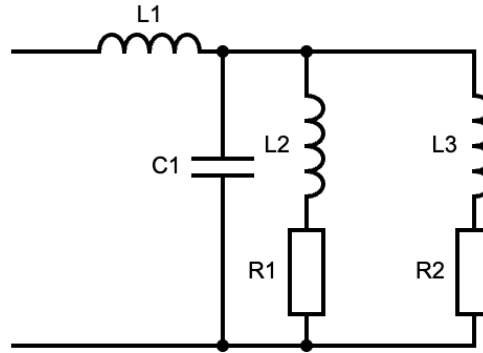
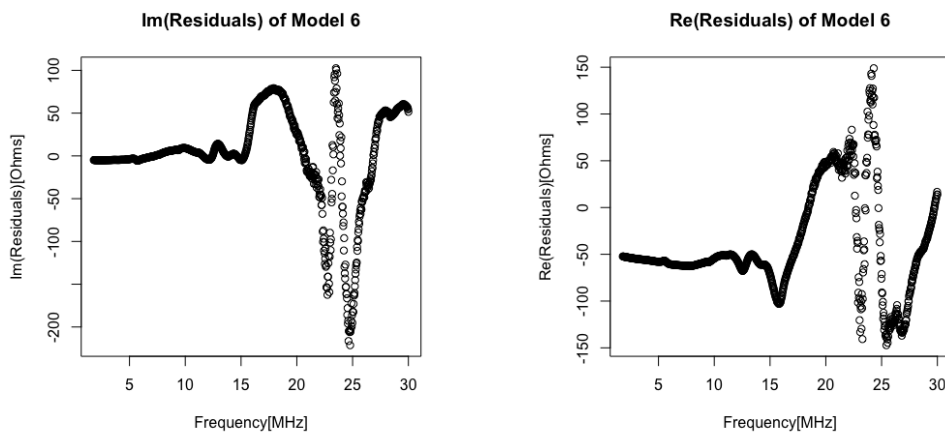


Figure 5.14: Model 6

Figure 5.15: Residual against Frequency graph of Model 6 testing the  $4\mu\text{H}$  data sets

At this stage, it is possible to performing the Kolmogorov-Smirnov test to check the normality of the residuals for current 6 model. The p-value of the residuals' real and imaginary part are compared for uniformity in the Table 5.8. The significance results for all existing models are less than  $2.2 \times 10^{-16}$ , which means there are substantial evidence to deny the hypothesis that the residuals followed the normal distribution.

However, even if the final p-value is very little, I expected model 6's significance results to be slightly larger than the p-values of the other models. The fact that the p-value of all models is tiny and hence equal to  $2.2 \times 10^{-16}$  indicates that

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
AIC	23561.17	22213.40	24695.37	22246.96	24701.41	22208.03

Table 5.7

Kolmogorov-Smirnov Test	Residual	Residual
Model 1: P- value	$<2.2 \times 10^{-16}$	$<2.2 \times 10^{-16}$
Model 2: P- value	$<2.2 \times 10^{-16}$	$<2.2 \times 10^{-16}$
Model 3: P- value	$<2.2 \times 10^{-16}$	$<2.2 \times 10^{-16}$
Model 4: P- value	$<2.2 \times 10^{-16}$	$<2.2 \times 10^{-16}$
Model 5: P- value	$<2.2 \times 10^{-16}$	$<2.2 \times 10^{-16}$
Model 6: P- value	$<2.2 \times 10^{-16}$	$<2.2 \times 10^{-16}$

**Table 5.8**

the predictive performance improvement gain from our model updates is not statistically significant.

We may anticipate that there will be a difference between the value predicted by our model and the actual measured data, as this error is inherent in the model's prediction. Residual's degree of volatility is determined by its variance, which is also known as a stochastic error. Nonetheless, according to Table 5.8, none of the six models currently under consideration satisfy the normality assumption. This also indicates that the systematic error of our model is significantly greater than the expected stochastic error.

Model 6 has the smallest AIC of the six tested models, with a value of 22264.71. Due to the fact that the AIC of all models are quite high, the AIC values of models 6, 2, and 4 are not significantly different (Difference of AIC is around 2000).

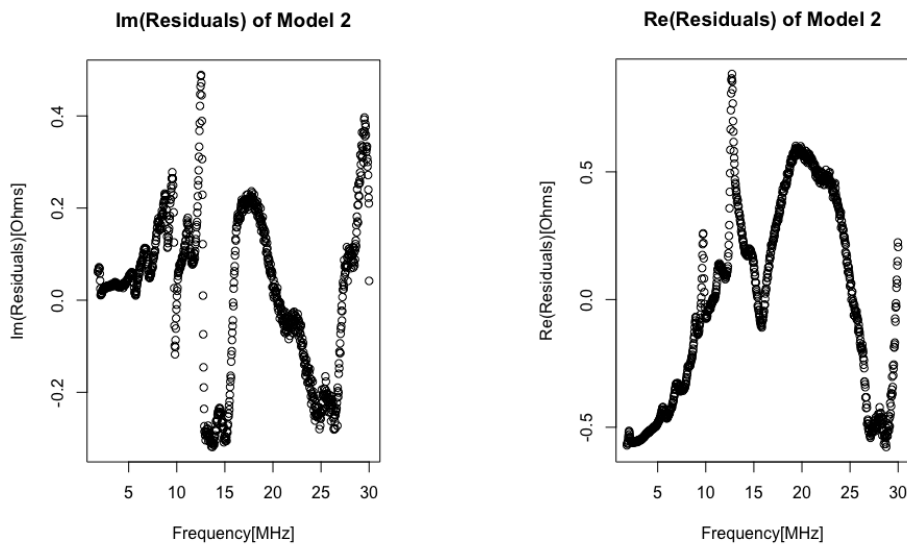
## 5.2 Simulation of 0.5 Microhenry to 2 Microhenry Cylindrical Inductor

	0.5 $\mu$ H dataset	1 $\mu$ H dataset	2 $\mu$ H dataset	4 $\mu$ H dataset
AIC of Model 1	4526.18	3103.84	14615.05	23561.17
AIC of Model 2	2065.27	270.54	11631.38	22213.40
AIC of Model 3	5915.34	4916.83	13511.53	24695.37
AIC of Model 4	2987.33	1543.83	10618.97	22246.96
AIC of Model 5	5917.33	4920.69	11574.61	24801.41
AIC of Model 6	4502.32	1807.58	11632.28	22208.03
Best Model	Model 2	Model 2	Model 4	Model 6

**Table 5.9:** AIC for six current models testing data sets ranging from 0.5 $\mu$ H to 4 $\mu$ H

The AIC value of the existing model's simulation of 0.5-4  $\mu\text{H}$  data is displayed in Table 5.9. We can see that model 2 simulates 0.5-1  $\mu\text{H}$  data the best. Models 4 and 6 performed optimally while modelling 2 and 4  $\mu\text{H}$  data, respectively. All of our models exhibit great predictive performance when simulating 1  $\mu\text{H}$  data set (Models have a mean AIC value of 3,600), and there is little variation in their AIC values. One probable explanation is that when simulating a 1  $\mu\text{H}$  coil, the number of turns of coils is substantially less than that of the other coils with a larger inductance. As a result, the former coil has a lower inter-winding capacitance (Inter-winding capacitance is illustrated in Figure 2.6), and is therefore easier to simulate. In contrast, the predictive performance of all models simulating 4  $\mu\text{H}$  and 2  $\mu\text{H}$  data are quite constrained.

Model 2 that simulating the 2 $\mu\text{H}$  data set has the lowest AIC through out(AIC=270.54). It's residuals graph is shown in Figure 5.16.



**Figure 5.16:** Residual against Frequency graph of Model 2 simulating the 2 $\mu\text{H}$  data sets

Figure 5.7 depicts the residual plot of Model 2 when simulating the data of 4  $\mu\text{H}$ . We can observe that the range of both the real and imaginary parts of residual is between 55 and -200 Ohms. Note that we are now simulating the 0.5  $\mu\text{H}$  data, we expect the relative error to be within (6.875, -25)[Ohms], which is approximately eight times smaller than the range depicted in Figure 5.7. However, the residual



plot degree of volatility depicted in Figure 5.16 is significantly lower than what we anticipated, the range of residual is between 0.5 Ohms and -0.2 Ohms. Yet, they are not a random scattering of points around (0,0), as the jitter of the points follows a very distinct pattern.

We have not yet discovered a model with superior performance. But, it is quite difficult to find the ideal model in real life. As often attributed [4] to George Box, "All models are wrong, but some are useful".

### 5.3 Modeling the Low Frequency Range

We can infer that the ability to predict inductor in around self-resonant frequency [24Mhz] is really limited by observing residual plots for  $4\mu H$  data set. However the prediction ability of models may be enhanced so long as they remain below self-resonance frequency. In order to confirm our claim, we will test the model with the best performance when simulating the varied data set from 0.5 to 4  $\mu H$  recorded in Table 5.9. Each model will now simulate the low frequency range of 1.8 to 15 MHz only. The result is shown below.

	Model 2 fitting 0.5uh data	Model 2 fitting 1 uh data	Model 4 fitting 2 uh data	Model 6 fitting 4uh data
AIC for [1.8-30Mhz]	2065.27	270.54	10618.97	22208.03
$E_{pp}$ [1.8-30Mhz]	1.02	0.131	5.30	11.1
$E_{pp}$ [1.8-15Mhz]	-0.611	-1.04	1.652	6.4
Resonant Frequency	101Mhz	86Mhz	35Mhz	24Mhz

**Table 5.10:** Comparison of the Error per point( $E_{pp}$ ) for the best model fitting the 0.5-4 $\mu H$  data sets with a frequency range of 1.8 MHz to 15 MHz and the entire frequency range(1.8-30MHz).

There are 1000 data sets for the entire frequency range of 1.8-30Mhz, and 500 data sets for the low frequency range of 1.8-15Mhz. The 2 data set have different length, therefore, I am applying Error per point  $E_{pp}$  to evaluate the performance of the same model at different frequency. Since Error per point is a relatively straightforward statistic for determining the amount of error at each data point. It is defined as:

$$E_{pp}(\theta) = \frac{1}{2v_{max}} \sum_v^{v_{max}} |X_{mes}(v) - X(v, \theta)|^2. \quad (5.14)$$

$v_{max}$  represent the maximum of frequency of the choosing data set.

If our model's predictive performance is consistent across the entire frequency range, then we might expect  $\ell(\hat{\theta}; D1) \approx 2\ell(\hat{\theta}; D2)$ , where  $D1$  and  $D2$  represent whole range frequency and low frequency data set respectively and the Error per point  $E_{pp}$  should be approximately consistent within the full frequency range.

By observing the Error per point ( $E_{pp}$ ) in Table 5.10 . When simulating the low frequency range of 0.5 and 1 data, the Error per pint of model 2 becomes negative, indicating that the error is reduced and the model's prediction is more accurate, although the improvement is not very significant. Small improvement can be explained by the skin effect, which essentially indicates that when the frequency of alternating current increases, the conductor's effective resistance increase either and causes variations in alternating current that affect impedance. Ken [28] elaborated in his report on how to simulate the skin effect of an inductor which requires more electromagnetic understanding.

The maximum frequency of data set is 30Mhz, and the resonant frequency of  $0.5\mu H$  and  $1\mu H$  data is 101Mhz and 86Mhz, respectively. As a result, model 2 does not need to simulate inductor performance at resonant frequency, so the model predictive performance is almost consistent across the entire frequency range.

However, the  $2\mu H$  and  $4\mu H$  data sets have resonant frequencies of 35Mhz and 24Mhz, respectively. Therefore, model must simulate inductor impedance at resonant frequency, which is extremely difficult to model. So there is a significant improvement in performance when the model simulates low frequency range by considering the quantity of  $E_{pp}$ .

3678



## Chapter 6

# Inductance for 2 inductors connected in series

This section will discuss how accurate it is to assert that measured impedance at  $1\mu\text{H}$  (Denoting by  $X_{1\mu\text{H}}$ ) Plus measured impedance at  $4\mu\text{H}$  ( $X_{4\mu\text{H}}$ ) = measured impedance at  $5\mu\text{H}$  ( $X_{5\mu\text{H}}$ ). Given the magnetic coupling effect, the total impedance ( $X_{total}$ ) of series connected inductors (with individual impedance  $L1=L2$  for example.) that are close enough to one other may not be  $X_{total} = 2L1$ . In the case of perfect coupling, where all of the magnetic flux of each inductor points in the same direction, resulting in zero magnetic flux loss, the total impedance  $X_{total} = 4L1$ . In addition, the concept with perfect anti-coupling action, all magnetic flux cancels out, then the total impedance  $X_{total} = 0$ .

I'm testing the coupling effect between series-connected inductors using the log-likelihood ratio test. The null and alternative hypotheses are:

$$H_0: X_{1\mu\text{H}} + X_{4\mu\text{H}} = X_{5\mu\text{H}} \text{ (Suggesting no coupling effect).}$$

$$H_1: X_{1\mu\text{H}} + X_{4\mu\text{H}} \neq X_{5\mu\text{H}} \text{ (Suggesting an existence of coupling effect).}$$

I will use the RLC model (Shown in Figure 2.2), which is the nested model and the first model I chose for this project to compute the log-likelihood function. And I will use the expression of  $RLC_{1\mu\text{H}}$  to denote RLC model modeling the  $1\mu\text{H}$  data.

$\ell_0(\theta_0)$  represent the log-likelihood functions under the Null hypothesis for the modeling of  $RLC_{4\mu\text{H}}$  &  $RLC_{1\mu\text{H}}$ . It's expression is shown in the formula below

where  $\theta_0 \in [0, \infty] \subseteq R^6$ , since there are two models with three-dimensional parameters each (which are inductance, capacitance, and resistance).

$$\begin{aligned} \ell_0(\theta_0) = & -\frac{1}{2} \left( \sum_v [X_{5\mu H}(v) - (RLC_{1\mu H}(v, \theta_0) + RLC_{4\mu H}(v, \theta_0))]^2 + \right. \\ & \left. \sum_v [X_{1\mu H}(v) - RLC_{1\mu H}(v, \theta_0)]^2 + \sum_v [X_{4\mu H}(v) - RLC_{4\mu H}(v, \theta_0)]^2 \right) \end{aligned} \quad (6.1)$$

The log-likelihood function of  $\ell_0$  simply expresses one thing: we are aiming to identify model's parameters that not only minimise the log-likelihood when modelling  $1\mu H$  and  $4\mu H$  data separately, but also minimise the log-likelihood when they are incorporate with each other in modelling the  $5\mu H$  data.

$\ell(\theta)$  represent the sum of the log-likelihood function for the Alternative hypothesis for the modeling for RLC of  $1\mu H$ ,  $4\mu H$ , and  $5\mu H$  data solely. It can be expressed as:

$$\begin{aligned} \ell(\theta) = & -\frac{1}{2} \left( \sum_v [X_{5\mu H}(v) - RLC_{5\mu H}(v, \theta)]^2 + \right. \\ & \left. \sum_v [X_{1\mu H}(v) - RLC_{1\mu H}(v, \theta)]^2 + \sum_v [X_{4\mu H}(v) - RLC_{4\mu H}(v, \theta)]^2 \right). \end{aligned} \quad (6.2)$$

Where  $\theta \in [0, \infty] \subseteq R^9$

Under the Null hypothesis, if inductors are uncoupled, then we would expect the log-likelihood function  $\ell(\theta)$  and  $\ell_0(\theta)$  has no significant difference, which indicating that the RLC model modeling the  $1\mu H$  and  $4\mu H$  inductor are statistically indistinguishable form the RLC model modeling the  $5\mu H$  inductor.

Test Statistics T, which is the 2 times of the difference between  $\ell(\theta)$  and  $\ell_0(\theta)$ , follows the Chi-square distribution with the degree of freedom of 3 [8].

$$T = -2(\ell(\theta) - \ell_0(\theta)) \sim \chi_{9-6=3}^2 \quad (6.3)$$

By applying the statistical optimisation to the Equation 6.1 and 6.2, we are be able to find the maximum log-likelihood estimate, and hence compute the quantity of  $\ell(\theta)$  and  $\ell_0(\theta)$ , which are -26533 and -24271 respectively. High-valued likeli-

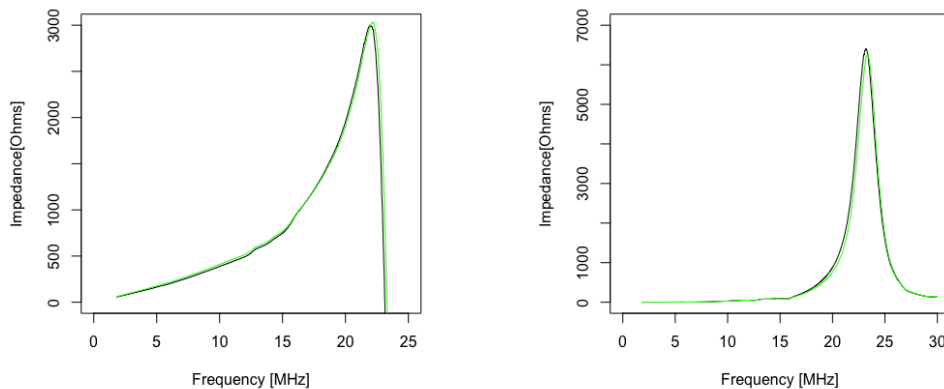
hood function implies that the data set is more consistent with the model than with a model with a lower likelihood function. In our context,  $\ell(\theta)$  has a relative larger value compared with  $\ell_0(\theta)$ .

If we choosing a significance level of 5% ( $\alpha=0.05$ ). The Null hypothesis will be rejected if:

$$T \geq \chi_{3,0.95}^2 \quad (6.4)$$

Since  $T = 4525$ , it means that it falls within the rejection region. Therefore, there is enough evidence to reject the null hypothesis ( $H_0$ ) and conclude that coupling effect exist with the 2 series-connected inductors.

Once we have determined that the coupling effect exists, we must know the extent of the effect. Related graph is shown below.



**Figure 6.1:** This graph compares the data set of a  $1\mu H$  inductor connected in series to  $4\mu H$  inductor (Shown in Green) with the theoretical  $5\mu H$  inductor (Shown in Black).

In the low frequency range, the modelling effect is not readily apparent, and the data between  $1\mu H$  and  $4\mu H$  is essentially identical to theoretical  $5\mu H$  data. However, at the high frequency interval, the data show little divergence. It is worthy to notes that the y-axis has a relatively large scale, slight devaiatons from the graph may result in a significant issues.

## Chapter 7

# Conclusion

In this study, I tested a total of six distinct nonlinear models and performed statistical analysis on data sets ranging from  $0.5\mu H$  to  $4\mu H$  in the entire frequency range and the low frequency range, respectively. All models did reasonably well in modelling the  $0.5\mu H$  and  $1\mu H$  data sets. While the predictive performance for current 6 nonlinear models modelling the  $2\mu H$  and  $4\mu H$  data sets are limited.

Due to the skin effect, the predictive performance of all models in the low frequency range has improved, and the performance of the models that simulate the  $2\mu H$  and  $4\mu H$  data sets has improved substantially, as they are no longer required to simulate the impedance near the resonant frequency.

The first nest model examined in this study is an RLC, which has only three parameters. As a result, if Bayesian optimization is not necessary at first, the log-likelihood function can be optimised with the BFGS approach alone (with reasonable starting values) to give me a relatively accurate parameter output. Nevertheless, when attempting more sophisticated models with higher dimensions, such as Model 2 and model 6, the role of BFGS becomes very limited, and the maximum log-likelihood estimate cannot be appropriately determined. The Bayesian optimisation method was used to successfully resolve the situation.

However, carrying out the bayesian optimisation is not a simple procedure. Before formally using the bayesian optimisation to the objective function, we need to find the constraint for the maximum log-likelihood estimate to ensure that Bayesian optimization is meaningful and efficient. This step is important since it make sure

the maximum likelihood can be accurately found in subsequent updates. Some electronic components may ~~do~~ not have any predictive power at all in the model, so the corresponding parameter show up as 0 during optimization. Therefore, I know from the prior knowledge that all of my parameters must be greater than or equal to 0. In addition, the exploratory data analysis I conduct at the beginning of the Section 5.1 also helped me in defining the parameter space. The result has showed that when simulating the  $4\mu H$  data set, using the model that contain  $4\mu H$  inductor would be a good choice.

## 7.1 Future work

In this project, I simulated individual inductors primarily, but this is insufficient. I have verified that the existence of coupling effect between two inductors using hypothesis test in the preceding Section 5.3. In this context, it is essential to simulate the inductor coupling effect because real-world appliances such as refrigerators and air conditioners contain a greater number of **I**nductors. As a result, accurately simulating the coupling effect is important and a future aim before modeling the series-connected inductor.

The inductor behaves differently at different temperatures. This study did not account for environmental factors such as temperature variation when measuring inductor impedance data. In the future, when modelling inductor, we can therefore also account for the surrounding temperature distribution during measurements.

Furthermore, because of the structured residual plots, the performance of our model in replicating 0.5 and  $1\mu H$  data sets that do not contain resonant frequency cannot be described as outstanding. One likely explanation is that the impedance of wires between electrical components is not considered in our inductor modelling. As a result, I believe we can attempt to model the wire impedance and discover an acceptable model. This model can be appropriately paired with our inductor candidate model in the future inductor modeling.



## Appendix A

# Appendix

In this project, Rstudio was the primary programming language utilised. The code that follows is the code I mostly utilised for this research for Optimisation problem. And I will add comments to each code to explain its purpose. While the code for the optimisation procedure of other models are in the same format with various parameters, I will just demonstrate the RLC model's code for our basic comprehension.

```
## Creat a function that automatically compute the log-likelihood.

LLH_RLC<-function(x,y,z) {
  list(

## Firstly, transfer the complex number into 2 x n matrix form,
## where n is the length of the dataset

X_mes_Re<-matrix(c(Re(inductance.table$Z[bit.subset])),nrow = 1),
X_mes_Im<-matrix(c(Im(inductance.table$Z[bit.subset])),nrow = 1),
X_mes<-rbind(X_mes_Re,X_mes_Im),

## Secondly, transfer the predicted impedance into 2 x n matrix form.

X_modeled_Re<-matrix(c(Re(LLH_RLC(inductance.table$f[bit.subset],
c(x,y,z))))),nrow=1),
X_modeled_Im<-matrix(c(Im(LLH_RLC(inductance.table$f[bit.subset],
c(x,y,z))))),nrow=1),
X_modeled<-rbind(X_modeled_Re,X_modeled_Im),
```



```


## Obatin the residuals
residuals<- X_mes-X_modeled,

## Construct the covariance matrix of the residuals
cov<-var(t(matrix(c(residuals[1,],residuals[2,]),nrow=2,
                    ncol = ncol(residuals),byrow = TRUE))),

## Determine the invser of the covariance matrix
cov_inv<-solve(cov),

## Compute the log-likelihood function.
Score=
log_likerlihood<-(-1/2)*sum(abs(apply(t(X_mes-X_modeled),1,function(x) x*%cov*%
+log(det(2*pi*cov))),
predict=c(x,y,z))
}

```



In the example code that I have shown above. LLH\_RLC is the log-likelihood function of RLC model. And I am simply using the Equation 5.9 to find the log-likelihood function. Moreover, the code "inductance.table\$Z [bit.subset]" means choosing the wanted impedance data sets according to the logical command "[bit.subset]" which simply allocate our different data sets using bit array method.

Now I can carry the Bayesian optimisation to the log-likelihood function.

```

RLC_par<-BayesianOptimization(LLH_RLC,bounds =
list(x = c(0,1),y=c(0,3),z=c(0,5)),
init_points = 20, n_iter = 50, acq = "ucb",
kappa = 2.576, eps = 0.0,verbose = TRUE)

```

Here, I define the log likelihood function of RLC model as an objective function and search for parameters that maximise the log likelihood. In Bayesian optimization, I

specify the interval in which I believe a parameter may exist, as well as the number of sampling and iterations.

After obtain the best estimate from the Bayesian Optimisation. Then, I construct a new function known as "LLH", which is essentially the log-likelihood function with an unknown parameter, so that it can be utilised with the BFGS optimisation method.

```
LLH<-function(parameter) {
  nll<-LLH_RLC(parameter[1],parameter[2],parameter[3])$Score
  return(-nll)
}
```

BFGS methods can be employed to more precisely find our maximum Likelihood estimates based on the results of Bayesian optimisation. Therefore, I use the "*RLC\_par\$Best\_Par*" as the starting value of BFGS method to locate the maximum Likelihood estimates.

```
opt.result.l2= optim(fn= LLH,par=c(RLC_par$Best_Par),
                    lower=c(0,0,0,0),method="BFGS",
                    control=count)

opt.result.l2
```



# Bibliography




- [1] (1950), S. O. Essentials of electricity for radio and television. 2nd ed.
- [2] ACQUAH, H. D.-G. Comparison of akaike information criterion (aic) and bayesian information criterion (bic) in selection of an asymmetric price relationship.
- [3] ACQUAH, H. D.-G. Comparison of akaike information criterion (aic) and bayesian information criterion (bic) in selection of an asymmetric price relationship.
- [4] ANDERSON, C. The end of theory: The data deluge makes the scientific method obsolete. *Wired magazine* 16, 7 (2008), 16–07.
- [5] ANSCOMBE, F. J. Examination of residuals.
- [6] BENDER, R., AUGUSTIN, T., AND BLETTNER, M. Generating survival times to simulate cox proportional hazards models. *Statistics in medicine* 24, 11 (2005), 1713–1723.
- [7] BLANCHARD, J. The history of electrical resonance. *The Bell System Technical Journal* 20, 4 (1941), 424–431.
- [8] CASELLA G, **B. R. L.** *Statistical inference*. Cengage Learning., 2021.
- [9] COOK, R. D., AND WEISBERG, S. *Residuals and influence in regression*. New York: Chapman and Hall, 1982.
- [10] CULLITY, B. D., AND GRAHAM, C. D. *Introduction to magnetic materials*. John Wiley & Sons, 2011.



- [11] DENNIS, J., AND SCHNABEL, R. B. Chapter i a view of unconstrained optimization. In *Optimization*, vol. 1 of *Handbooks in Operations Research and Management Science*. Elsevier, 1989, p. 8.
- [12] DETKA, K., GORECKI, K., AND ZAREBSKI, J. Modeling single inductor dc–dc converters with thermal phenomena in the inductor taken into account. *IEEE Transactions on Power Electronics* 32, 9 (2016), 7025–7033.
- [13] DONG TAN, F., VOLLIN, J., AND CUK, S. A practical approach for magnetic core-loss characterization. *IEEE Transactions on Power Electronics* 10, 2 (1995), 124–130.
- [14] FLETCHER, R. Practical methods of optimization (2nd ed.).
- [15] FLETCHER, R. *Practical methods of optimisation*, 2 ed. Chichester ; New York : Wiley, 1987.
- [16] FRANK J.FABOZZI, SERGIO M. FOCARDI, S. T. R., AND ARSHANAPALLI, B. G. *The Basics of Financial Econometric: Tools, Concepts, and Asset Management Applications*. John WileySons, Inc., 2014.
- [17] FRAZIER, P. I. A tutorial on bayesian optimization. *arXiv preprint arXiv:1807.02811* (2018).
- [18] GEISSER, S. *Predictive Inference*. New York, NY: Chapman and Hall., 1993.
- [19] GUPTA, M. S. Georg simon ohm and ohm’s law. *IEEE Transactions on Education* 23, 3 (1980), 156–162.
- [20] HARRINGTON, R. F. *Introduction to electromagnetic engineering*. Courier Corporation, 2003.
- [21] HSIAO, C. Identification. *Handbook of econometrics 1* (1983), 223–283.
- [22] JAN G. DE GOOIJER, K. K. Some recent developments in non-linear time series modelling, testing, and forecasting,.

- [23] KAKANI, S. L. *Electronics Theory and Applications*, 1 ed. New Age International, 2005.
- [24] KAO, K. C. *Dielectric phenomena in solids*. Elsevier, 2004.
- [25] KIPNIS, N. A law of physics in the classroom: The case of ohm's law. *Science & Education* 18, 3 (2009), 349–382.
- [26] KNIGHT, D. W. The self-resonance and self-capacitance of solenoid coils. *DW Knight* (2013).
- [27] KOLSTER, F. The effects of distributed capacity of coils used in radio-telegraphic circuits. *Proceedings of the Institute of Radio Engineers* 1, 2 (1913), 19–27.
- [28] KUNDERT, K. Modeling skin effect in inductors. *The Designer's Guide* (2006).
- [29] LI, Z., LIU, Y., YIN, P., PENG, Y., LUO, J., XIE, S., AND PU, H. Constituting abrupt magnetic flux density change for power density improvement in electromagnetic energy harvesting. *International Journal of Mechanical Sciences* 198 (2021), 106363.
- [30] LIN, T., ZHOU, K., CAO, Y., AND WAN, L. A review of air-core coil sensors in surface geophysical exploration. *Measurement* 188 (2022), 110554.
- [31] LIU, D.C., N. J. On the limited memory bfgs method for large scale optimization.
- [32] MCELREATH, R. Statistical rethinking: A bayesian course with examples in r and stan. *CRC Press*. (2016).
- [33] MCLYMAN, C. . Transformer and inductor design handbook, third edition (3rd ed.).
- [34] MCLYMAN, C. W. T. *Transformer and inductor design handbook*. Dekker New York, NY, USA, 1988.

- [35] MORICONI, R. K. K., AND DEISENROTH., M. High-dimensional bayesian optimization with projections using quantile gaussian processes. *arXiv preprint arXiv:1807.02811* (2019).
- [36] MYUNG, I. J. Tutorial on maximum likelihood estimation. *Journal of mathematical Psychology* 47, 1 (2003), 90–100.
- [37] NOCEDAL, J., AND WRIGHT., S. Numerical optimization / by jorge nokedal, stephen wright. 2nd ed. 
- [38] PASKO, S. W., KAZIMIERCZUK, M. K., AND GRZESIK, B. Self-capacitance of coupled toroidal inductors for emi filters. *IEEE Transactions on Electromagnetic Compatibility* 57, 2 (2015), 216–223.
- [39] PLATT, C. *Electronics (Second edition)*. Maker Media, 2015.
- [40] POWELL, R. G. *Introduction to electric circuits*. Arnold, 1995, p. 44.
- [41] POWELL, R. G. *Introduction to electric circuits*. Arnold, 1995, p. 11.
- [42] PURCELL, E. M., AND MORIN, D. J. *Electricity and Magnetism*, 3 ed. Cambridge University Press, 2013.
- [43] SEARS, F. *University Physics (6th ed.)*. Addison Wesley, 1982.
- [44] SNOEK, J., LAROCHELLE, H., AND ADAMS, R. P. Practical bayesian optimization of machine learning algorithms. *Advances in neural information processing systems* 25 (2012).
- [45] SVOBODA, J. A., AND DORF, R. C. *Introduction to electric circuits*(3rd ed.). Don Fowley (2013).
- [46] VELEY, V. F. C. *The benchtop electronics reference manual* (1st ed.). 229–232.
- [47] WAGENMAKERS, E.-J., AND FARRELL, S. Aic model selection using akaike weights. *Psychonomic bulletin & review* 11 (2004), 192–196.

- [48] WANG, H., VAN STEIN, B., EMMERICH, M., AND BACK, T. A new acquisition function for bayesian optimization based on the moment-generating function. In *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)* (2017), pp. 507–512.
- [49] XU, H., KAMADA, H., NOMURA, S., CHIKARAISHI, H., TSUTSUI, H., AND ISOBE, T. A simple calculation method for center magnetic flux density of a magnetic core electromagnet with a wide air gap. *IEEE Transactions on Applied Superconductivity* 32, 6 (2022), 1–6.
- [50] YAN, Y. rbayesianoptimization: Bayesian optimization of hyperparameters. *R package version 1, 0* (2016). 