

Statistical Variability Analysis of SilGeo

by

Srijan Pabbi

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Applied Science
in
Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2023

© Srijan Pabbi 2023

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

The electronics manufacturing and supply industry is rife with the unchecked circulation of counterfeit and tampered components. Although developments have been made to counter this issue, the methods often involve visual analysis, X-ray imaging, destructive testing or extensive functional verification of individual components. Applying one or more of these solutions to verify electronic components requires a significant investment in time and capital.

In this thesis, we present a statistical evaluation of SilGeo, a non-intrusive counterfeit electronics detection technology presented by Moreno et al. The primary purpose of this evaluation is to progress towards transforming SilGeo from a successful research result into a standardised verification method in industry applications.

Considering SilGeo as a measurement system, we focus on quantifying the variability due to the hardware assembly and determining repeatability and reproducibility. We use Design of Experiments concepts and hardware domain knowledge to identify main effects and generate assembly configurations to be tested. Accordingly, we apply the statistical technique analysis of variance (ANOVA) to obtain the variability results with individual factor contributions.

Acknowledgements

I extend my sincere gratitude to Professor Sebastian Fischmeister, my supervisor, for his invaluable support, guidance, and the opportunity to work under his mentorship as a member of the Embedded Software Group. Additionally, I am deeply grateful to Dr. Carlos Moreno for his exceptional expertise, personal guidance, and unwavering support throughout the course of my thesis, enriching my academic journey with insightful discussions and encouragement at every step.

A big thanks to Kavish Mehta from the Embedded Software Group for going above and beyond, spending over a hundred hours as a second operator, playing a key role in collecting measurement data. This effort significantly strengthened the depth and accuracy of our research.

Dedication

I dedicate this thesis to my family, whose unwavering support and love have been my steadfast anchor throughout my academic pursuit. Their encouragement has fueled my determination, and I am deeply grateful for their enduring presence at every step of this journey.

Table of Contents

Author's Declaration	ii
Abstract	iii
Acknowledgements	iv
Dedication	v
List of Figures	ix
List of Tables	xi
List of Abbreviations	xiii
1 Introduction	1
1.1 Motivation	3
1.2 Contribution	4
2 Background	6
2.1 Two-Port Networks	6
2.1.1 Scattering Parameters	7
2.2 SilGeo Hardware Fingerprinting	8
2.2.1 Mechanism of Fingerprinting: The SONAR Analogy	9

2.2.2	Nature of the Fingerprint	9
2.2.3	Hardware Anatomy of SilGeo	12
2.3	Analysis of Variance (ANOVA)	15
2.3.1	Definitions and Terminology	17
2.3.2	ANOVA Results	22
3	Our Approach	23
3.1	Measurement System Analysis	23
3.1.1	Measurement System Variation	24
3.1.2	Key Results	27
3.2	Sources of Variability	27
3.2.1	Cause and Effect Diagram	28
3.2.2	Sources of Variability in SilGeo	28
3.2.3	From SilGeo Perspective	35
3.3	Capability Study (Gage R&R)	36
3.3.1	Planning GRR	36
3.3.2	Expectation and Analysis	37
3.4	Design of Experiments	38
3.4.1	Basic Principles	40
3.4.2	Design Approaches	42
3.4.3	Orthogonal Arrays	50
3.5	Design of Experiments for SilGeo	55
4	Case Study	60
4.1	Experiment	60
4.1.1	Setup	60
4.1.2	Test Protocol	65
4.1.3	Operator	69

4.2	Distance Metric	69
4.3	Observations and Results	72
4.3.1	Gage R&R	72
4.3.2	Variability Analysis	78
5	Conclusion	86
5.1	Recommendations	87
5.2	Future Work	89
	References	91
	Glossary	95

List of Figures

1.1	Anvil (SilGeo) Product Prototype	4
2.1	Two-Port Network	6
2.2	Representation of a Sweep in Silicon Geometry (SilGeo)	10
2.3	Sample S-Parameter Capture from the Vector Network Analyzer (VNA) . .	11
2.4	Block Diagram for Hardware Setup	12
2.5	Example of a 40-Pin Target and Receptacle	13
2.6	Z80 40-Pin Adapter Board	14
2.7	Multiplexing Block Diagram	15
3.1	Effect of Bias and Variance (Bulls Eye Plot) [1]	24
3.2	SilGeo Assembly Components Connection Diagram	29
3.3	SilGeo Variability Cause and Effect Diagram	34
3.4	The L_8 Linear Graphs [2]	48
3.5	Factors and Interactions	56
4.1	SilGeo Assembly	61
4.2	Z80 IC X-ray	62
4.3	Counterfeit Z80 IC X-ray	62
4.4	P9370A Keysight Streamline USB Vector Network Analyzer, 4.5 GHz [3] .	63
4.5	Acquisition Controller and Data Capture Screen	65
4.6	Target Insertion Process	68

4.7	Main Effects Plot of SilGeo Assembly	83
5.1	Potential Impact of Remedial Actions on Sources of Variability in SilGeo	89

List of Tables

2.1	Two-Port Network Parameters	7
2.2	Format of Analysis of Variance (ANOVA) Results	22
3.1	Forbidden Region for Gage Repeatability and Reproducibility (GRR) (Marked by $\times$)	37
3.2	Estimate of Variance Components	38
3.3	$2^{(4-1)}$ Factorial Experiment	45
3.4	Alias Table for a Half-Fractional 4-Factor Experiment	46
3.5	L_8 Taguchi Orthogonal Array [2]	47
3.6	Triangular Interaction Table for Two-Level Orthogonal Arrays [2]	49
3.7	OA ($16, 2^{15}$) Orthogonal Array	51
3.8	Mutually Exclusive Sets of Column Triplets in the 2^k Designs [4]	52
3.9	Level Increase	52
3.10	OA($16, 2^{12}4^1$) Orthogonal Array	53
3.11	Degrees of Freedom Allocations	57
3.12	Defining Relations and Designs	58
3.13	Designed Experimental Array	59
4.1	Experimental Array	66
4.2	Experiment Order	66
4.3	Two-Way ANOVA Table: Operational Variability (S00)	74

4.4	Two-Way ANOVA Table: Operational Variability (S01)	74
4.5	Two-Way ANOVA Table: Operational Variability (S10)	74
4.6	Two-Way ANOVA Table: Operational Variability (S11)	74
4.7	Variance Components (S00)	75
4.8	Variance Components (S01)	75
4.9	Variance Components (S10)	76
4.10	Variance Components (S11)	76
4.11	Summary Statistics (Nominal Targets)	77
4.12	Summary Statistics (Counterfeit Targets)	77
4.13	ANOVA Table: Assembly Variability (S00)	80
4.14	ANOVA Table: Assembly Variability (S01)	80
4.15	ANOVA Table: Assembly Variability (S10)	81
4.16	ANOVA Table: Assembly Variability (S11)	81
4.17	Summary Statistics (Nominal Targets)	84
4.18	Summary Statistics (Counterfeit Targets)	85

List of Abbreviations

AIAG Automotive Industry Action Group [23](#)

ANOVA Analysis of Variance [xi](#), [xii](#), [6](#), [15–17](#), [21](#), [22](#), [37](#), [40](#), [43](#), [69](#), [72](#), [74–76](#), [78–81](#), [84](#)

ASQ American Society for Quality [23](#)

DIP Dual In-line Package [12](#), [13](#), [61](#), [63](#)

DOE Design of Experiments [23](#), [38–42](#), [46](#), [50](#)

DOF Degrees of Freedom [21](#), [45](#), [49](#), [51–53](#), [79](#)

DUT Device Under Test [8](#), [30](#), [41](#)

EM Electromagnetic [9](#)

EMI Electromagnetic Interference [30](#), [32](#), [88](#)

ESRI Electrochemical Safety Research Institute [1](#)

ETN Engineering Technology News [2](#)

EV Electric Vehicle [2](#)

GRR Gage Repeatability and Reproducibility [xi](#), [26](#), [27](#), [36](#), [37](#), [41](#), [71](#), [72](#), [75–78](#), [86–88](#)

GWh Giga Watt hour [2](#)

IC Integrated Circuit [10](#), [12](#)

MD Mahalanobis Distance [70](#), [71](#)

MDs Mahalanobis Distances [70](#)

ML Machine Learning [90](#)

MSA Measurement Systems Analysis [23](#), [26–28](#)

OA Orthogonal Array [47–51](#), [53](#), [55](#)

OEM Original Equipment Manufacturer [1](#)

OVAT One-Variable-At-A-time [38](#), [39](#)

PC Personal Computer [14](#)

PCB Printed Circuit Board [10](#), [28](#)

S-parameters Scattering parameters [7](#), [10](#), [11](#), [13](#), [15](#), [28](#), [63](#), [70–72](#), [78](#), [79](#), [89](#)

SilGeo Silicon Geometry [ix](#), [x](#), [2–6](#), [8–14](#), [23–26](#), [28–30](#), [33–37](#), [39–41](#), [55](#), [56](#), [60–62](#), [67](#), [77](#), [78](#), [83](#), [86–90](#)

SMP Subminiature push-on [13](#), [31](#)

SONAR Sound Navigation And Ranging [9](#)

SSD Solid State Drive [2](#)

SSDs Solid State Drives [2](#)

UL Underwriters Laboratories [1](#), [2](#)

VNA Vector Network Analyzer [ix](#), [10–12](#), [14](#), [28](#), [30](#), [35](#), [40](#), [55](#), [56](#), [61–64](#), [66](#), [67](#), [69](#), [78](#), [82](#), [84](#), [86](#)

ZIF Zero Insertion Force [12](#), [13](#)

Chapter 1

Introduction

The quest for convenience, connectivity, and automation has made electronic devices all-pervasive, significantly transforming industries such as healthcare, manufacturing, energy, automotive, agriculture, and aerospace. The demand for such electronic devices has grown exponentially and is expected to continue growing. Presently, supply chains are strained more than ever before, and the issue has only been exacerbated by the advent of the COVID-19 pandemic. Like any other competitive industry, the electronic devices and components industry also faces a constant risk of malicious actors. The industry supply is rife with unauthorized replicas or imitations of genuine devices, produced without [Original Equipment Manufacturer \(OEM\)](#) authorization, designed to mimic the appearance and functionality of legitimate counterparts. These devices are sold on the open market, often at lower prices than legitimate parts.

The negative effect of such devices is more profound than it appears on the face of it. Counterfeit products are often significantly inferior in overall quality and performance compared to genuine products, made with substandard materials and lacking proper quality control in the manufacturing process. Not only are the consumers deceived into buying a non-genuine product, but their premature failure or subpar performance gets inadvertently attributed to the [OEM](#), who then loses both revenue and reputation on account of such devices. Furthermore, these counterfeit devices pose a considerable threat to consumer safety due to the lack of veritable data on their conformance to mandatory consumer safety standards established by appropriate governing bodies.

A research article from [Underwriters Laboratories \(UL\) Electrochemical Safety Research Institute \(ESRI\)](#) [5] discusses a study of counterfeit lithium-ion batteries. The primary safety hazard that lithium-ion batteries pose is the risk of thermal runaway and

[deflagration](#). Once the critical threshold temperature of the cell is exceeded, a thermal runaway sets in, which can reach extreme temperatures. This extreme heat is often accompanied by toxic smoke release, electrolyte leakage, and fire, which can cause severe injuries and lead to larger fires similar to what happened in a documented case of counterfeit BlackBerry phone batteries [6]. Counterfeit batteries may use inferior material, not include built-in safety mechanisms, and be manufactured in non-ideal conditions, leading to heightened consumer safety risks. The article citing a [UL](#) study published in [Engineering Technology News \(ETN\)](#) [7] claims that the current energy demand from lithium-ion based storage is approximately 400 [Giga Watt hour \(GWh\)](#) and is expected to grow to more than 2000 [GWh](#), the majority of it being passenger [Electric Vehicle \(EV\)](#). This makes the lithium-ion battery market incredibly lucrative to counterfeiters. As detailed with evidence in the study [7], counterfeit batteries lack protective assembly features, lack performance and capacity, and fail most non-destructive and destructive safety tests.

The risk of physical harm on an individual level from counterfeit devices is only one of the many aspects of this industry-wide issue. Counterfeit device usage also impacts society by threatening safety in the aerospace industry. In 2015, the CEO of a US government defence contractor, Harry Krantz LLC, pleaded guilty to compromising the Department of Defence's supply chain. The accused was involved in falsely remarking unapproved microprocessor chips [8]. Many of these microprocessor chips were used to assemble U.S. Military and Commercial helicopters. The introduction of such counterfeit devices in defence and aerospace infrastructure endangers not only the humans involved in the operation of the equipment but also the nation's defence mission readiness and efficacy of any threat response by handicapping the performance of the military equipment.

On the consumer electronics front, fake Samsung [Solid State Drives \(SSDs\)](#) have been reported to be in circulation [9]. Reports indicate that the counterfeit [SSDs](#) are visually identical in appearance but fail to meet specified performance and can be identified by Samsung's Magician Software. Samsung uses its own designed and manufactured [Solid State Drive \(SSD\)](#) controller and NAND memory storage chips with an onboard [DRAM](#). Whereas the counterfeit [SSDs](#) lack such design features.

It is clear that these malicious devices pose severe threats to the industry. This area presents great potential for novel technological solutions to tackle this issue. [SilGeo](#) is one such solution presented by Moreno et al. [10].

1.1 Motivation

[SilGeo](#) by Moreno et al. [10] is one of the first physics-based hardware assurance technologies. It started as a prototype with exceedingly positive results in detecting counterfeit FTDI ICs, as demonstrated by Vibien in the thesis titled *SilGeo A Method for the Detection of Counterfeit, Compromised, or Tampered Electronic Devices* [11]. From then forward, all efforts have been towards refining and productising [SilGeo](#). Moreno et al. have demonstrated more positive results using [SilGeo](#) to detect non-nominal (counterfeit/malicious) devices with Z80 microprocessors [12]. Figure 1.1 showcases Anvil, the product form of [SilGeo](#), developed with Palitronica Inc.

The vision is to present [SilGeo](#) as one of the standard verification techniques in applications such as transport security, supply chain assurance, counterfeit detection and defective device detection. One of the most crucial goals in achieving the vision for [SilGeo](#) is to gain clarity on its scalability as a hardware assurance solution. [SilGeo](#)'s successful technical demonstrations with various target classes can be viewed as proof of concepts with a scope limited to a single measurement system assembly. The following questions present themselves when using multiple instances of [SilGeo](#) to independently perform hardware verification:

- Repeatability and Reproducibility of a Measurement (Variability in Re-measuring) made using [SilGeo](#).
- The Intrinsic (Assembly) Variability in [SilGeo](#) instances.
- Compatibility and Comparability of Data between [SilGeo](#) instances.



Figure 1.1: Anvil (SilGeo) Product Prototype

1.2 Contribution

The term variability of a measurement system is defined by both its short-term and long-term characteristics. To concretely define the scope of the problem in this thesis, we limit our study of variability to factors affecting the system in the short term. The reasons to follow this approach are as follows:

- As a first study of this kind on [SilGeo](#), we aim to incrementally tackle the problem of quantifying variability and isolating the main sources.
- The factors affecting short-term variability contribute to the long-term variability but vice versa may not be true. Therefore, isolating the two into separate studies is necessary.
- Long-term variability study involves equipment for accelerated ageing and causes irreversible changes to the system.

Furthermore, we explore the variability in operation and variability due to assembly separately. To summarise, with this thesis, we make the following contributions:

- Model the potential sources of variability in [SilGeo](#).
- Develop an approach to measure and quantify the operational and assembly variability.
- Apply the developed method to a case study employing [SilGeo](#) and quantify the variability.
- Provide recommendations to counter the variability-causing sources.

Chapter 2

Background

This chapter discusses details about the concepts and workings of [SilGeo](#) and builds a mathematical foundation for the [ANOVA](#) method that will become the basis of the variability analysis.

2.1 Two-Port Networks

In network theory, a port is defined as a pair of terminals to which when a voltage is applied, the current is conserved, i.e. the input and the output current are equal [\[13\]](#). A two-port network [Figure 2.1](#), is described by the relationship between V_1 , V_2 , I_1 , I_2 . Any two can act as dependent variables and be determined from the values of the remaining.

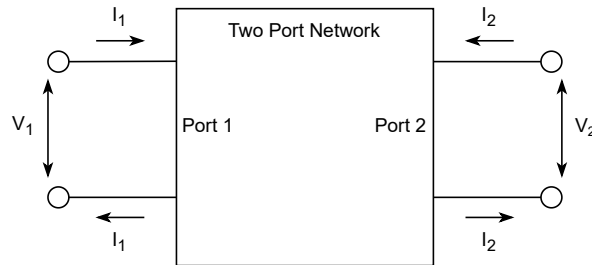


Figure 2.1: Two-Port Network

There can be a total of 6 combinations of unknowns, each relevant to various practical

scenarios. Table 2.1 showcases the different parameters used in literature to represent the network given the combination of known values [13].

Parameter	Unknowns	Relation
Z	(V_1, V_2)	$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{bmatrix} \cdot \begin{bmatrix} I_1 \\ I_2 \end{bmatrix}$
Y	(I_1, I_2)	$\begin{bmatrix} I_1 \\ I_2 \end{bmatrix} = \begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix} \cdot \begin{bmatrix} V_1 \\ V_2 \end{bmatrix}$
H	(V_1, I_2)	$\begin{bmatrix} V_1 \\ I_2 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} \cdot \begin{bmatrix} I_1 \\ V_2 \end{bmatrix}$
$ABCD$	(V_1, I_1)	$\begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \cdot \begin{bmatrix} V_2 \\ -I_2 \end{bmatrix}$

Table 2.1: Two-Port Network Parameters

2.1.1 Scattering Parameters

The measurement of Z and Y parameters poses a challenge as they require the circuits to be terminated with open or short. As stated by Steer [14], this may cause oscillations and undesired behaviour in the case of active circuits. Adopting a measurement characteristic with high cohesion with practical RF circuit properties, such as maximum power transfer and resistive termination is desirable.

Scattering parameters are such a metric, involving measurement of RF in both forward and reflected direction. This is described by the relationship between V_1^+ , V_1^- , V_2^+ , and V_2^- . Reflection coefficient Γ , the voltage ratio of the cumulative emergent (reflected) and the incident wave, is the **Scattering parameters (S-parameters)** for a one-port network. Similarly, for a two-port network, the **S-parameters** and the transformation relation are given as follows:

$$\begin{aligned}
 S_{11} &= \left. \frac{V_1^-}{V_1^+} \right|_{V_2^+=0} & S_{12} &= \left. \frac{V_1^-}{V_2^+} \right|_{V_1^+=0} & \begin{bmatrix} V_1^- \\ V_2^- \end{bmatrix} &= \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \cdot \begin{bmatrix} V_1^+ \\ V_2^+ \end{bmatrix} \\
 S_{21} &= \left. \frac{V_2^-}{V_1^+} \right|_{V_2^+=0} & S_{22} &= \left. \frac{V_2^-}{V_2^+} \right|_{V_1^+=0} & & (2.1)
 \end{aligned}$$

2.2 SilGeo Hardware Fingerprinting

SilGeo is a technology that aims to tackle the problem of counterfeiting in hardware devices by utilizing the nature of the physical construction of a [Device Under Test \(DUT\)](#) (the target). As we discussed, in the industry, several methods have evolved to counter this problem. Still, none of the methods come close to providing a comprehensive solution to the problem. Either they depend on visual analysis of a target and its electrical connections, verification of their device functional characteristics, or destructive testing and imaging of a subset of the targets.

As suggested in Vibien’s thesis [11], for many years, methods to verify pieces of data and software have existed and worked reliably without posing any loopholes and challenges. This method involves the use of [checksum](#) to independently verify any data, file, or software by employing the concept of one-way cryptographic [hashing](#). This underlying concept in the use of [checksum](#) ensures that any tampering with a piece of data will change the value of the calculated checksum, thereby providing a unique fingerprinting mechanism. This fingerprint can be compared with values from authentic sources to verify the integrity and authenticity of the piece of data at the time of use.

In hardware, this analogy dictates that any dissimilarity would translate into an intrinsic physical difference between the target and the original device. As presented in Vibien’s thesis [11], a counterfeit device by nature is different from a genuine device; therefore, it is justified to claim: *If a certain device is counterfeit, then it has an intrinsic physical difference in comparison to a genuine device.* The contrapositive of this statement is as follows: *if a certain device is physically equivalent to a genuine device or doesn’t have intrinsic physical and geometric differences, then it is not a counterfeit.* With this premise of physics-based fingerprinting, using SilGeo, we capture a fingerprint of a target device, and that fingerprint can be then compared with the fingerprint of a known genuine device to determine the conformity of the target.

The fundamental idea behind the approach is to obtain fingerprint data pertaining to an individual target device when treated as a black box. In an ideal scenario, this fingerprint data, minus the variability in the measurement process, is unique for a given individual target and could, in principle, be used to isolate that particular target. The comparison of the captured fingerprint data to a known genuine device is essentially a distance calculation to quantify the magnitude of the intrinsic difference into a confidence interval or a similarity score. This indicates that SilGeo can not only be used for counterfeit sifting but potentially also for applications such as defect and damage detection, tamper detection, supply chain integrity, ageing characterization and other applications.

2.2.1 Mechanism of Fingerprinting: The SONAR Analogy

To gain an intuitive understanding of the underlying concept behind the operation of **SilGeo**, we look at how a **Sound Navigation And Ranging (SONAR)** works and what principles are used in detection and imaging using sonar.

SONAR imaging is a widely used technique in **oceanography** for underwater geological features, **hydrographic surveying** used to map the seafloor and create navigation charts, marine navigation, and many more areas [15]. Outside of the underwater application areas, **SONAR** is also used in acoustic cave mapping [16]. Imaging using **SONAR** typically involves sound wave generation (ping or a sweep), sound wave reflection off the physical objects and landscape geometry, capturing and storing the reflected response, processing the response and interpretation [17]. The premise behind the success of **SONAR** is that the physical shape and arrangement of the landscape or the desired area under scan determines the response, and that response can be used for imaging after signal processing [15] as it encodes the underlying physical and geometrical features possessed by the scan target.

Light or the **Electromagnetic (EM)** waves in the visible band of the Spectrum, emitted by a source and reflected off the surroundings, are captured by the human eye and interpreted by the human brain to form the abstract concept of vision [18]. Looking at this from a system design perspective, the captured light is merely an electrical response sent to the brain via the optic nerve.

With **SilGeo** as showcased by Vibien [11] and Moreno et al. [10], we obtain the *response to a stimulus* that encodes the physical characteristics of a target device that is electrical in nature. This is done by electrically injecting a stimulus of **RF** signal into the target device and capturing the response of the reflected and propagated waves after interacting with the device. This data, encoded with the target device’s intrinsic physical characteristics and geometry, is deemed as the hardware fingerprint of the target device.

2.2.2 Nature of the Fingerprint

This section introduces details relevant to building a base that will help in understanding **SilGeo** as a measurement system and interpretation of the results obtained from the data manipulation and statistical techniques applied in this thesis; in addition to that, it will also help explain the many inherent challenges of the measurement method and the nature of the data itself.

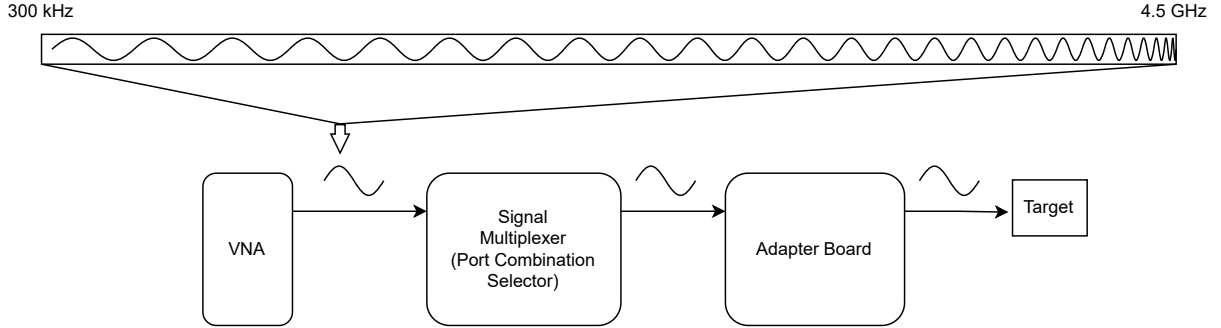


Figure 2.2: Representation of a Sweep in [SilGeo](#)

Stimulus to the Target Device: [SilGeo](#) uses [RF](#) signals as a stimulus to incite a response from the target device. [RF](#) waves propagating through a medium (in this case, the transmission line and the target are affected by the properties of the medium within which they travel; characteristics such as the varying and overall electrical impedance (which includes resistance, capacitive reactance and inductive reactance), the relative permittivity of the materials encountered, and dimensional geometry. These effects vary with the frequency of the injected signal. Therefore, a frequency sweep using a signal generator is the logical requirement from the stimulus.

Measurement of Response: The next crucial feature is to capture the response of the signal sweep, as shown in Figure 2.2. The target device, as we consider it, is a black box and, therefore, a two-port network system. For any given frequency in the sweep, one port is used for injecting the signal and capturing the reflected response, and the other for capturing the propagated signal response and vice versa. [SilGeo](#) uses the [S-parameters](#) as the response data captured using a two-port [VNA](#). A sample of the captured data is shown in Figure 2.3.

Connection to a Target Device: An electrical Target device can have many external interfaces, such as multiple pins in the case of semiconductor [Integrated Circuit \(IC\)](#)s, or electrical contacts of/on a [Printed Circuit Board \(PCB\)](#) in the case of [Shop-Replaceable Unit \(SRU\)](#)s, or user interfaces with multiple electrical contacts such as USB and Ethernet etc... in case of [Line-Replaceable Unit \(LRU\)](#)s. With [SilGeo](#), the variety in this availability of electrical interfaces is considered a bonus to the application of the technique. We will discuss in 2.2.3 how each target class (interface class) as a whole requires a unique [receptacle](#) to connect. Each of the electrical contacts connected to [SilGeo](#) is considered a port.

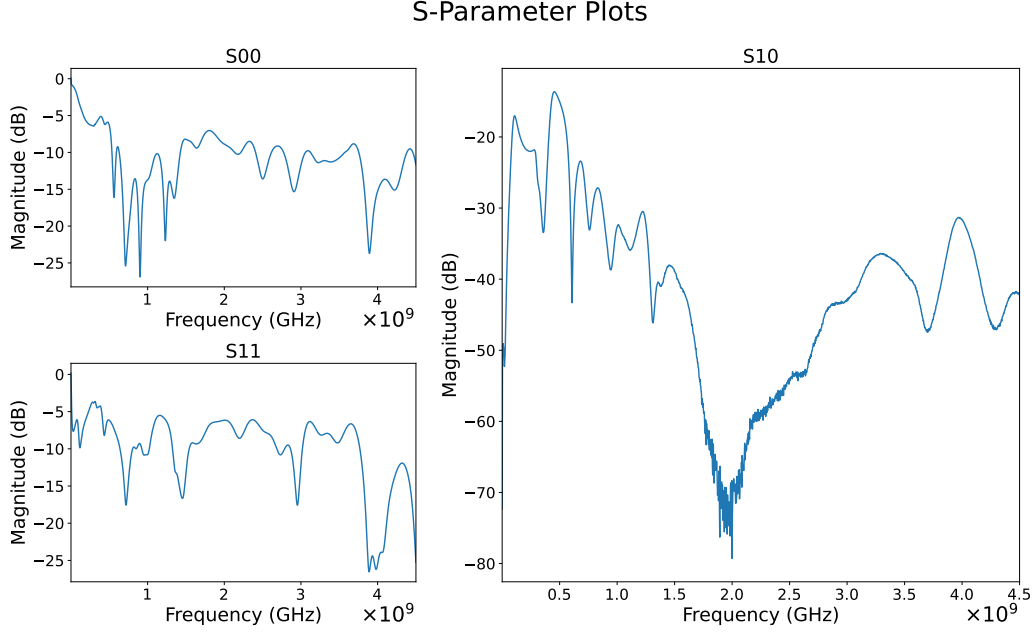


Figure 2.3: Sample S-Parameter Capture from the [VNA](#)

The technique of using multiple port connections for a single target can be intuitively understood by associating it with acoustic cave mapping of a cave with multiple entrances. For each port combination, [SilGeo](#) captures four [S-parameters](#), corresponding to reflection and transmission between the two ports, respectively. And this data is the full hardware fingerprint produced [SilGeo](#).

S-parameter Data: As discussed in Section 2.2.3, [SilGeo](#) is an assembly of several components. Apart from the [VNA](#), all the components are in a direct path of the stimulus signal sweep as shown in Figure 2.2. This causes all the components in the chain to have an effect on the [S-parameters](#) data recorded, and hence the response data is conflated with the effect of all the factors in this signal chain.

With [SilGeo](#), the primary purpose is not to discern the individual effects of each component factor but to compare the cumulative response of a target with the cumulative response of a reference. The issue that remains at hand is that it cannot be assumed that the effect of each component factor is uniform between different instances of the assembled measurement system. Hence, the main goal of this thesis is to quantify and evaluate the

variability aspects of the complete system.

2.2.3 Hardware Anatomy of SilGeo

This section discusses the implementation of [SilGeo](#) as a measurement system. The crux of this section is that [SilGeo](#) as a measurement system is an assembly of four main components (represented as a block diagram in [Figure 2.4](#)); the adapter board (connects to the target device), the signal multiplexer, the [VNA](#) multiplexer interconnect cables.

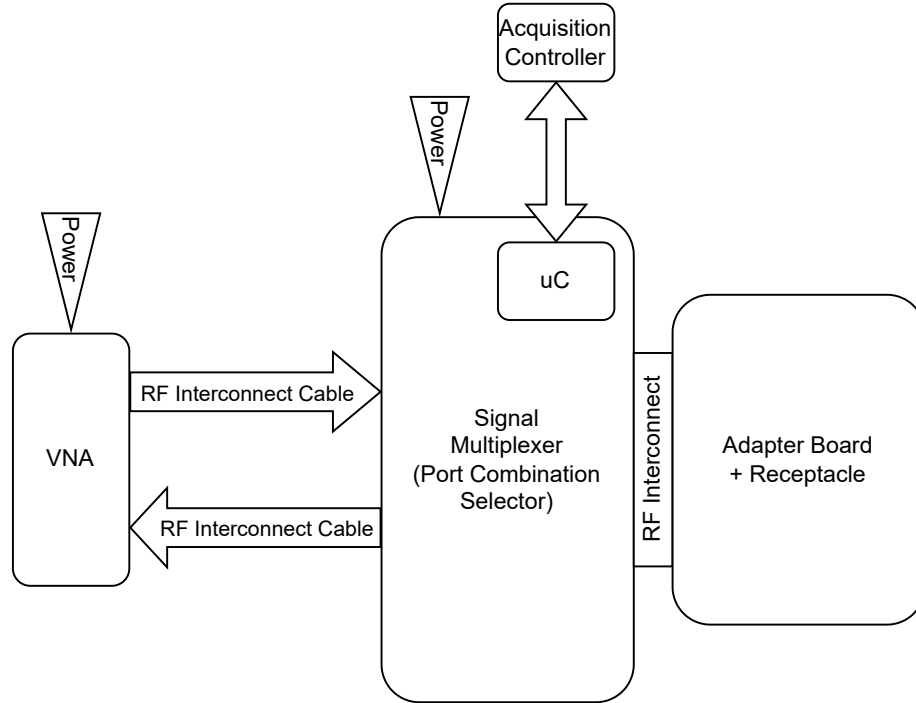


Figure 2.4: Block Diagram for Hardware Setup

Target Adapter (Receptacle): The Target Adapter is a piece of hardware designed specifically around a class of target devices. For context, a target device is an electrical component, for example, a semiconductor [IC](#) packaged as a 40-pin [Dual In-line Package \(DIP\)](#). A receptacle capable of accepting 40-pin [DIP ICs](#) connects this target with the measurement setup. Further in this thesis, we will discuss a case study using a 40-pin [Zero](#)

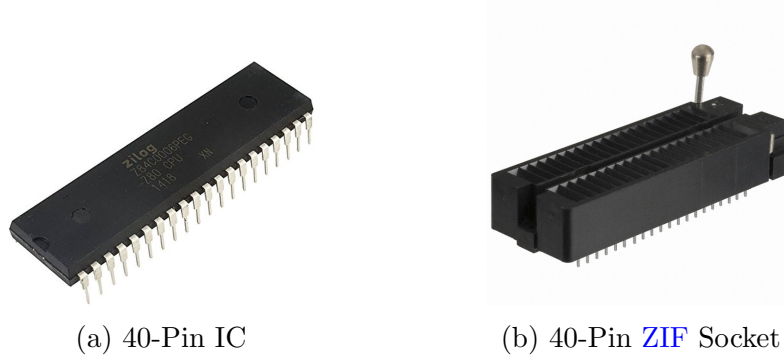


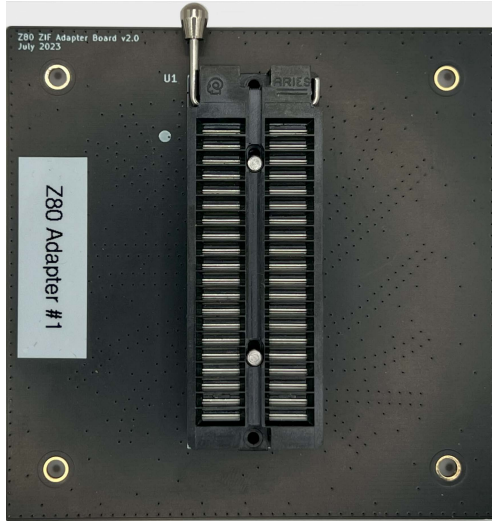
Figure 2.5: Example of a 40-Pin Target and Receptacle

[Insertion Force \(ZIF\)](#) socket (Figure 2.5b) receptacle capable of housing a 40-pin [DIP](#) Z80 microprocessor (Figure 2.5a).

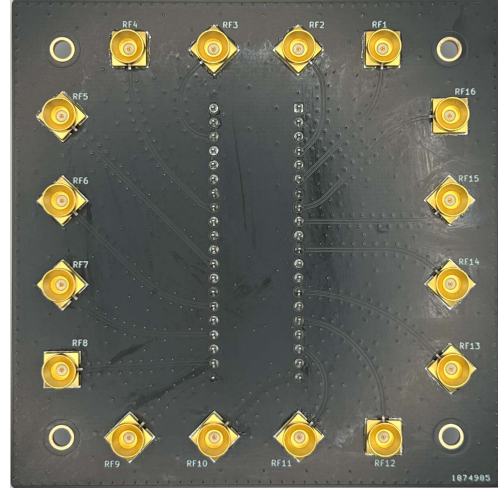
The purpose of using such a receptacle is to offer multiple re-connections without any significant degradation of the receptacle itself. The receptacle is unique for each interface class. It has to be connected with the [multiplexing](#) hardware with a level of indirection in order to create a generalized multiplexer design that connects to a variety of receptacles. Thus the necessity of using an adapter board to house the receptacle. All adapters now have a common interconnect to the multiplexer, but each adapter class has a unique receptacle capable of interfacing with a compatible target device class.

In the current implementation of [SilGeo](#), the design decision is to have the capability of connecting to a total of 64 port combinations with a configuration of 8×8 split in the [multiplexing](#) architecture. Therefore, each adapter consists of 16 [RF-capable Subminiature push-on \(SMP\)](#) connectors as a design standard to connect with the [multiplexing](#) hardware (Figure 2.6).

It is important to note that, of all the other component factors, the receptacles for most targets (sourced as mass-manufactured products) are not for [RF-capable](#) connectivity. Therefore, we expect them to greatly affect the variability of the [S-parameters](#) measurement. As we discussed, the other components in the signal chain are designed and manufactured to an [RF](#) specification. Although variabilities are expected in any assembly component, it is valid to assume that the variations in the receptacle are perhaps the most significant contributor to the overall variability in measurement. In upcoming Sections 3.2 and 3.5, we discuss more on this aspect and accordingly design the experiments.



(a) Top (Receptacle Side)



(b) Bottom (Interconnect Side)

Figure 2.6: Z80 40-Pin Adapter Board

The Signal Multiplexer: The next important design component of [SilGeo](#), is the [RF-capable multiplexing](#) hardware. As discussed previously, the current architecture of [SilGeo](#) considers the target device as a two-port network black box from the perspective of a [VNA](#). Hence, to measure the 8×8 port combinations, the [VNA](#) stimulus and response signals must be multiplexed by using bi-directional [RF-capable](#) analog signal multiplexers. Before each sweep of a port combination the [VNA](#) signal has to be routed to the desired port by using the [multiplexing](#) architecture shown in [Figure 2.7](#).

This [multiplexing](#) topology is controlled by using a microcontroller, which then receives instructions from a user [Personal Computer \(PC\)](#) or the acquisition orchestration system in this scenario, running the acquisition controller program. The acquisition controller commands the signal multiplexer with instructions to select an adapter connection (referred to as the port) for each of the [VNA](#)'s connections. Then, the acquisition controller triggers the [VNA](#) to capture a sweep and store this data corresponding to this port combination, and the process repeats for each of the 64 (8×8) port combinations.

Interconnect RF Cables: This part of the assembly significantly influences the data. The quality of the cable and the length play a role in determining the signal integrity at the multiplexer board. The cable interconnect part specifications are to remain the same for capturing fingerprints that are to be compared. As discussed with the target receptacle,

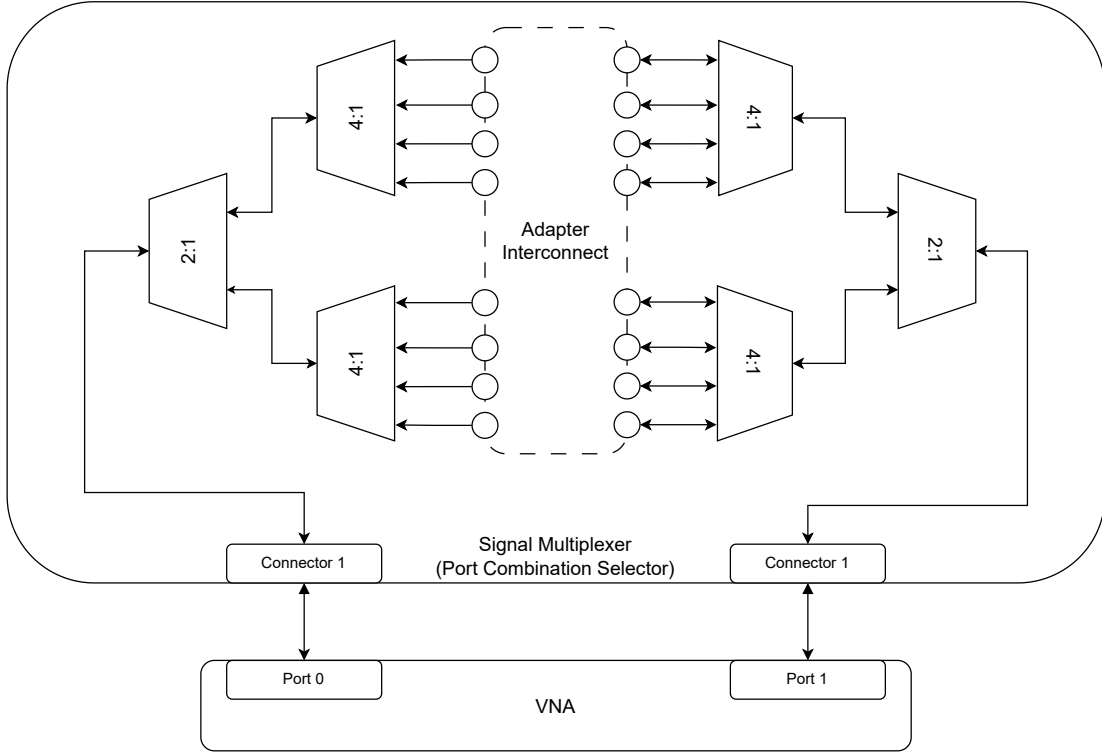


Figure 2.7: Multiplexing Block Diagram

the effect of the [RF](#) interconnect cable is expected to be the same for cables of equivalent certification, specification, and [S-parameters](#).

2.3 Analysis of Variance (ANOVA)

In statistics, [ANOVA](#) is a standard technique employed for data analysis in experiment design and hypothesis testing. It provides a structured and comprehensive way to determine the statistical significance of differences among the groups in data [\[19\]](#). Unlike the t-test, this technique allows us to deal with two or more independent variables simultaneously. [ANOVA](#) not only helps us determine the individual effects of each variable separately but also their interactions [\[20\]](#).

By definition, [ANOVA](#) follows an underlying structural model, where each observation is the function of the total mean, the deviation due to factors and the error (within-group

deviation). Taking an example of a dataset with n captures per level of one factor [20], we can say that each observation will be represented as follows:

$$x_{ij} = \mu + \alpha_i + \epsilon_{ij} \quad (2.2)$$

where,

x_{ij} is the i^{th} observation for group j (level of factor).

μ is the global mean.

$\alpha_j = \mu_j - \mu$ is the deviation of the mean of group j (μ_j) from the global mean (μ).

ϵ_{ij} represents the random error or deviation from the group mean for observation i .

For the underlying model in ANOVA to be applied to a dataset, it is important to understand the purpose and conformance to the following assumptions:

- **Normality of Data** is an assumption that considers the data populations to be distributed normally around the group mean. In other words, any data within a certain group is a random sample from the normal distribution $X \sim \mathcal{N}(\mu_i, \sigma_\epsilon^2)$
- **Homogeneity of Variance** (homoscedasticity) is important when performing hypothesis testing in ANOVA. It is assumed that the variances, and thus the distributions of distinct groups on a continuous scale, are equal, similar or equivalent. The concept of homogeneity of variance assures us that the distributions of outcomes within each group can be compared consistently. Failure to establish similarity in this context among independent groups may lead to the emergence of misleading or spurious results from F-statistic.
- **Independence of Observations:** This assumption is crucial to how data is presented to the analysis method. For a pair of observations within a group, the deviation of one from the group mean is independent of the deviation of the other. In other words, in a within-group, each observation of a target must be obtained by a random sample from the distribution.

Note: It is suggested by [20], in practice, ANOVA is a robust statistical procedure where the assumptions of normality and homogeneity can be violated to a certain degree with a minor impact on results. According to [20], if the largest variance is no more than four times the smallest, the ANOVA results are most likely valid. Additionally, the crucial detail is that unequal sample sizes and non-conformance of the homogeneity assumption together negatively affect the analysis's robustness.

Aside from our primary goal of determining variability, an important aspect of [ANOVA](#) is the statistical significance of differences in grouping based on levels of each factor. This is done by performing the F-test on the data with the following null hypothesis. For a factor with k levels:

$$\begin{aligned} H_0 : \mu_1 = \mu_2 = \dots = \mu_k \\ H_0 : \mu_i = \mu_j \quad \forall i, j \in \{0 \dots k\} \end{aligned} \quad (2.3)$$

The inverse of the null hypothesis is that at least one group in the same factor differs from the other groups.

$$H_A : \exists i, j \in \{0 \dots k\} : \mu_i \neq \mu_j \quad (2.4)$$

2.3.1 Definitions and Terminology

Consider the case of a dataset with two factors, A and B , with n and k levels, respectively, with each observation captured r times. Therefore, the following equation can represent each observation in the dataset [\[21\]](#):

$$x_{ijm} = \mu + \alpha_i + \beta_j + \delta_{ij} + \epsilon_{ijm} \quad \begin{cases} i = 1, 2, \dots, n \\ j = 1, 2, \dots, k \\ m = 1, 2, \dots, r \end{cases} \quad (2.5)$$

where,

- μ is the overall mean effect
- α is the effect of i^{th} level of A
- β is the effect of j^{th} level of B
- δ is the interaction effect of i^{th} level of A and j^{th} level of B .
- ϵ is a random error component.

Factor: is defined as an influence or variation in the experiment.

Group: is defined as a set of units to which the same variation or the same level of a factor is applied.

Sum of Squares

A sum of Squares is a statistical metric used to quantify the deviation of data points from the mean within a dataset. The following equation gives the total sum of squares for the

given use case:

$$\begin{aligned}
SS_T &= \sum (X - \overline{X_{\text{global}}})^2 \\
&= \sum_{i=1}^n \sum_{j=1}^k \sum_{m=1}^r (X_{ijm} - \overline{X})^2
\end{aligned} \tag{2.6}$$

where X denotes the observations and $\overline{X}$ or $\overline{X_{\text{global}}}$ refer to the overall mean of the data.

As detailed in [19], the total sum of squares SS_T can be decomposed into its component sum of squares representing the factors, interactions and errors in the data. The first step is to write it as follows:

$$SS_T = SS_{\text{Bet}} + SS_W \tag{2.7}$$

where,

SS_{Bet} is the sum of square deviations of group means about the global mean. *Bet* is derived from the word *between*.

SS_W is the sum of square deviations of the observations about the group mean. *W* is derived from the word *within*, also refers to the experimental error.

The between-group square deviations can be calculated using the following formula:

$$\begin{aligned}
SS_{\text{Bet}} &= \sum r (\overline{X_{\text{group}}} - \overline{X_{\text{global}}})^2 \\
&= \sum_{i=1}^n \sum_{j=1}^k r (\overline{X_{ij.}} - \overline{X})^2
\end{aligned} \tag{2.8}$$

Decomposition of Variance

Moreover, the between-group square deviations can be broken down further into the following terms for two factors and their interaction term [19, 21].

$$SS_{\text{Bet}} = SS_A + SS_B + SS_{A \times B} \tag{2.9}$$

The formulae for the square deviations of individual factors and the interaction follow [22, 20].

$$SS_A = \sum_{i=1}^n kr (\overline{X_{i..}} - \overline{X})^2 \quad (2.10)$$

$$SS_B = \sum_{j=1}^k nr (\overline{X_{.j.}} - \overline{X})^2 \quad (2.11)$$

$$SS_{AB} = SS_{Bet} - (SS_A + SS_B) \quad (2.12)$$

$$\text{Alternatively, } SS_{AB} = \sum_{i=1}^n \sum_{j=1}^k r (\overline{X_{ij.}} - \overline{X_{i..}} - \overline{X_{.j.}} + \overline{X})^2 \quad (2.13)$$

Using the above formulae and the decomposition of square deviations, the within-group square deviation or the error can be calculated as follows:

$$SS_W = SS_T - SS_A - SS_B - SS_{A \times B} \quad (2.14)$$

$$\text{Alternatively, } SS_W = \sum_{i=1}^n \sum_{j=1}^k \sum_{m=1}^r (X_{ijm} - \overline{X_{ij.}})^2 \quad (2.15)$$

In practice, the interaction term is calculated using Equation 2.12. Subsequently, the within-group deviation or the error term is calculated by Equation 2.14.

Correction Factor/Calculation Formulae

The formulae to calculate square deviations defined in the previous part pose an enormous challenge when scaling to experiments with three or more factors. It is mathematically feasible to realise the conceptual formulae, but due to the complexity of terms resulting from three or more factors, they lose clarity. Moreover, the conceptually realised formulae involve redundant calculations when algebraically simplified. To obtain the practical calculation formulae, consider the total sum of squares of the dataset.

$$\begin{aligned}
SS_T &= \sum (X - \overline{X_{\text{global}}})^2 \\
&= \sum_{i=1}^n \sum_{j=1}^k \sum_{m=1}^r (X_{ijm} - \overline{X})^2 \\
&= \sum_{i=1}^n \sum_{j=1}^k \sum_{m=1}^r (X_{ijm}^2 - 2X_{ijm} \overline{X} + \overline{X}^2) \\
&= \sum_{i=1}^n \sum_{j=1}^k \sum_{m=1}^r (X_{ijm}^2) - \sum_{i=1}^n \sum_{j=1}^k \sum_{m=1}^r (2X_{ijm} \overline{X}) + \sum_{i=1}^n \sum_{j=1}^k \sum_{m=1}^r (\overline{X}^2) \\
&= \sum_{i=1}^n \sum_{j=1}^k \sum_{m=1}^r (X_{ijm}^2) - 2\overline{X} \sum_{i=1}^n \sum_{j=1}^k \sum_{m=1}^r (X_{ijm}) + nkr \overline{X}^2 \\
&= \sum_{i=1}^n \sum_{j=1}^k \sum_{m=1}^r (X_{ijm}^2) - 2nkr \overline{X} \sum_{i=1}^n \frac{1}{n} \sum_{j=1}^k \frac{1}{k} \sum_{m=1}^r \frac{X_{ijm}}{r} + nkr \overline{X}^2 \\
&= \sum_{i=1}^n \sum_{j=1}^k \sum_{m=1}^r (X_{ijm}^2) - 2nkr \overline{X} \sum_{i=1}^n \frac{1}{n} \left(\sum_{j=1}^k \frac{1}{k} \left(\sum_{m=1}^r \frac{X_{ijm}}{r} \right) \right) + nkr \overline{X}^2 \\
&= \sum_{i=1}^n \sum_{j=1}^k \sum_{m=1}^r (X_{ijm}^2) - 2nkr \overline{X} \overline{X} + nkr \overline{X}^2 \\
&= \sum_{i=1}^n \sum_{j=1}^k \sum_{m=1}^r (X_{ijm}^2) - nkr \overline{X}^2
\end{aligned}$$

From the expression of SS_T , we define the correction factor CF as follows:

$$\begin{aligned}
CF &= nkr \overline{X}^2 \\
&= nkr \left(\sum_{i=1}^n \frac{1}{n} \left(\sum_{j=1}^k \frac{1}{k} \left(\sum_{m=1}^r \frac{X_{ijm}}{r} \right) \right) \right)^2 \\
&= \frac{1}{nkr} \left(\sum_{i=1}^n \sum_{j=1}^k \sum_{m=1}^r X_{ijm} \right)^2
\end{aligned} \tag{2.16}$$

Let the [Degrees of Freedom \(DOF\)](#) for the full experiment be f_T . Therefore, the correction factor for the general case is defined as.

$$CF = \frac{(\sum X)^2}{f_T + 1} \quad (2.17)$$

The formulae for squared deviations can be redefined using the correction factor are as follows:

$$SS_T = \sum (X^2) - CF, \quad SS_A = \sum_i^n \frac{(\overline{X_{i..}})^2}{r(f_A + 1)} - CF, \quad SS_B = \sum_j^k \frac{(\overline{X_{.j.}})^2}{r(f_B + 1)} - CF \quad (2.18)$$

Note: Similar to the algebraic simplification for two factors, the correction factor can be derived and proved to be the same for any number of factors.

Degrees of Freedom

The concept of degrees of freedom is hard to define, yet it is fundamental to conducting various statistical analyses. The degrees of freedom in statistical analysis can be described as the number of independent and fair comparisons that can be made in a given dataset [23]. For [ANOVA](#), this translates to the number of unique comparisons that can be made within the group. For instance, let the i^{th} level of a n level factor A be A_i . In [ANOVA](#), a unique comparison within A can be written as $(A_i \text{ and } A_l : i, l \in \{1...n\}, l \neq i)$. Since the total number of unique comparisons is one less than the number of levels of A , the degree of freedom f_A is $n - 1$.

Calculating the degrees of freedom for each factor and the complete dataset, it can be shown that the following equation holds [19]. This property of degrees of freedom can be used to verify the correct assignment to each SS term.

$$f_T = f_A + f_B + f_{A \times B} + f_W \quad (2.19)$$

Mean Sum of Deviations Squares and F-ratio

In the context of [ANOVA](#), *mean square* refers to the variance (variability) within and between groups in an experimental design. It is used to assess the differences in means

among groups and within groups.

$$MS_A = \frac{SS_A}{f_A} \quad (2.20)$$

$$MS_B = \frac{SS_B}{f_B} \quad (2.21)$$

$$MS_{AB} = \frac{SS_{AB}}{f_{AB}} \quad (2.22)$$

$$MS_W = \frac{SS_W}{f_W} \quad (2.23)$$

Using the mean squares of the factors we can perform the F-test to identify the significance of variability (mean squared deviation) due to each factor. The F-test considers the variability due to the factor and the variability due to the experimental error. Once fit in the F-distribution, we find if the variability ratio of factor to the error is statistically significant and generate the p-value from its cumulative distribution function.

$$F_* = \frac{MS_*}{MS_W} \quad (2.24)$$

$$\text{p-Value} = 1 - \text{cdf}(F_*, f_*, f_W) \quad (2.25)$$

2.3.2 ANOVA Results

After the calculations, the [ANOVA](#) results are tabulated in the format shown in Table [2.2](#).

Table 2.2: Format of [ANOVA](#) Results

Source	DF	SS	MS	Percent	F	p
A	f_A	SS_A	MS_A	Percent _A	F_A	
B	f_B	SS_B	MS_B	Percent _B	F_B	
AB	f_{AB}	SS_{AB}	MS_{AB}	Percent _{AB}	F_{AB}	
Error	f_E	SS_E	MS_E	Percent _E		

Chapter 3

Our Approach

In this chapter, we delve into the central aspect of our methodology, aiming to quantify the variability within [SilGeo](#). Building upon the concepts elucidated in Chapter 2, we outline a comprehensive plan for conducting experiments and analysing the captured data. Our discussion involves two evaluation approaches for variability. The first one evaluates [SilGeo](#) as a measurement system. The second involving [Design of Experiments \(DOE\)](#) is a work towards quantifying variability in [SilGeo](#) as an assembly of components.

3.1 Measurement System Analysis

[SilGeo](#) is used as a measurement instrument in a data-driven novelty detection system to sift counterfeit targets from a collection of targets of the same type. It is pointed out in the [Measurement Systems Analysis \(MSA\)](#) Reference Manual by [Automotive Industry Action Group \(AIAG\)](#) [24] that the effectiveness of any data-driven approach depends on the quality of the measurement data. Similarly, the performance of this novelty detection system in sifting counterfeit targets relies on the quality of measurement data produced by [SilGeo](#). This section discusses techniques established in the industry to evaluate the quality of a measurement system.

To better understand the discussion in the section, we must define the following terms in the context of [MSA](#) Reference Manual [24] adopted by [AIAG](#) and Practical Attribute and Variable Measurement Systems Analysis [25] by [American Society for Quality \(ASQ\)](#):

Measurement: It is defined as assigning numeric quantities to tangible items to represent relations among them regarding specific attributes. The process of assigning such numbers is called the measurement process, and the quantity is the measurement value.

Gage: A Gage in the context of a measurement system is the device that is used to perform the measurement process and obtain the measurement value. In this thesis, [SilGeo](#) is the Gage that is used to capture data of hardware fingerprints of the targets.

Measurement System: A measurement system is the collection of gages, applied standards, operating procedures, used fixtures, software, personnel, measuring environment, data format, assumptions and measurement units for the characteristic being measured.

3.1.1 Measurement System Variation

According to the Measurement Systems Analysis Reference Manual [24], the quality of data is characterised based on its *bias*(location error) and *variance*. For a given measurement system, the *bias* is defined as the difference between the obtained average measurement and the expected average measurement. On the other hand, *variance* is defined as the spread in the observed data. It is also stated that high variation is the most common cause of low-quality data [24]. Figure 3.1 serves as a visualisation of the effect of both bias and variance on the data.

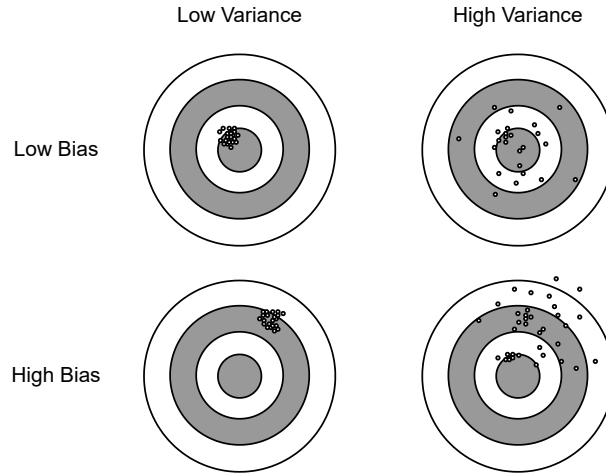


Figure 3.1: Effect of Bias and Variance (Bulls Eye Plot) [1]

When evaluating a measurement system, the following fundamental issues must be addressed:

- **Sensitivity:** The measurement system must possess the capability to provide adequate discrimination for the desired use case. In the case of detecting counterfeit electronics, the measurement system must produce data that reflects the differences in the internal physical structure of the hardware targets.
- **Stability:** This issue deals with the behaviour and consistency of the variation due to the measurement system. Determining whether the variation in the measurement is due to normal causes or a special case/outlier.
- **Statistical Properties:** A measurement system must maintain consistency in statistical properties such as *errors*, *bias* and *variance* over the expected use-case of measurement targets.

Note: As presented in the thesis by Vibien [11], [SilGeo](#) has adequate sensitivity to be applied to detecting counterfeit targets FTDI targets. Furthermore, [SilGeo](#) can also be applied successfully to Z80 microprocessors, as presented by Moreno et al. [12]. In this thesis, we will keep our focus on statistical properties, limiting the analysis to the domain Z80 use case.

Main Categories

The variability due to a measurement system can be classified into two broad categories: width variation and location variation [24].

1. **Location Variation:** Also known as location error, it is represented as bias in the measurement data. The term location variation originates from the fact that due to a change in bias, there is a change in the mean of the data distribution produced by a measurement system. It can further be broken down into the following subcategories:
 - **Accuracy and Bias:** Accuracy as a concept refers to the degree of correctness of the measured value with respect to the reference value. On the other hand, bias is the measure of systematic error; it is the difference between the observed average and the reference value.
There is no notion of a reference value in the detection of counterfeits. To detect counterfeits, a hardware fingerprint is compared to a known legitimate model created from fingerprints of one or more legitimate targets. Therefore, in this case, [SilGeo](#) must only reproduce the measurement(for example, a reference target). Any deviation among instances of [SilGeo](#) would be captured in the assembly variability of the instrument.

- **Stability:** is defined as the drift or total variation in the bias over a long period. In this thesis, we focus on the characteristics of **SilGeo** in the short term.
 - **Linearity:** is defined as the behaviour of bias throughout the operating range of the measurement instrument.
In the case of **SilGeo**, linearity loses the conventional meaning. It can be interpreted as the behaviour of bias for the domain/subset of targets considered valid given a specific use case (for example, the set of any microprocessor labelled as Z80 in a use-case detecting Z80 counterfeits).
2. **Width Variation:** Width is the degree of spread in the data. the term originates from the fact that any change in the variance of the data is a change in the width of the distribution produced by the measurement system.
- **Precision:** is generally defined as the expected variation of repeated measurements.
 - **Repeatability:** is defined as the variation in measurements obtained by one operator using one measurement instrument measuring a particular characteristic of the same target multiple times.
 - **Reproducibility:** is defined as the variation in measurements obtained by more than one operator using one measurement instrument measuring a particular characteristic of the same target.
 - **GRR:** is an estimate of the combined effect of repeatability and reproducibility. One of the focal points of the thesis. By definition, **GRR** is:

$$\sigma_{\text{GRR}}^2 = \sigma_{\text{repeatability}}^2 + \sigma_{\text{reproducibility}}^2 \quad (3.1)$$

- **Sensitivity:** refers to the degree of responsiveness and capability to discern differences in the measured feature.
- **Consistency and Uniformity:** Consistency can be described as the change in measurement variation or repeatability over a long term of operation. Uniformity on the other hand, can be described as the degree of homogeneity of repeatability over the operating domain.

Note: Reproducibility is a significant factor in measurement instruments, especially influenced by the operator's skill. It is either considered null for automated processes or insignificant where the operator is not a major source of variation. In the **MSA** Reference Manual [24], reproducibility is referred to as the average variation between-systems or between-conditions of a measurement.

3.1.2 Key Results

Condensing the properties defining the quality of measurements obtained from a system, the following three metrics are used to evaluate a measurement system fully [24]:

- **Capability:** This is defined as the estimate of variation due to combined measurement errors (both random and systematic) for a short-term assessment. The following components quantify it:
 - Uncorrected bias or linearity
 - Repeatability and Reproducibility (GRR)

The capability of a system can be quantified as:

$$\sigma_{\text{capability}}^2 = \sigma_{\text{bias(linearity)}}^2 + \sigma_{\text{GRR}}^2 \quad (3.2)$$

- **Performance:** is defined as the net effect of all significant and quantifiable sources of variation in readings taken a relatively longer period of time. Performance includes capability in conjunction with the long-term measures of variability:
 - Capability
 - Stability and Consistency

It is quantified as follows:

$$\sigma_{\text{performance}}^2 = \sigma_{\text{capability}}^2 + \sigma_{\text{stability}}^2 + \sigma_{\text{consistency}}^2 \quad (3.3)$$

- **Measurement Uncertainty:** is the single term used to describe the quality of the *measurement value*. The difference between Uncertainty and MSA is that the latter is used to understand the end-to-end measurement process and promote variation reduction in each aspect, whereas the former is concerned with the resulting *measurement value* with respect to the true value.

3.2 Sources of Variability

A crucial aspect of designing an experiment to quantify variability is identifying any possible sources of variability in the system. Conceptually, the source of variability in a measurement system can be divided into two subcategories: noisy and controllable. The idea behind accounting for sources of variability is to identify all controllable sources, isolate their effect on the overall variability, and then figure out as many noisy sources as possible to avoid their effects to the best of our ability.

3.2.1 Cause and Effect Diagram

In [MSA](#), the variability in a system is studied using a Cause and Effect Diagram. The [MSA](#) Reference Manual [24] recommends following the acronym *S.W.I.P.E.* as a guideline to identify the majority of elements responsible for overall variability. *S.W.I.P.E.* stands for Standard, Workpiece, Instrument, Person and Procedure, and Environment.

The sample diagram on Page 17 of The MSA Guide [24] represents a standard cause and effect diagram, also known as the *ishikawa* or the *fishbone* diagram for measurement systems. This type of analysis approach is used to identify and analyse potential causes pertaining to a problem, which in this case is the measurement system variability.

1. The first step is to identify the main issue. The main issue to investigate is the variability in the system due to various causes, to be placed at the head of the diagram.
2. The next step is to identify the main categories of causes. For measurement systems, we use *S.W.I.P.E.* for causes.
3. Third, we identify any subcategories to each main cause by brainstorming, applying domain knowledge, and standards for measurement systems analysis [24].

3.2.2 Sources of Variability in SilGeo

Continuing the discussion on Hardware Fingerprinting using [SilGeo](#) Section 2.2, we recall that [SilGeo](#) employs a stimulus of [RF](#) frequency signals to capture the [S-parameters](#). The signals generated using a [VNA](#) are routed through a multiplexer to a target receptacle on the adapter [PCB](#). Using this information, we discuss the sources of variability for each component of the [SilGeo](#) measurement system. Figure 3.2 shows individual assembly components separated and depicts the details of their interconnection to form the [SilGeo](#) gage.

Variability in VNA and Cable As an independent [RF](#) measurement instrument, a [VNA](#) (Vector Network Analyzer), a test instrument, is used to analyze the electrical behaviour of high-frequency components and networks, typically antennas, filters, and amplifiers. Although as a high-precision instrument, the [VNA](#) used for this study is factory-calibrated and robust to most variability factors. Nevertheless, we list possible sources of variability in the [VNA](#).

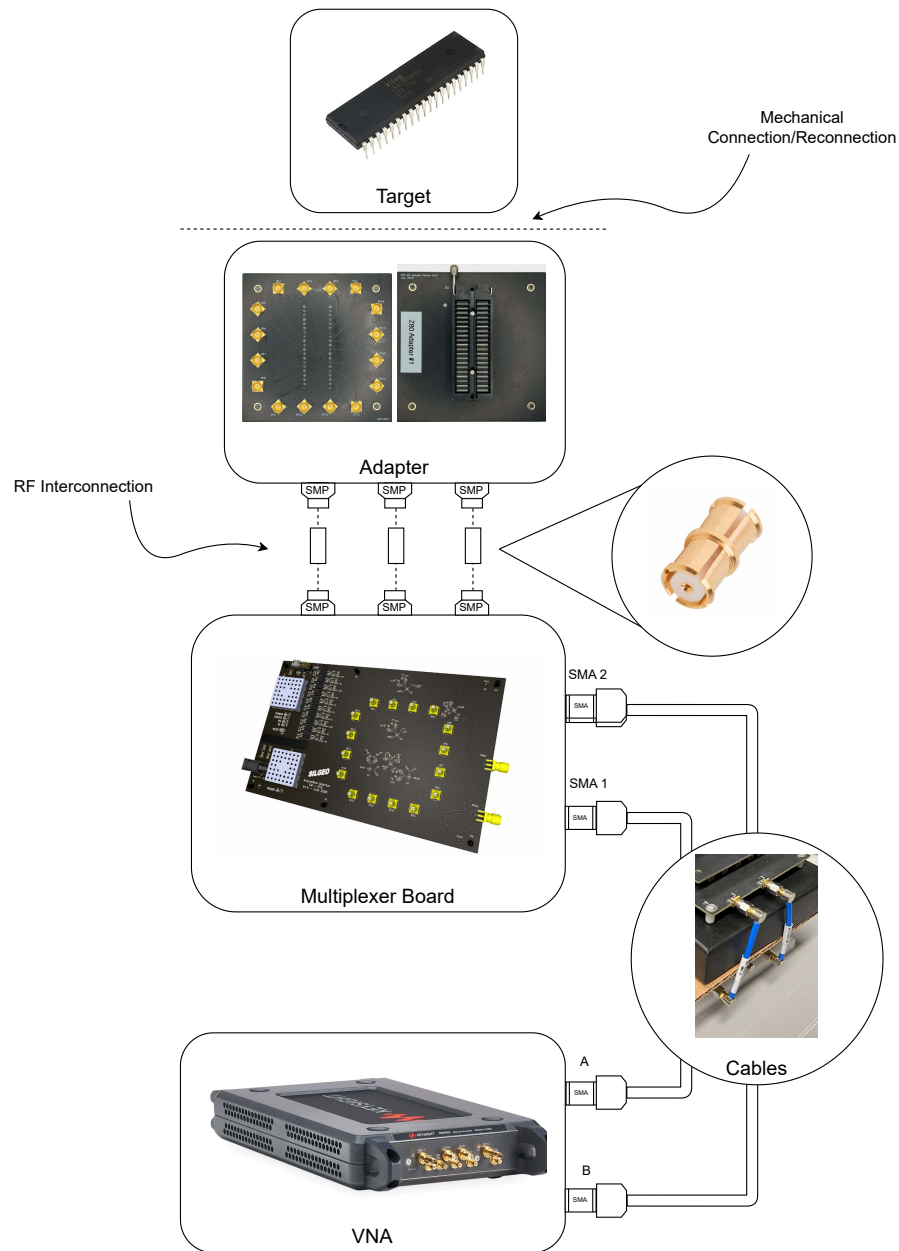


Figure 3.2: [SilGeo](#) Assembly Components Connection Diagram

- **Connection Variability:** The cables and connectors that connect the [VNA](#) to the [DUT](#) can introduce variability due to losses, impedance mismatches, and phase shifts. Moreover, repeated connections and disconnections of test cables can cause connector wear, resulting in degraded performance and variability.
- **Cross-talk:** Improper isolation and shielding in conductors can result in electromagnetic coupling and cause cross-talk, affecting the measurements.
- **Calibration Errors:** Calibrating the [VNA](#) is critical to accurate measurements. Errors in the calibration process, such as incorrect calibration standards, incomplete calibration, or drift in calibration standards, can introduce variability.
- **Frequency Stability:** [VNAs](#) performing high-frequency sweeps with high resolution rely on stable frequency sources. Variations in the frequency source or reference oscillator can lead to frequency inaccuracies and measurement errors.
- **Temperature Drift:** Temperature fluctuations can affect the characteristics of electronic components and cables, leading to measurement variability. To counter this, the [VNA](#) used in the experiments for this thesis involves a warm-up phase to reach operating temperature.
- **Interference and Environmental Factors ([Electromagnetic Interference \(EMI\)](#)):** RF interference from the background, nearby equipment and other environmental factors may introduce noise and affect [VNA](#) measurements.
- **Nonlinearities:** Nonlinear behaviour in the [VNA](#)'s components or the DUT can lead to measurement variability, especially when dealing with high-power signals.
- **Instrument Age:** Long-term wear and tear over time may cause mechanical imperfections in connectivity and general drift in the calibration state, adversely affecting the accuracy of measurements.
- **User Technique:** The technique and skill of the operator using the [VNA](#) can impact measurements, especially for manual operations such as assembling the test setup.

Multiplexer As discussed in Section [2.2](#), the multiplexer board designed by Vibien [\[11\]](#) is a standalone piece of hardware electronics that performs the function of high-frequency signal routing in [SilGeo](#). It consists of a microcontroller to handle communication and process commands, several power supplies, [RF-capable](#) connectors to interface with [VNA](#) and

adapter boards, and high-frequency analog multiplexers. Due to the variety of components, the following factors may contribute to the variability in the multiplexer board:

- **RF Interconnects:** Due to the repeated use of the adapter board mounting, the **SMP** connectors are subjected to wear and tear. With each actuation, small imperfections may be introduced in the connectivity that accumulate over time and cause variability.
- **Mechanical Stresses:** Improper mounting of the board to other components and chassis, subjection to shock and vibrations while handling and assembling the experiment, can lead to permanent deviation.
- **Component Ageing:** Characteristics of passive components and semiconductors can change over time due to ageing, affecting the **RF** circuit's performance.
- **RF:** Losses from reflections, insertion loss, and skew resulting from impedance mismatches, crosstalk, and other factors are suspected to cause variability in the measurements.
- **Power Supply:** any unfiltered noise and ripple in power rails of the analog multiplexers can adversely affect their **RF** performance.
- **Printed Circuit Board:**
 - **Dielectric Variations:** The dielectric constant of PCB materials can vary, affecting the characteristic impedance of transmission lines. Inconsistencies in dielectric properties can lead to impedance mismatches.
 - **Copper Foil Variability:** Variations in the thickness and surface roughness of copper foils used on the PCB can impact electrical performance, especially at high frequencies.
 - **Manufacturing Processes:** Variations in the PCB manufacturing process, including etching, drilling, and plating, can affect the board's electrical properties and introduce variability.
 - **Ground Plane Integrity:** Proper grounding is crucial in **RF** PCBs. Variations in ground plane continuity or the presence of ground loops can introduce unwanted interference.
 - **Solder Joint Quality:** Variations in soldering quality, such as solder bridging, cold solder joints, or excess solder, can impact the electrical connections and **RF** performance.

- Component Placement: Variations in the placement of components, especially in [RF](#) circuits with critical component spacing, can affect signal propagation and [RF](#) coupling.
- Environmental factors:
 - [Electromagnetic Interference \(EMI\)](#): Variations in [EMI](#) shielding and susceptibility can affect the PCB's ability to resist external interference and maintain [RF](#) signal integrity.
 - Temperature Effects: [RF](#) PCBs can be sensitive to temperature changes. Temperature variations can lead to thermal expansion and contraction, affecting the electrical properties of components and PCB traces.
 - Environmental Factors: Exposure to moisture, humidity, and contaminants can lead to variation in the PCB's electrical properties over time.
- Design: Flaws in design can lead to defects such as non-linearity in measurements and signal integrity issues.

Adapter Board The highlight of the adapter board is the target receptacle, but the point of connection of the receptacle is also where the impedance mismatch begins. In addition to the effects the adapter has in common with the multiplexer, the following are additional sources of variability in the adapter:

- Receptacle: The [RF](#) mismatch in the receptacle leads to a loss of signal integrity when transmitting through to the target.
- Operation: Additionally, repeated actuation of the receptacle causes wear and tear, leading to a further mismatch and potential drift.

Target Under Test The variation in target devices is an external source of variability that is not in our control. Few reasons of this variation

- Physical Characteristics and Compatibility: Due to geometrical, electrical or random inconsistencies in insertion, a target may pose significant variation in connecting with the measurement setup.
- Manufacturing: Each type of target, especially semiconductor ICs, has intrinsic variations due to factors such as a large number of manufacturing processes involved, the location of raw materials, contamination, etc.

- **Ageing:** The age of a target can alter its internal physical characteristics. In the case of a semiconductor IC, the die could suffer radiation effects, temperature cycling, dielectric degradation, oxide breakdown, etc. and cause variation in the measurements taken with [SilGeo](#).

Operator In many experiments, precise measurements and data collection are crucial. The operator must follow established procedures to obtain accurate and reliable data. This includes using appropriate instruments, calibrating equipment, and minimising sources of measurement error. We enumerate such concerns as follows:

- **Training and Qualification:** Lack of training for complex and specific technical tasks may cause inconsistencies in the experimental setup, equipment usage, and interpreting results.
- **Standardisation:** Protocols and procedures are designed to ensure consistency and reduce variability. This may include maintaining detailed records of operator activities, including the conditions under which measurements were made, any deviations from the protocol, and other relevant information.
- **Operator Effects:** In some experimental designs, the operator influence on the results is explicitly studied. This is done to assess whether the human factor contributes to the observed variability in the data.
 - **Reproducibility:** For scientific experiments, it is essential that the results can be reproduced by different operators or in different laboratories. Reproducibility helps confirm the validity of the findings.
 - **Operator Bias:** Researchers also consider the possibility of operator bias, where an operator’s personal beliefs or expectations may unintentionally influence the results. Steps are taken to minimise such bias.

Environment Environmental effects are the primary source of noise factors in variability; they include temperature, humidity, vibration, background radiation etc. These factors lead to the ageing of the system, which can appear in the form of oxidation, measurement drift and connectivity mismatches.

We now draw out an Ishikawa cause-effect diagram for [SilGeo](#), summarising the sources of variability impacting the system. Taking the reference of the cause-effect diagram Figure I-B 1 from MSA Reference Manual, Figure [3.3](#) maps the potential cause and effect of variability in [SilGeo](#).

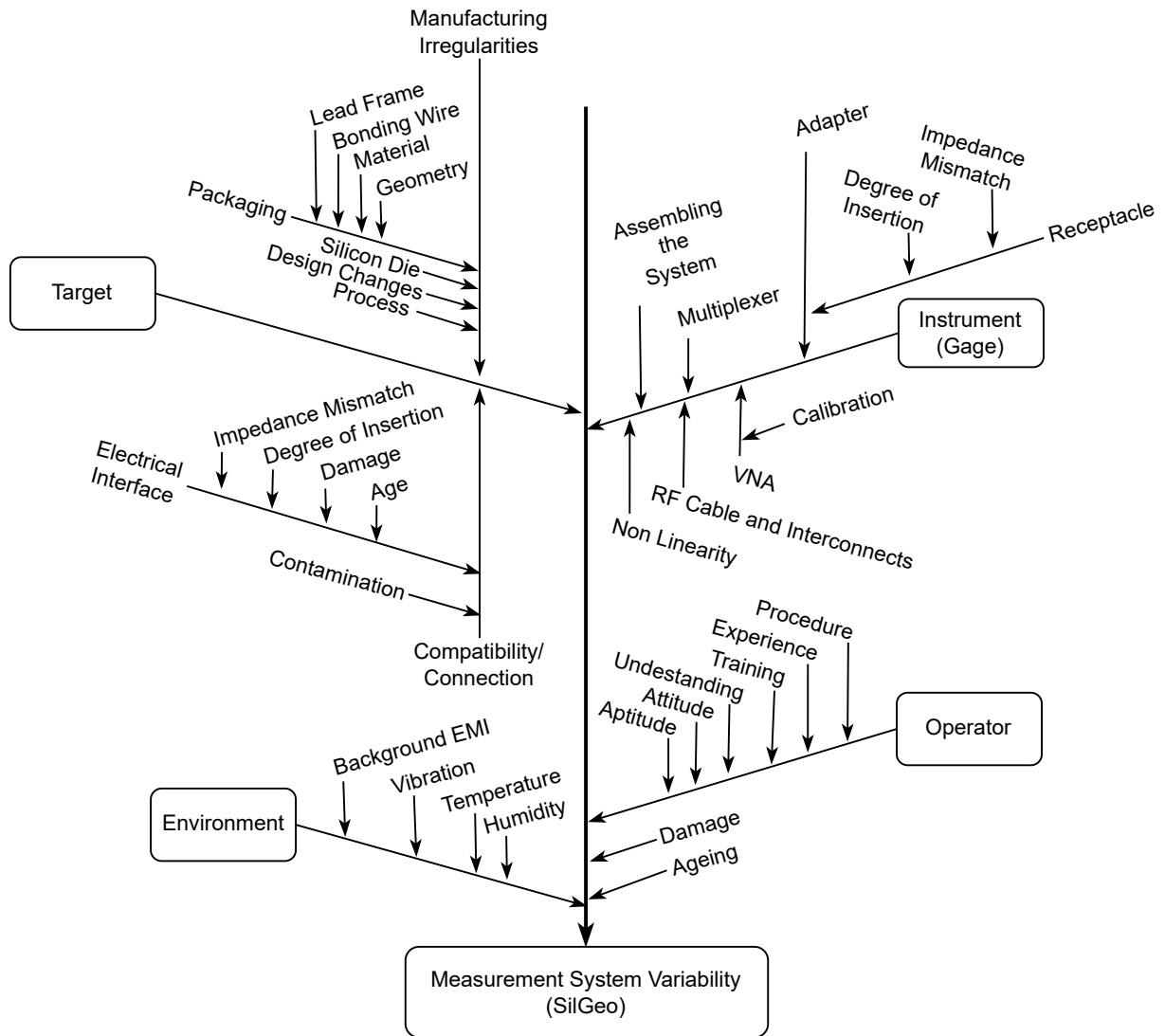


Figure 3.3: SilGeo Variability Cause and Effect Diagram

3.2.3 From SilGeo Perspective

Having identified the possible sources of variability and defined the main focus of a measurement system capability, we now move towards isolating the feasible from the theoretical. In context to [SilGeo](#), considering veritable expectations, we formulate the approach starting with identifying the controllable factors for experimentation. Only a few of the factors discussed in the previous section can be considered controllable. All the remaining factors are placed under the category of noisy factors.

We will plan and perform the variability experiment by considering the following controllable qualitative factors:

- Assembly
 - [VNA](#): Vector Network Analyser
 - [RF](#) Cable
 - Multiplexer
 - Adapter + Receptacle
- Operation
 - Target IC
 - Operator

These variability sources will be under the influence of noisy factors, but using proper experiment design techniques will allow us to place these effects under experimental error. As showcased in Figure [2.4](#) and explained in Section [2.2.3](#), the [VNA](#) is responsible for the stimulus. The cable connects with the [VNA](#) and the multiplexer. The adapter only connects with the multiplexer and accepts the target into the receptacle. Hence as we will see further in this thesis Section [3.5](#), the interactions of importance originate from these connections.

As we discussed in Section [2.2.3](#), the [RF](#) Cables, [VNA](#), multiplexer board and Adapter board PCB, even though are under the influence of wear-tear and noises, they have been designed to respect the [RF](#) signal integrity from the [VNA](#). This fact is not necessarily true for the target receptacle, which is sourced from third-party manufacturers. This is of concern and gives us strong grounds to suspect the influence of the adapter board in the variability to be of the highest significance. For this, as we will see in Section [3.5](#), we

explore more possible variations for it and observe if these concerns are veritable in the analysis Section 4.3.2.

From this point in the thesis, we focus on discussing how to plan, design approaches, and execute them to quantify the overall variability in [SilGeo](#) and the effect of each factor on it.

3.3 Capability Study (Gage R&R)

For the capability assessment, we only consider the variation caused by the measurement gage (repeatability) and the personnel performing the measurement capture (reproducibility). The lack of a standardised measurement reference in [SilGeo](#) eliminates the need for a bias study. Hence, we will perform the R&R study on [SilGeo](#).

3.3.1 Planning GRR

As recommended in [25], the following are the key issues that should be considered in the planning phase of a [GRR](#) study:

- **Number of Appraisers/Operators:** An automatic gage is considered to have a single operator and is considered to have no operator effect on the measurement. With a gage that requires personnel to interact and perform measurements, MSA Reference Manual [24] recommends having a minimum of two operators, preferably three, depending on availability.
- **Number of samples:** MSA Reference Manual [24] recommends using at least 10 samples and a rule of thumb. The recommendation is to select enough samples so that $(\text{number of samples}) \times (\text{number of operators})$ exceeds 15. Table 3.1 serves as a guide to choosing the number of samples, operators, and trials to be performed to achieve a reasonable outcome from the [GRR](#) study within use case-specific constraints.
- **Number of trials:** In this context, a trial is one measurement of all the samples by each operator. Similar to the number of operators, as a bare minimum, an experiment should have two trials, with three or more, to increase statistical soundness. MSA Reference Manual [24] recommends using Table 3.1 as a guide in conjunction with the number of operators and samples.

In the context of [SilGeo](#) and the use case of quantifying the variance, this study is performed on two types of targets of the same part number: one set of 20 known legitimate targets sourced from two trusted vendors (10 each), and a second set of 10 targets known to be physically different (X-ray) and suspected to be counterfeits. The number of operators for this study is 2; therefore, with three trials each, we exceed the minimum recommendation of 14 degrees of freedom to base the estimate of equipment variation by a significant margin.

Table 3.1: Forbidden Region for [GRR](#) (Marked by $\times$)

Num trials (r)	k operators $\times$ n targets												
	4	5	6	7	8	9	10	11	12	13	14	15	≥ 16
2	$\times$	$\times$	$\times$	$\times$	$\times$	$\times$	$\times$	$\times$	$\times$	$\times$	$\times$	$\times$	
3	$\times$	$\times$	$\times$	$\times$									
4	$\times$												

3.3.2 Expectation and Analysis

The detailed process for the complete experiment is described in Section 4.1.2. For the [GRR](#) study, we only consider the Operator and the Part, hence the study is performed separately for each experimental setup. The results from [GRR](#) applied to all experiments are combined and averaged.

The outcome of the [GRR](#) study for any given experimental setup is calculated using the [ANOVA](#) technique, helping us quantify the following [24] in Table 3.2:

Using the results in Table 3.2, the [GRR](#) is calculated as follows:

$$\text{GRR} = \sqrt{\tau^2 + \gamma^2 + \omega^2} \quad (3.4)$$

Additionally, we use the following to quantify the percentage contribution to the total variability due to each component:

$$\text{Total Variation (TV)} = \sqrt{\text{GRR}^2 + \text{PV}^2} \quad (3.5)$$

$$\% \text{ of Total Variation} = 100 \left(\frac{\sigma_{(\text{component})}}{\sigma_{(\text{total})}} \right) \quad (3.6)$$

$$\% \text{ Contribution to Total Variance} = 100 \left(\frac{\sigma_{(\text{component})}^2}{\sigma_{(\text{total})}^2} \right) \quad (3.7)$$

Table 3.2: Estimate of Variance Components

Variance Component	Estimate
Equipment (EV)	$\tau^2 = 2MS_{\epsilon}$
Interaction (INT)	$\gamma^2 = \frac{MS_{OP} - MS_{\epsilon}}{r}$
Operator (OV)	$\omega^2 = \frac{MS_O - MS_{OP}}{nr}$
Part (PV)	$\sigma^2 = \frac{MS_P - MS_{OP}}{kr}$

3.4 Design of Experiments

[Design of Experiments \(DOE\)](#), a powerful and systematic approach first introduced by Ronald Fischer in the 1920s [2], is used in various fields to optimize processes, solve problems, and gain deeper insights into complex systems. [DOE](#) is a means to find the relationship between various factors (variables) and responses to them. It is, therefore, important to express their relationship precisely and efficiently by an equation. If the interactions between variables (factors) are significant, it is important to include the terms of interactions in the equation.

The conventional approach to experimentation is often referred to as The [One-Variable-At-A-time \(OVAT\)](#) or *trial-and-error* approach. It involves changing one factor or variable at a time while keeping all others constant. This approach has several limitations when compared to the [DOE](#) [23]:

- **Inefficiency:** [OVAT](#) requires examining the impact of a single factor while assuming all others remain constant, which may not reflect real-world conditions. Moreover, it involves a large number of experiments to understand the effects of multiple factors and their interactions. Making it relatively resource-intensive and time-consuming.
- **Missed Interactions:** [OVAT](#) often fails to capture interactions between variables. In reality, factors often interact with each other, meaning the combined effect of two factors is not simply the sum of their individual effects. It may identify the impact of individual factors but does not provide a comprehensive understanding of how these factors work together to affect the outcome.

- Risk of Sub-optimal Solutions: Since [OVAT](#) does not consider the combined effects of factors. Adjusting one factor to optimize a response may inadvertently worsen the response when other factors are taken into account.

On the other hand, the [DOE](#) approach is a systematic and statistically rigorous method for experimentation. It offers several advantages over the conventional approach by being more efficient and requiring fewer experiments to analyse the effects of multiple factors and interactions by efficiently exploring the design space. It also allows the application of optimisation solutions by providing comprehensive insights.

At its core, [DOE](#) involves planning and conducting experiments to uncover the relationships between input factors (variables) and the output responses (outcomes) of a system. For [SilGeo](#), we use [DOE](#) to plan experiments, capture data and quantify the variability and contribution of several controllable factors, as discussed in the previous sections. In standard literature, experimental problems can be categorised into the following [4]:

- Comparing Factors: The primary objective is to assess multiple factors and identify the most significant ones. In [SilGeo](#), the goal is to determine if there are variations in the data capture of the same target using different setups. If differences exist, it's essential to discern the nature of these distinctions and identify the main cause.
- Variable Selection: When a system involves several variables, but only a subset of them is significant, a screening experiment is employed to pinpoint these crucial variables. It minimizes the use of degrees of freedom in estimating error variance and higher-order interactions. Once the pivotal variables are identified, a subsequent experiment can be undertaken to conduct a more comprehensive investigation of their effects.
- Exploring the Response Surface: is the next step after identifying the main effects. It is essential to investigate how the main factors influence the response. Examining the relationship between the reaction and these variables is often referred to as exploring the response surface.
- System Optimization: When the primary focus revolves around enhancing the efficiency of a system or manufacturing process, the response surface becomes a valuable tool for optimisation. An intelligent and sequential application of [DOE](#) can guide the experiment towards the optimal variable settings.
- Enhancing System Robustness: An essential aspect of quality improvement is ensuring the system remains robust in the face of noise factors. This robustness is typically

achieved by selecting control factor settings that render the system less sensitive to such variations. These experiments systematically induce noise factors and assess the variance among the noise replicates for a specific controllable factor setting.

3.4.1 Basic Principles

Nature of Factors In DOE, factors can be categorized into two main types based on their nature: qualitative (categorical) factors and quantitative (numeric) factors [23]. These two types of factors have distinct characteristics and are treated differently in both design and analysis phases of the experiment.

- **Qualitative Factors:** Also known as categorical factors or discrete factors, are variables that have distinct categories or levels without any inherent numerical value. They represent attributes, characteristics, or categories that are not measured on a numeric scale, for example, types of material, colour, etc. Qualitative factors are typically manipulated by assigning levels to different categories. In the case of SilGeo, the variation due to factors such as the VNA, Mux, Cables, and Adapter is qualitative in nature.
- **Quantitative Factors:** These are variables that have numeric values and can be measured on a continuous or numeric scale. These factors represent quantities that can be measured or quantified, for example, temperature (°C), time taken, pressure, length, etc. Quantitative factors are manipulated by changing their numeric values within a specified range. In the case of SilGeo, there are no quantitative factors.

In the context of DOE, the choice of whether a factor is qualitative or quantitative depends on the nature of the factor and its relevance to the experiment. Analyzing qualitative and quantitative factors typically involves different statistical techniques and experimental designs. For example, when conducting a DOE with qualitative factors, methods like ANOVA for categorical data or chi-square tests are used. For quantitative factors, multiple regression analysis or ANOVA for continuous data.

Treatment In experimental design, when there are multiple factors involved, a *treatment* typically refers to a specific combination of levels or conditions from those multiple factors. In other words, a treatment is a unique set of factor levels that represents a particular experimental condition.

Experimental Unit An experimental unit is a generic term that refers to a basic unit, such as material, animal, person, machine, or time period, to which a treatment is applied [4]. It is important to note that an experimental unit is a term assigned to an instance of the main subject of the experimental study. Its definition can drastically change depending on the type of study.

For example, in the case of a medical drug study, the experimental unit may be the white mice, and each unit is applied to certain treatments (levels of the drug). On the other hand, this definition changes when we consider a GRR study, where the experimental unit is defined as the target/sample, which is being measured or evaluated to assess the repeatability and reproducibility of a measurement system.

In this thesis, as discussed, our focus is bifurcated into evaluating SilGeo and its R&R as a measurement system. In both of these cases, the experimental unit is the target or the DUT. In GRR, the experimental units help evaluate variability due to target, operator, and equipment. For the Overall variability, the same experimental unit helps evaluate the effect of varying each factor as discussed in Section 3.2.3.

Design Principles To improve the efficiency of experimentation, three fundamental principles need to be considered in the DOE: replication, randomisation, and blocking [23, 4].

- Replication: It is defined as the application of treatments to multiple experimental units. The term *repetition* is different from *replication*. The former refers to multiple measurements from the same experimental unit, whereas the latter defines involving multiple readings from different experimental units. Replication increases the reliability and validity of experimental results by obtaining a more precise estimate of the effect due to factors and interactions.
- Randomization: is the concept by which the effect of experimental bias is reduced. In the conventional sense, it is defined as the random allocation of experimental units to treatment groups. However, it can be extrapolated into randomizing the order of performing the experiments and capturing target data.
- Blocking: is a method of eliminating the effects of batch-to-batch, day-to-day, in other words, block-to-block variations. It is hard to reproduce the same experimental conditions, operator temperament, and environmental factors when a significant time has passed. The idea is to arrange similar experimental runs into groups and complete the run within the block.

Design Resolution (Higher is Better) It is defined as the degree to which the main effects are aliased with the interaction effects. Resolution is R if not p -factor is aliased or confounded with another effect containing less than $R - p$ factors [23]. This definition will become clear in the fractional factorial experiment part of Section 3.4.2. The resolution is represented with Roman numerals.

- Resolution III Designs: No main effects are confounded with each other. The main effects are confounded with second-order interactions, and second-order interactions may be confounded with each other.
- Resolution IV Designs: No main effects are confounded with each other and second-order interactions. Second-order effects may be confounded with each other.
- Resolution V Designs: No main effects are confounded with each other or second and third-order interactions. Second-order interactions are confounded with third-order interactions.

3.4.2 Design Approaches

DOE encompasses various approaches and strategies for planning and conducting experiments to efficiently and effectively investigate the effects of multiple factors on a response variable. Different design approaches are chosen based on the specific objectives of the study, the number of factors to be studied, resource constraints, and the nature of the problem.

Conventional Approach

This section discusses the most plain designs and defines relevant terminology that we will use in characterizing the advanced approaches.

Full Factorial Experiments Many experiments involve the study of the effects of two or more factors. When all possible combinations of factor levels are tested, it's called a full factorial design. Each factor has one or more levels, representing the different settings or values of that factor. For example, if temperature is a factor, its levels might be low, medium, and high.

Design Space/Response Surface: In a full factorial design, if there are k factors, each with 2 levels, a full factorial experiment would involve all 2^k possible combinations of levels. The experimental space grows exponentially in the number of levels.

Matrix Notation: A full factorial design is often represented using matrix notation, where each row in the matrix corresponds to an experimental run, and each column corresponds to a factor level. The entries in the matrix indicate the factor levels for each run.

Interactions: Full factorial designs allow you to assess not only the main effects of each factor but also the interactions between factors. Interactions occur when the combined effect of two or more factors differs from the sum of their individual effects.

Advantages

- **Comprehensive Understanding:** Full factorial designs provide a complete understanding of how each factor and interaction affects the response variable. This is valuable for gaining insights into complex systems.
- **Statistical Analysis:** The data from full factorial experiments can be analyzed using statistical methods, such as [ANOVA](#), to quantify the significance of factor effects and interactions.
- **Efficient for Smaller Designs:** For experiments with a relatively small number of factors or levels, a full factorial design can be practical and insightful.

Challenges

- **Resource-Intensive:** Conducting experiments for all possible combinations can be resource-intensive, especially when there are many factors or levels.
- **Large Experiments:** As the number of factors and levels increases, the number of experimental runs grows exponentially, making the design impractical for large systems.
- **Limited by Resources:** In practice, researchers often resort to fractional factorial designs or other approaches when resource constraints prevent testing all combinations.

In an ideal case, one would always prefer to perform a full factorial experiment. However, it is not always feasible due to time and resource constraints. In that case, the approach of fractional factorial experimentation is used.

Fractional Factorial Experiments These designs involve testing a subset of the full factorial combinations while still allowing for the estimation of main effects and some interactions. Fractional factorial designs or other specialized designs are used to reduce the number of experimental runs while still providing valuable information about the effects of factors.

Factorial Notation: Fractional factorial experiments are often represented using exponential notation. For example, a $2^{(k-p)}$ design represents a fractional factorial experiment with k (2 level) factors where only a fraction $1/2^p$ of the full factorial runs are conducted. The value of p determines the fraction of runs included.

Resolution: The resolution of a fractional factorial design indicates the highest order of interactions that can be estimated with the design. Commonly used fractional factorial designs include $2^{(k-p)}$ designs with resolutions III, IV, or higher. A higher resolution indicates the ability to estimate more interactions, but it often comes at the cost of a larger number of runs.

Main Effects and Interactions: An effect is defined as the application of Factors and their interactions in the experimental array. Each column is referred to as an effect. Fractional factorial experiments focus on estimating main effects (the effects of individual factors) and some interactions (the effects of combinations of factors). Typically, lower-order interactions are prioritized (e.g., two-factor interactions), while higher-order interactions may not be estimable in designs with lower resolutions.

Half-Fractional Designs: A special case of fractional factorial designs is the half-fractional design, which tests only half of the possible combinations. This is common in $2^{(k-1)}$ designs and can be useful when a full 2^k experiment is too large. Half-fractional designs allow for the estimation of main effects and two-factor interactions.

Fundamental Principles of Factorial Effects According to Wu et al. [4], three fundamental principles are used to justify the development of fractional factorial experiments.

- **Effect Hierarchy Principle:** This principle suggests that when resources are scarce, priority should be given to the estimation of lower-order effects (interactions).
 - Lower-order effects are more likely to be important than higher-order effects.
 - Effects of the same order are equally likely to be important.
- **Effect Sparsity Principle:** The number of relatively important effects in a factorial experiment is small. As stated by Wu et al. [4], it is also referred to as the *Pareto*

Table 3.3: $2^{(4-1)}$ Factorial Experiment

S.No.	Factors						
	A	B	AB	C	AC	BC	D = ABC
1	–	–	+	–	+	+	–
2	+	–	–	–	–	+	+
3	–	+	–	–	+	–	+
4	+	+	+	–	–	–	–
5	–	–	+	+	–	–	+
6	+	–	–	+	+	–	–
7	–	+	–	+	–	+	–
8	+	+	+	+	+	+	+

Principle in Experimental Design, following a Pareto distribution, where the most contribution is only due to the vital factors.

- Effect Heredity Principle: In order for an interaction to be significant, at least one of its parent main effects should be significant.

Defining Relation To explain the concept of confounding and aliasing, let us consider a 2^3 experiment and convert it into a $2^{(4-1)}$ half-fractional design. In the experiment, each factor has 1 DOF, and so does each interaction. For two-level multi-factorial experiments, columns 1, 2, and 4 are generator columns for which each level is represented by sign – or +, and the interactions can be represented by performing scalar multiplication of signs ($- \times - = +$, $- \times + = -$, and $+ \times + = +$). Table 3.3 represents the full factorial experiment of factors A , B , and, C with interactions [23].

From the fundamental principles, we can assume here that third-order interactions are not significant. Hence, we get to introduce a new main factor, D . By introducing a new main factor, we have effectively avoided using the 2^4 full factorial space and are now working with the half fraction of that. The result of this introduction is that the effects of factor D are now indistinguishable from the effect of ABC . The factor D is now considered aliased or confounded with $A \cdot B \cdot C$ interaction.

$$D = ABC \quad \text{or} \quad I = ABCD$$

where I represents the identity column or the column of all +

Table 3.4: Alias Table for a Half-Fractional 4-Factor Experiment

Effect	Alias
A	BCD
B	ACD
C	ABD
D	ABC
AB	CD
AC	BD
BC	AD
AD	BC
BD	AC
CD	AB
ABC	D
ABD	C
ACD	B
BCD	A

Here, the relation $I = ABCD$ is the defining relation. By using this relation, we can find out all the aliased factors in the experiment. In addition to that, by definition, the resolution of the fractional factorial is IV since the defining relation is a word of length 4 [4]. Following the defining relation, the calculated aliasing pattern is shown in Table 3.4

Taguchi Approach

Taguchi's approach to DOE is focused on robust product design and quality improvement. These methods aim to find factor settings that minimise the variation in the response variable, making the process or product less sensitive to variations in factors or external conditions.

Taguchi Orthogonal Arrays The experimental array in the Taguchi approach is a fractional factorial array, but the choice of combinations is derived by combining orthogonal Latin Squares [26]. These orthogonal arrays for 2 level factors as defined by Taguchi, are denoted by $L_k(2^{k-1}) : k \in \{4, 8, 12, 16, 32, 64\}$ [2]. Table 3.5 is the L_8 Taguchi array, with 3 factors and their interactions.

Table 3.5: L_8 Taguchi Orthogonal Array [2]

Run	1	2	3	4	5	6	7
1	1	1	1	1	1	1	1
2	1	1	1	2	2	2	2
3	1	2	2	1	1	2	2
4	1	2	2	2	2	1	1
5	2	1	2	1	2	1	2
6	2	1	2	2	1	2	1
7	2	2	1	1	2	2	1
8	2	2	1	2	1	1	2
Factor Effects	a	b	a	c	a	b	a
			b		c	c	b
							c

Experimental Characteristics Before the design process, the following questions must be answered:

- Number of Factors: This refers to the number of main controllable factors that will be varied in the experiment.
- Number of Levels: Assuming the number of levels for each factor is the same, this number is the number of variations in the factor.
Note: In Taguchi [Orthogonal Array \(OA\)](#)s, the levels/elements are generally represented by symbols $(1, 2, \dots, s)$. Whereas, in traditional [OAs](#), the representation is $(0, 1, 2, \dots, s - 1)$, s being the max level number.
- Importance of Interaction: In most cases, the factors affecting the experimental unit are non-interacting. However, if there are interacting factors or their interaction is unknown, deciding the interactions to be considered is crucial as it impacts the effective number of factors.

Linear Graphs Linear Graphs are an important part of Taguchi's approach to experiment design. Linear graphs are visual representations of triangular tables or matrices that

depict the assignment of interactions between all columns of an array. They serve as a useful tool to simplify the often complex assignment of factors and interactions, particularly when dealing with factors having different numbers of levels or complicated experiments, like pseudo factor design, to an orthogonal array [2].

A linear graph assigns factors to points and interactions to connecting segments. Figure 3.4a and 3.4b are examples of linear graph interpretations that encode relations represented by L_8 Taguchi OA.

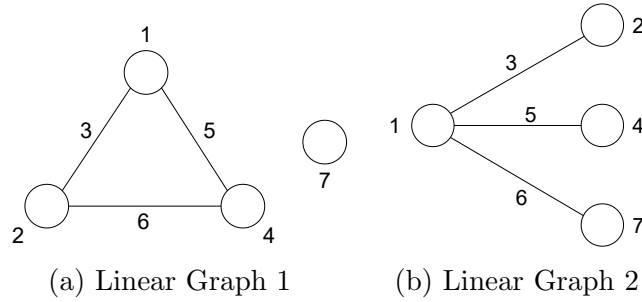


Figure 3.4: The L_8 Linear Graphs [2]

The linear graph in Figure 3.4a indicates that the interaction between column 1 and column 2 is represented by column 3. Similarly, the interaction between column 1 and column 4 by column 5. The third-order interaction column 7 is represented as an independent point outside of the triangle. On the other hand, Figure 3.4b is used in cases where the requirement dictates the study of the interaction between one main factor and three factors.

Another way to represent the interacting columns in the experimental array is to refer to the two-level triangular table for OAs created by Taguchi, to aid the experiment designer. In a triangle table, each row and column represents the column number in the experimental array. The corresponding cell value is the column number where the interaction is represented in the array.

As we can see in Table 3.6 and 3.5, the column 3 (C_3) represents the interaction of column 1 (C_1) and 2 (C_2). The L_8 array forms the following defining relation: (Where a $\cdot$ is representing interaction)

$$C_3 = C_1 \cdot C_2, \quad C_5 = C_4 \cdot C_1, \quad C_6 = C_2 \cdot C_4, \quad C_7 = C_1 \cdot C_2 \cdot C_3$$

$$I = C_1 \cdot C_2 \cdot C_3 = C_4 \cdot C_1 \cdot C_5 = C_2 \cdot C_4 \cdot C_6 = C_1 \cdot C_2 \cdot C_3 \cdot C_7$$

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
(1)	3	2	5	4	7	6	9	8	11	10	13	12	15	14
	(2)	1	6	7	4	5	10	11	8	9	14	15	12	13
		(3)	7	6	5	4	11	10	9	8	15	24	13	12
			(4)	1	2	3	12	13	14	15	8	9	10	11
				(5)	3	2	13	12	15	14	9	8	11	10
					(6)	1	14	15	12	13	10	11	8	9
						(7)	15	14	13	12	11	10	9	8
							(8)	1	2	3	4	5	6	7
								(9)	3	2	5	4	7	6
									(10)	1	6	7	4	5
										(11)	7	6	5	4
											(12)	1	2	3
												(13)	3	2
													(14)	1

Table 3.6: Triangular Interaction Table for Two-Level Orthogonal Arrays [2]

Applying the Method After the factors and level decisions have been made, the overview of the steps of design using the Taguchi approach is as follows [26]. Let the experiment have 5 two-level factors (A, B, C, D, E) with two interactions ($A \cdot C$) and ($B \cdot C$).

- **Degrees of Freedom (DOF):** Each factor has two levels; thus, the degree of freedom for each individual factor is 1 ($f_* = \text{number of levels} - 1$). The degree of freedom for the interactions is also 1 ($f_{A \times C} = f_A \times f_C$). Since the total degree of freedom is the sum of all constituents (Equation 2.19), the value comes out to be 7 in this case.
- **Selection of a Taguchi Orthogonal Array:** The DOF of the appropriate Taguchi array is greater than equal to the DOF of the experiment. In the example case, the minimum sized Taguchi OA is L_8 with a degree of freedom of exactly 7. The choice of the array is highly dependent on any potential interaction investigation.
- **Column Assignment:** As suggested in [26], in designing experiments with interactions, it is necessary first to identify which columns should be reserved for studying interactions. This is done by identifying interactions from the triangular table, as shown in Table 3.6 or one of the representative linear graphs. The key is to leave free columns for each of the main factors. For the current example.
 - Factor C is common in both interactions $A \cdot C$ and $B \cdot C$. Assign C to column 2 (Figure 3.4a), intersecting vertex for columns 1 and 4.

- Assign A and B to columns 1 and 4 respectively.
- The remaining columns can be assigned to the factors D and E .
- Increasing number of Levels: There are standard OAs available with mixed factorial design. However, they may not always be the most economical or suitable for a particular task. In such cases, Taguchi recommends modifying the standard arrays to fit the specific requirements. As we will detail in Section 3.4.3, a two-level column can be easily converted to a four or eight-level column, and a four-level column can be upgraded to an eight-level column, and so on. This approach can also be applied to Taguchi OAs.
- Decreasing number of levels: If the requirement is to use an odd levelled factor, it is possible to decrease levels using the dummy treatment. Following is an overview of the steps involved; more can be learned from the book *A Primer on The Taguchi Method* [26]:
 - Choose an appropriate OA, based on the number of levels in factors.
 - Assign all the factors to the OA.
 - Choose any one level for a factor whose levels we want to reduce.
 - Replace that level in the experimental array with another level of the same factor. Making it a dummy level.

3.4.3 Orthogonal Arrays

It is well-known in the literature that the Taguchi DOE approach is derived from conventional factorial designs and orthogonal Latin squares [26, 27]. As Kacker et al. [27] present in their paper, the Taguchi catalogue can include generic orthogonal arrays, and thereby expanding the flexibility in experiment design.

An orthogonal array (more specifically, a fixed-element orthogonal array) of t strength denoted by $\text{OA}(N, s_1^{m_1} \dots s_\gamma^{m_\gamma}, t)$ is an $N \times m$ matrix ($m = \sum_{i=1}^{\gamma} m_i$) whose columns s_i symbols such that for any t columns all combinations of symbols appear with the same frequency. In the context of DOE, this property is called the orthogonality of the matrix [4, 27].

For $\gamma = 1$, the arrays are referred to as symmetrical OAs, and for any value of $\gamma > 1$ are called asymmetrical or mixed-level OAs. The resolution of a 2^{k-p} or 3^{k-p} design is the same as the strength of the design t plus 1 when viewed as an orthogonal array.

Generator Columns: In an orthogonal array of type $OA(N, 2^k)$, the generator columns represent the full factorial space, and the generated columns are the interaction effects while maintaining the orthogonality of the array.

To demonstrate the processes followed in the manipulation and application of OAs in the experiment design, we take the example of a 15 DOF array $OA(16, 2^{15})$, represented in Table 3.7.

Table 3.7: $OA(16, 2^{15})$ Orthogonal Array

	Col Numbers														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Compos	1	2	12	3	13	23	123	4	14	24	124	34	134	234	1234
S.N.	1	2	12	3	13	23	123	4	14	24	124	34	134	234	1234
1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
2	-1	-1	-1	-1	-1	-1	-1	1	1	1	1	1	1	1	1
3	-1	-1	-1	1	1	1	1	-1	-1	-1	-1	1	1	1	1
4	-1	-1	-1	1	1	1	1	1	1	1	1	-1	-1	-1	-1
5	-1	1	1	-1	-1	1	1	-1	-1	1	1	-1	-1	1	1
6	-1	1	1	-1	-1	1	1	1	1	-1	-1	1	1	-1	-1
7	-1	1	1	1	1	-1	-1	-1	-1	1	1	1	1	-1	-1
8	-1	1	1	1	1	-1	-1	1	1	-1	-1	-1	-1	1	1
9	1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	1
10	1	-1	1	-1	1	-1	1	1	-1	1	-1	1	-1	1	-1
11	1	-1	1	1	-1	1	-1	-1	1	-1	1	1	-1	1	-1
12	1	-1	1	1	-1	1	-1	1	-1	1	-1	-1	1	-1	1
13	1	1	-1	-1	1	1	-1	-1	1	1	-1	-1	1	1	-1
14	1	1	-1	-1	1	1	-1	1	-1	-1	1	1	-1	-1	1
15	1	1	-1	1	-1	-1	1	-1	1	1	-1	1	-1	-1	1
16	1	1	-1	1	-1	-1	1	1	-1	-1	1	-1	1	1	-1

Mix Factor Experiments

In this thesis, we will make use of one four-level column and three two-level columns for the main factors. In this section, we discuss the standard approach to increasing the number of levels in an orthogonal array to form a mixed-factor experiment. This approach, when combined with dummy treatment, leads to a high level of flexibility in experiment design.

In two-level OAs, column triplets of the form $(\alpha, \beta, \alpha\beta)$ can be interpreted as mutually exclusive orthogonal sub-arrays of two factors and their interaction, and a seven-column

combination of the form $(\alpha, \beta, \alpha\beta, \gamma, \alpha\gamma, \beta\gamma, \alpha\beta\gamma)$ can be interpreted as similarly, and so on. These columns represented by the groups are mutually exclusive because an orthogonal array remains an orthogonal array when one of its columns is replaced with an orthogonal array whose rows have a one-to-one correspondence with the elements of the replaced column, thus maintaining the array's orthogonality property. Based on this, the following steps perform a level increase from two-level to four-level:

1. First step is to choose a mutually exclusive triplet of the form $(\alpha, \beta, \alpha\beta)$. Table 3.8 offers a reference to select mutually exclusive columns for level increase.

k	Sets of Columns
3	(1, 2, 12)
4	(1, 2, 12), (3, 4, 34), (13, 24, 1234), (23, 124, 134), (123, 14, 234)
5	(1, 2, 12), (4, 5, 45), (234, 1235, 145), (1234, 135, 245), (134, 15, 345), (14, 25, 1245), (124, 235, 1345), (24, 35, 2345), (34, 125, 12345)
6	(1, 2, 12), (3, 4, 34), (13, 24, 1234), (23, 124, 134), (123, 14, 234), (5, 6, 56), (15, 26, 1256), (125, 16, 256), (25, 126, 156), (35, 46, 3456), (345, 36, 456), (45, 346, 356), (135, 246, 123456), (12345, 136, 2456), (245, 12346, 1356), (235, 1246, 13456), (1345, 236, 12456), (1245, 1346, 2356), (1235, 146, 23456), (2345, 1236, 1456), (145, 2346, 12356)

Table 3.8: Mutually Exclusive Sets of Column Triplets in the 2^k Designs [4]

2. Replace the three columns with a new column by the following rules in Table 3.9.

Table 3.9: Level Increase

α	β	$\alpha\beta$	New Factor
-1	-1	-1	0
-1	1	1	1
1	-1	-1	2
1	1	1	3

3. Repeat steps for all required level increases. (A similar process can be followed for level increases to 8.)

By successive replacements arrays $OA(16, 2^{12}4^1)$, $OA(16, 2^94^2)$, $OA(16, 2^64^3)$, $OA(16, 2^34^4)$, $OA(16, 4^5)$ can be created from $OA(16, 2^{15})$. It is important to note that the total DOF

in the array are unchanged, and the selection of an OA can be done by solely working with the DOF of the base array. Table 3.10 is the outcome of applying this technique to $OA(16, 2^{15})$ creating $OA(16, 2^{12}4^1)$.

Table 3.10: $OA(16, 2^{12}4^1)$ Orthogonal Array

S.N.	Col Numbers												
	New	4	5	6	7	8	9	10	11	12	13	14	15
1	0	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
2	0	-1	-1	-1	-1	1	1	1	1	1	1	1	1
3	0	1	1	1	1	-1	-1	-1	-1	1	1	1	1
4	0	1	1	1	1	1	1	1	1	-1	-1	-1	-1
5	1	-1	-1	1	1	-1	-1	1	1	-1	-1	1	1
6	1	-1	-1	1	1	1	1	-1	-1	1	1	-1	-1
7	1	1	1	-1	-1	-1	-1	1	1	1	1	-1	-1
8	1	1	1	-1	-1	1	1	-1	-1	-1	-1	1	1
9	2	-1	1	-1	1	-1	1	-1	1	-1	1	-1	1
10	2	-1	1	-1	1	1	-1	1	-1	1	-1	1	-1
11	2	1	-1	1	-1	-1	1	-1	1	1	-1	1	-1
12	2	1	-1	1	-1	1	-1	1	-1	-1	1	-1	1
13	3	-1	1	1	-1	-1	1	1	-1	-1	1	1	-1
14	3	-1	1	1	-1	1	-1	-1	1	1	-1	-1	1
15	3	1	-1	-1	1	-1	1	1	-1	1	-1	-1	1
16	3	1	-1	-1	1	1	-1	-1	1	-1	1	1	-1

Mix Factor Interactions

The next logical step is to decompose the degrees of freedom in each factor into contrast relations. These contrasting relations define the interactions between the experimental factors. The contrast relation for mixed-level factor interaction can be identified using the *linear-quadratic system* presented by Wu et al. [4].

The need for defining contrast mechanisms is due to the complexity introduced in interactions between non-binary (more than two levels) factors with differing levels. Interactions between two-level factors A and B have $DOF = 1$, therefore, the interaction $A \times B = AB$. However, the interactions involving factors with three or more levels have a $DOF > 1$, resulting in $A \times B \neq AB$.

For the purposes of this thesis, we limit our discussion to the mixed-level interaction effects of four and two-level *qualitative* factors. In the following technique, factor A has four levels and factor B has two levels:

Linear-quadratic system is used to parameterise the interaction effects of a quantitative factor into effects classified as linear, quadratic, cubic and so on, depending on the degrees of freedom [4].

The interaction of A with four levels and B with two levels with a $DOF = 3$ is decomposed into linear, quadratic and cubic interactions, each having a $DOF = 1$. From Table 3.9, we can derive the contrasts of factor A , with weights of -1 or $+1$ since A is a qualitative factor.

$$\begin{aligned} A_1 &= (-1, -1, 1, 1) \\ A_2 &= (1, -1, -1, 1) \\ A_3 &= (-1, 1, -1, 1) \end{aligned} \tag{3.8}$$

The contrasts (A_1, A_2, A_3) , each with $DOF = 1$ follow $(\alpha, \alpha\beta, \beta)$, with A_2 representing the quadratic effect, and A_1, A_3 are highly correlated to the linear and cubic effects respectively [4].

Considering the factor B with two levels ($DOF = 1$), the interaction effects can be represented as follows:

$$\begin{aligned} A \times B : (AB)_{ij}(x, y) &= A_i(x) \cdot B_j(y) \\ \text{Levels} : x &\in \{0, 1, 2, 3\}, y \in \{0, 1\} \\ \text{Contrast} : i &\in \{1, 2, 3\}, j \in \{1\} \end{aligned} \tag{3.9}$$

Using this relation, we have the following contrasts in our orthogonal array:

$$\begin{aligned} (AB)_{11} &: \text{linear-by-linear interaction effect} \\ (AB)_{21} &: \text{quadratic-by-linear interaction effect} \\ (AB)_{31} &: \text{cubic-by-linear interaction effect} \end{aligned}$$

Note: The interaction effects between two factors with an equivalent number of levels (> 2) are calculated by using modulus arithmetic when the number of levels is a prime number, and Galois Field arithmetic in other cases [27].

3.5 Design of Experiments for SilGeo

Continuing our discussion from Section 3.2.3, we aim to vary the assembly factors of SilGeo. To reiterate, the following factors form the SilGeo assembly: (Note that each factor will be interchangeably referred to by the letter assigned from this point forward in the thesis)

- VNA (C)
- Cable (D)
- Multiplexer (B)
- Adapter & Receptacle (A)

The design and assembly imply that VNA is factory calibrated, and the multiplexer and cables are manufactured with impedance matching and RF compliance. Therefore, we in our study, assume that the most variability in measurements can be attributed to the Adapter and Receptacle combination, where the receptacle is not impedance-matched and involves mechanical actuation to connect with the target.

Choice of Approach: For the study in this thesis, we adopt the OA approach. As described by Kacker et al. [27], the OA approach is intuitive to understand and is practically similar to Taguchi. Using the OA approach, we can easily present the defining relations, increase the number of levels, identify mixed-factor interactions and, most importantly, avoid referring to reference tables defined in the Taguchi approach. The second reason to use the OA approach is that it is generic and can be applied to Taguchi and all other cases. In contrast, the Taguchi approach results from empirical analysis done in an industrial environment to improve the quality of the manufacturing process [26].

Number of Levels

There are four factors under consideration for the experiment design. A 2^4 full factorial experiment results in a total of 16 configurations. Referring to the assembly block diagram in Figure 2.4, and applying domain knowledge with respect to the operation of SilGeo, we claim the following:

- VNA is directly or indirectly connected to each assembly component.

- Cable is connected to the [VNA](#) and the multiplexer
- Adapter is connected to the Multiplexer.

Each factor thus has a main effect and interacts with other factors that affect it. Figure 3.5 represents the important interactions in the [SilGeo](#) setup inspired by the Linear Graph representation in the Taguchi Approach.

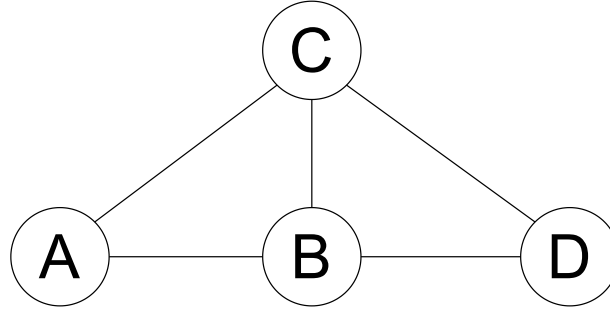


Figure 3.5: Factors and Interactions

The experiment has the following factors and interactions: A , B , C , D , $A \times B$, $A \times C$, $B \times C$, $B \times D$, $C \times D$. With two levels for each factor, the factors and interactions account for $DOF = 9$ degrees of freedom. The experimental array of a 2^4 full factorial experiment accounts for $DOF = 15$ with all factors (first), second and third-order interactions between the factors. Since the relevant interactions in this case are second-order, we have a 6 degrees of freedom flexibility.

To take advantage of this, we increase the number of levels in factor A (Adapter Board) from two to four. The following Table 3.11 represents the allocation of factors in the experiment:

Orthogonal Array Approach

This approach is derived from the analysis strategy for $2^m 4^n$ experiments presented by Wu et al. [4].

Level Increase: The first step in designing the variability experiment for [SilGeo](#) is to follow the steps described in Section 3.4.3 and increase the levels corresponding to one factor in a $OA(16, 2^{15})$ array and generate a $OA(16, 2^{12} 4^1)$ array as represented in Table 3.10.

Component	Factors and Interactions	DOF
Adapter	A	3
Multiplexer	B	1
VNA	C	1
Cable	D	1
	$A \times B$	3
	$A \times C$	3
	$B \times C$	1
	$B \times D$	1
	$C \times D$	1
Total		15

Table 3.11: Degrees of Freedom Allocations

Factor Assignment: After a level increase, the next step is to assign the main factors to the columns corresponding to the remaining generator columns (1, 2, 3, 4) in the OA $(16, 2^{12}4^1)$ array. This leaves the main factor D without assignment.

- Assign factor A the new level increased column, such that $A = (A_1, A_2, A_3) = (1, 12, 2)$.
- Assign factors with direct interaction to A a generator column. Therefore, the assigned combos are $B = 3$ and $C = 4$, corresponding to column numbers 4 and 8, respectively.

Interaction: The next step is to follow the steps described in Section 3.4.3 and generate the interaction columns for linear, quadratic and cubic interaction for factor A with factors B and C , and linear interactions of the remaining as allocated in the Table 3.11.

Interaction Column Assignment: The generated interactions are in the form of combinations of generator columns of the OA $(16, 2^{15})$ array. They are mapped to the column numbers and combos as per Table 3.7. The following are the corresponding combos:

$$\begin{aligned}
A &= (A_1, A_2, A_3) = (1, 12, 2), & B &= 3, & C &= 4, & BC &= 34 \\
A_1B &= 13, & A_2B &= 123, & A_3B &= 23, \\
A_1C &= 14, & A_2C &= 124, & A_3C &= 24
\end{aligned} \tag{3.10}$$

Defining Relation and Minimum Aberration: To consider the assignment of factor D, we discuss the concept of *minimum aberration* designs. The defining relations for a $2^m 4^1$ experiment can be categorised into two types.

- Type 0: Involving only two-level factors.
- Type 1: Involving one $DOF = 1$ component of a four-level factor with other two-level factors.

Due to the level increase $I = A_1 A_2 A_3$, any combination of two A_i 's can be replaced with the third. As explained by Wu et al. [4], aliasing in defining relations of type 1 is less serious due to higher degrees of freedom. Moreover, one of the A_i can be represented as a cubic contrast effect in interaction terms, making it less significant than the other two. Let for a design d , let $A_{i0}(d)$ and $A_{i1}(d)$ be the number of type 0 and type 1 words, respectively having length i , and $A_i(d) = \{A_{i0}(d), A_{i1}(d)\}$ is the word-length pattern of d .

Using this information, we define the minimum aberration experiment as: *For any two $2^m 4^1$ designs d_1 and d_2 of the same run size and r be the smallest integer such that $A_r(d_1) \neq A_r(d_2)$. If $A_{r0}(d_1) < A_{r0}(d_2)$ or $A_{r0}(d_1) = A_{r0}(d_2)$, $A_{r1}(d_1) < A_{r1}(d_2)$, then d_1 is said to have less aberration than d_2 . A design d has minimum aberration if no other design has less aberration than d .*

We are mainly interested in the main effects and only a few second-order interactions, so a resolution of III or IV suffices. Table 3.12 lists all possible configurations of designs and their word patterns.

Resolution	D	Defining Relation	BD	CD	A_{30}	A_{31}	A_{41}
III	13	$I = A_1 BD$	1	134	0	1	0
III	23	$I = A_3 BD$	2	234	0	1	0
III	123	$I = A_2 BD$	12	1234	0	1	0
III	14	$I = A_1 CD$	134	1	0	1	0
III	24	$I = A_3 CD$	234	2	0	1	0
III	124	$I = A_2 CD$	1234	12	0	1	0
III	34	$I = BCD$	4	3	1	0	0
IV	134	$I = A_1 BCD$	14	13	0	0	1
IV	234	$I = A_3 BCD$	24	23	0	0	1
IV	1234	$I = A_2 BCD$	124	123	0	0	1

Table 3.12: Defining Relations and Designs

By applying the definition of minimum aberration designs to Table 3.12, we conclude that any minimum aberration design has at least one A_i of factor A confounded with a second-order interaction. Therefore, we choose one of the minimum aberration configurations. The following is the chosen configuration, and Table 3.13 represents the designed experiment array:

$$D = 23, \quad B = 3, \quad BD = 2, \quad CD = 234, \quad I = A_3BD \quad (3.11)$$

	A_1	A_3	A_2	A	B	C	D	A_1B	A_2B	A_3B	A_1C	A_2C	A_3C	BC	CD	BD
Col No.	1	2	3		4	8	15	5	7	6	9	11	10	12	14	11
Combo	1	2	12		3	4	23	13	123	23	14	124	24	34	234	2
1	-1	-1	1	0	-1	-1	1	1	-1	1	1	-1	1	1	-1	-1
2	-1	-1	1	0	-1	1	1	1	-1	1	-1	1	-1	-1	1	-1
3	-1	-1	1	0	1	-1	-1	-1	1	-1	1	-1	1	-1	1	-1
4	-1	-1	1	0	1	1	-1	-1	1	-1	-1	1	-1	1	-1	-1
5	-1	1	-1	1	-1	-1	-1	1	1	-1	1	1	-1	1	1	1
6	-1	1	-1	1	-1	1	-1	1	1	-1	-1	-1	1	-1	-1	1
7	-1	1	-1	1	1	-1	1	-1	-1	1	1	1	-1	-1	-1	1
8	-1	1	-1	1	1	1	1	-1	-1	1	-1	-1	1	1	1	1
9	1	-1	-1	2	-1	-1	1	-1	1	1	-1	1	1	1	-1	-1
10	1	-1	-1	2	-1	1	1	-1	1	1	1	-1	-1	-1	1	-1
11	1	-1	-1	2	1	-1	-1	1	-1	-1	-1	1	1	-1	1	-1
12	1	-1	-1	2	1	1	-1	1	-1	-1	1	-1	-1	1	-1	-1
13	1	1	1	3	-1	-1	-1	-1	-1	-1	-1	-1	-1	1	1	1
14	1	1	1	3	-1	1	-1	-1	-1	-1	1	1	1	-1	-1	1
15	1	1	1	3	1	-1	1	1	1	1	-1	-1	-1	-1	-1	1
16	1	1	1	3	1	1	1	1	1	1	1	1	1	1	1	1

Table 3.13: Designed Experimental Array

Chapter 4

Case Study

Until this point in the thesis, we have explored relevant literature and established the groundwork for a variability study. In this chapter, we execute the plan to perform a variability study of [SilGeo](#). The primary objective of this study is to quantify the variability of [SilGeo](#). The secondary objective is to form an independent target feasibility evaluation mechanism and determine how effective the method is for the particular target.

Note: The term *variation* in technical representations refers to the *standard deviation*, and the term *variability* refers to the *variance*.

4.1 Experiment

As discussed in Section [3.5](#), the experimental array we follow in performing the experiment is shown in Table [4.1](#). This section discusses how we use this array to perform experiments and any considerations in the experimental setup and protocol.

4.1.1 Setup

The measurement system shown in Figure [4.1](#) comprises of the [SilGeo](#) assembly, the targets, the data acquisition system, and the operator.

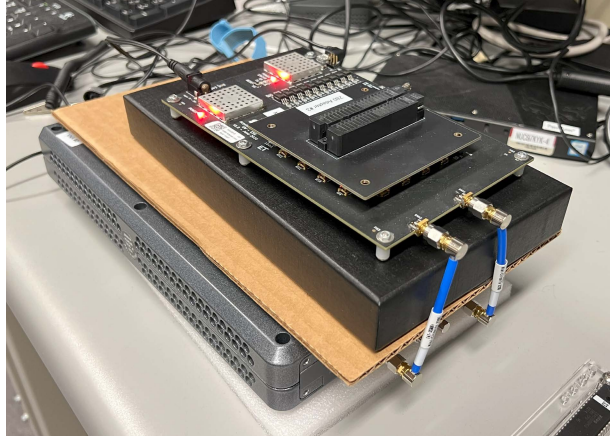


Figure 4.1: SilGeo Assembly

Targets

The experimental units in this study are the hardware targets of one category. The category is defined by the part number and the physical characteristics of the target. In this case study, we focus on *Z84C0008PEG* (referred to as Z80) targets. All the targets have the following properties:

- All nominal targets have been sourced from a reputable supplier.
- The target's internal geometry (excluding the silicon Die) has been visually inspected by us using X-ray imaging (Figure 4.2). We have deemed the internal die-to-pin connections of nominal targets as visually identical.
- Targets have the part number *Z84C0008PEG* printed on the package.
- The targets are in the standard 40-pin DIP type package.
- The pin pitch and target width of all targets are the same.
- Counterfeit targets have been sourced from untrusted sources and inspected with X-ray, confirming the dissimilarity from the nominal targets (Figure 4.3).

SilGeo Assembly

From the discussion in Section 2.2, we know that SilGeo is built from four main components, namely VNA, Cables, Multiplexer Board, and Adapter Board.

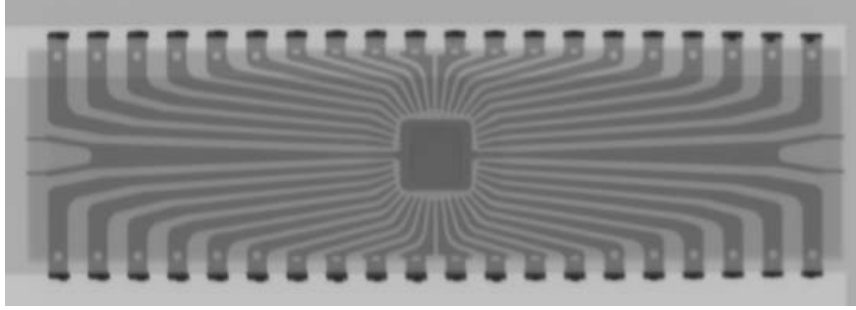


Figure 4.2: Z80 IC X-ray

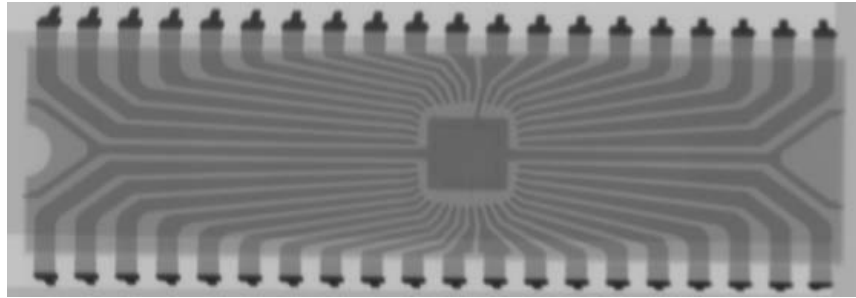


Figure 4.3: Counterfeit Z80 IC X-ray

VNA (Vector Network Analyser): From the discussion in the Section **refer the msa section**, we infer that it is the best strategy to use a calibrated high-quality instrument in variability studies. Therefore, to evaluate the variability in [SilGeo](#), we use the two-port factory-calibrated [P9370A Keysight Streamline USB Vector Network Analyzer, 4.5 GHz](#) [3] (showcased in Figure 4.4). We use two VNAs of this type representing the two levels for the factor of [VNA](#). The following are the setup parameters used in the experiment:

- Stimulus:
 - Frequency Sweep: Linear 300 kHz - 4.5 GHz.
 - Power: -5 dBm.
 - Sampling points: 1601.
 - IFBW: 10 kHz.
- Capture:
 - Trigger: Single (Manual).



Figure 4.4: P9370A Keysight Streamline USB Vector Network Analyzer, 4.5 GHz [3]

- **S-parameters:** S_{00} , S_{01} , S_{10} , S_{11} .
- Time (per sweep): 172 ms.

Cables: For connecting the **VNA** with the multiplexer board, we use two **RF-capable** coaxial cables. Each pair is labelled to correspond to the level of the two-level factor in the experimental array and a second label to match it with the **VNA** port number. To keep the setup compact, we use the Mini-circuits **141-3SMR+** [28] coaxial **RF** cable with the following specifications:

- Connector A: SMA, Male, 90°.
- Connector B: SMA, Male, 90°.
- Length: 3 inch.
- Impedance: 50 Ω .
- Maximum Frequency: 18 GHz.

Multiplexer: We use the multiplexer v1.0 designed by Vibien et al in June 2020 [11]. We used two multiplexer boards, each marked corresponding to the two levels in the factor in the experimental array.

Adapter: We use the ZIF 40 **DIP** adapter board v2.0 designed and manufactured as part of this work. The receptacle used is the 40-pin ZIF socket by Aries Electronics (PN: 40-6554-10). similar to the other components, we use four adapter boards, one for each of the four levels in the experimental array.

Environmental Setup: The environmental setup made for the experiment involved keeping a stable room temperature of 23 °C.

Workstation PC: For interfacing with the [VNA](#) and running the acquisition controller script for capturing data we use an Intel NUC with Windows 10 installed and the following additional software:

- Keysight IO Libraries Suite 2023.
- Keysight PXI/USB [VNA](#) Firmware Release A.15.75.26.

Communication Cables:

1. Two USB 3.0 type-C cables are used to connect the Keysight [VNA](#) with the workstation PC to communicate with the acquisition controller.
2. Two micro-USB cables are used to connect the multiplexer board with the workstation PC to communicate with the acquisition controller.

Tools:

- Torque Wrench.
- Tweezers.

ESD Protection: For ESD protection, we use a hand ESD strap connected to ground.

Data Acquisition

We performed the data acquisition programmatically by using a *C#* program. The program communicates with the Keysight chassis windows driver to configure the setup, trigger the sweep and capture the data in the [Touchstone Format](#) [29]. We also use the script to perform Serial communication (UART) with the Multiplexer board and send commands to switch between port combinations. A sample screenshot of the acquisition controller operating with the interactive [VNA](#) software is shown in Figure 4.5.

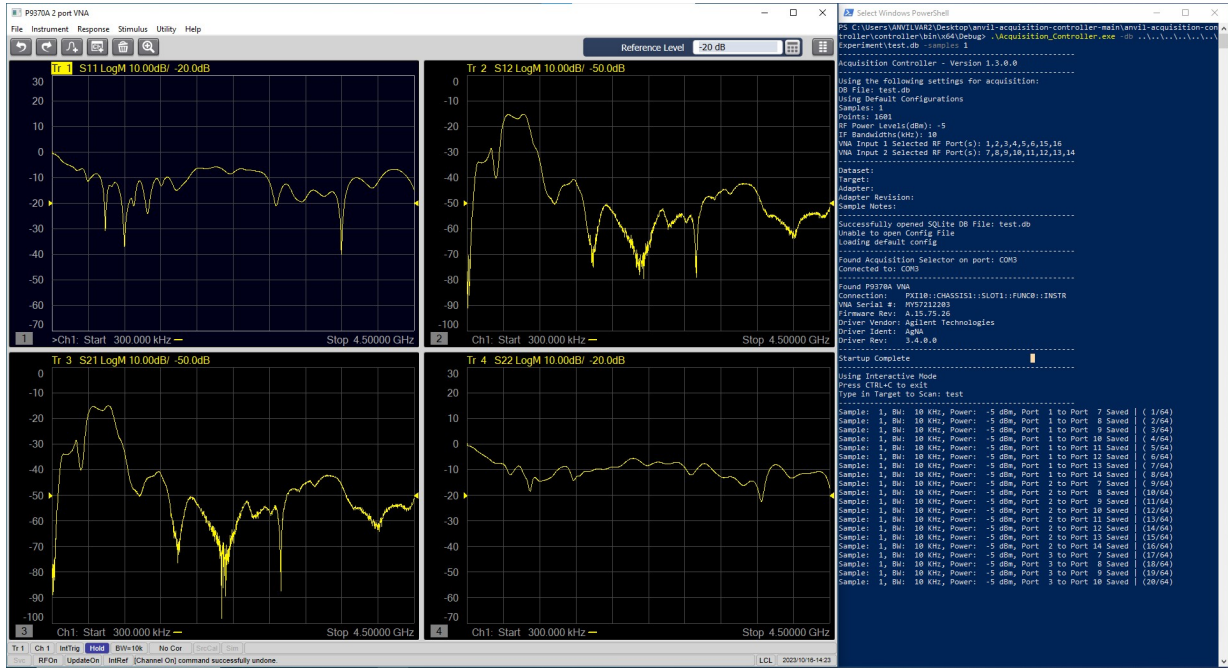


Figure 4.5: Acquisition Controller and Data Capture Screen

4.1.2 Test Protocol

For defining the experiment procedure, we will derive from the experimental array designed in Section 3.5. The new experimental array is converted to levels of type $\{1, 2, 3, 4, \dots\}$, as marked on the physical components involved in the experiment. Table 4.1 shows the experimental array followed in this study.

	A_1	A_3	A_2	A	B	C	D	A_1B	A_2B	A_3B	A_1C	A_2C	A_3C	BC	CD	BD
Col No.	1	2	3		4	8	15	5	7	6	9	11	10	12	14	11
Combo	1	2	12		3	4	23	13	123	23	14	124	24	34	234	2
1	1	1	2	1	1	1	2	2	1	2	2	1	2	2	1	1
2	1	1	2	1	1	2	2	2	1	2	1	2	1	1	2	1
3	1	1	2	1	2	1	1	1	2	1	2	1	2	1	2	1
4	1	1	2	1	2	2	1	1	2	1	1	2	1	2	1	1
5	1	2	1	2	1	1	1	2	2	1	2	2	1	2	2	2
6	1	2	1	2	1	2	1	2	2	1	1	1	2	1	1	2
7	1	2	1	2	2	1	2	1	1	2	2	2	1	1	1	2
8	1	2	1	2	2	2	2	1	1	2	1	1	2	2	2	2
9	2	1	1	3	1	1	2	1	2	2	1	2	2	2	1	1
10	2	1	1	3	1	2	2	1	2	2	2	1	1	1	2	1
11	2	1	1	3	2	1	1	2	1	1	1	2	2	1	2	1
12	2	1	1	3	2	2	1	2	1	1	2	1	1	2	1	1
13	2	2	2	4	1	1	1	1	1	1	1	1	1	2	2	2
14	2	2	2	4	1	2	1	1	1	1	2	2	2	1	1	2
15	2	2	2	4	2	1	2	2	2	2	1	1	1	1	1	2
16	2	2	2	4	2	2	2	2	2	2	2	2	2	2	2	2

Table 4.1: Experimental Array

The configuration is formed by using the qualitative labels assigned to factors $ABCD$. Since there are two VNAs involved in this evaluation, to progress efficiently, we perform two experiments simultaneously with two workstations in the manner shown in Table 4.2.

Block Number	VNA-1	VNA-2
1	1112	3221
2	3211	1122
3	1211	3122
4	3112	1221
5	2111	4222
6	4212	2121
7	2212	4121
8	4111	2222

Table 4.2: Experiment Order

Preparation

Since the full study requires the capturing data with each setup configuration in the experimental array, the assembly will be performed before the data capture. The following steps must be taken to perform the [SilGeo](#) assembly:

1. Gather one of each assembly part [VNA](#), Cable Group, and Multiplexer corresponding to the current experimental configuration and the markings.
2. Place the [VNA](#) on the Table and partially fasten the Connector A on both cables in the group with their corresponding [VNA](#) port numbers.
3. Place the multiplexer board on top of the [VNA](#) and partially fasten Connector B on both cables to the corresponding SMA connectors of the multiplexer (Note: [VNA](#) Port1 SMA connects to Multiplexer SW1 SMA on the PCB).
4. After partial fastening of the cable, align the two such that they are in a flat resting position with equal mechanical stress on both ports.
5. Use the torque wrench to fasten the SMA connectors to torque spec.
6. Identify the correct alignment of the adapter with respect to the multiplexer and place it on the multiplexer board.
7. Apply even pressure on opposite corners until the adapter board completely clicks in and connects to the AMP connectors on the multiplexer.
8. Connect the communication cables for the multiplexer and the [VNA](#) to the workstation PC.
9. Power up the setup and let it run for at least 60 minutes to reach a stable operating temperature for the [VNA](#) as per the datasheet [\[3\]](#).

Data Capturing Process

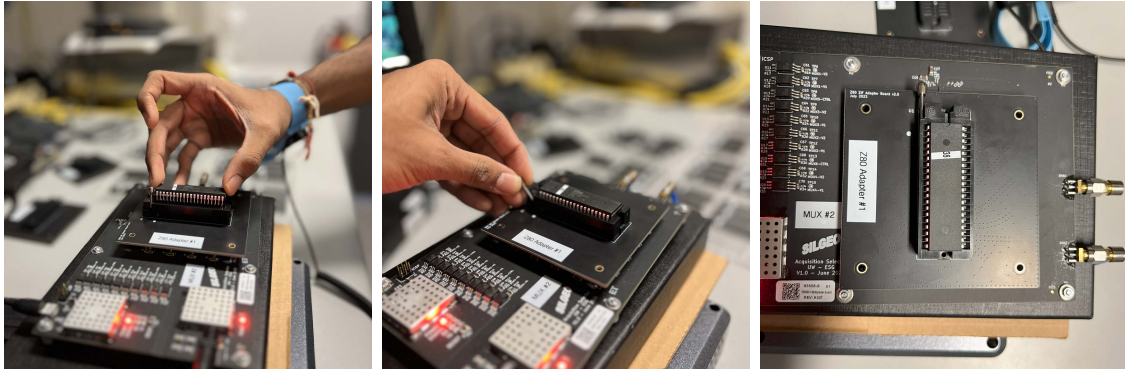


Figure 4.6: Target Insertion Process

After the Assembly is complete, the acquisition controller script is executed with parameters and three repetitions for each measurement. Each sweep takes approximately two seconds, therefore, making the total capture time for one target to be about two minutes (given the presence of 64 port combinations). Figure 4.6 shows the operator inserting the target in the setup. The following steps are performed to capture the data for the targets:

1. Run the acquisition automation script with the parameters discussed in the experimental setup. any communication errors will be posted by the script.
2. Attach the anti-static strap to the wrist.
3. Un-latch the receptacle on the adapter board and manually place the target in the receptacle.
4. Ensure that the target is correctly seated and latch the receptacle.
5. Enter the target ID in the acquisition prompt and press enter.
6. After approximately two minutes, the program will indicate that the capture of all 64 port combinations is complete.
7. Repeat steps 3-6 for all targets.
8. Exit the script by either a keyboard interrupt or closing the terminal.

4.1.3 Operator

In the context of experiments, the term operator refers to the individual or entity responsible for executing the experimental plan or protocol as designed. This experiment involves performing physical tasks, measurements and manipulating equipment. The operator's role is critical in ensuring the experiment is conducted accurately, consistently, and by established protocol. Operator-related factors can significantly impact the quality and reliability of experimental results, as discussed in Section 3.2. We adhere to two operators experimenting independently following the recommendations of the MSA Reference Manual [24] and Gage R&R Guide [25]:

4.2 Distance Metric

One of the main requirements for an ANOVA analysis is obtaining a scalar value corresponding to each measurement. In this section, we discuss how we use the concept of Mahalanobis Space to convert a multi-dimensional measurement vector into a scalar value.

Defining The Measurement

As we have discussed, the multiplexer board performs the action of switching between different port combinations and routing the stimulus signal from the VNA to different ports on the Adapter. Although the data is captured for multiple port combinations, we make the decision to obtain the results from the analysis separately for each port combination. The main reasons for this approach are:

- A measurement made from one port combination is incompatible for comparison with a different port combination.
 - Each port combination causes a physically different connection point on the target.
 - The physical routing of the stimulus signal is different in each case.
- The data must be concatenated to consider all port combinations together. This exacerbates the problem of high dimensionality and causes the covariance matrix to become non-invertible. Keeping them separate, on the other hand, makes the analysis simpler.

- Different port combinations may be affected differently by the factors and produce different magnitudes of variability. Combining the data may mask this non-linear behaviour among port combinations.

Therefore, we only consider one of the [S-parameters](#) of one port combination at a given moment for our data analysis.

Distance Calculation

Due to the nature of this data, there is merit to considering the correlation in the data. This makes the [Mahalanobis Distance \(MD\)](#) the most appropriate choice for the situation. The univariate data [MD](#) is intuitively understood as the number of standard deviations from the mean. For multivariate data, on the other hand, this is not the case, but regardless, [MD](#) encodes the covariance properties in the data. The formula for [MD](#) in multivariate data is given by the following equation [30].

$$MD_k = \sqrt{(X_k - \bar{X})^T \Sigma^{-1} (X_k - \bar{X})} \quad (4.1)$$

where,

X_k is the column vector of the k^{th} observation in the data with n observations $k \in [1, n]$.

Σ is the covariance matrix.

$\bar{X}$ is a column vector for the mean of the data used to calculate the covariance matrix.

We construct a Mahalanobis space, which is the collection of [Mahalanobis Distances \(MDs\)](#) calculated using a covariance matrix. Our experiment involves two separate analyses with independent goals. Hence, we calculate the [MD](#) in two different contexts. The following process is followed for [MD](#) for the Variability Analysis:

1. First, we isolate the data for one port combination containing four [S-parameters](#) for each observation.
2. Second, we only consider the data captured for nominal (Type-1) targets.
3. Next, we strip away the phase data and retain the log magnitude of the [S-parameters](#).
4. We calculate the average of the data for all [S-parameters](#) respectively (this includes all targets). This will act as the mean in calculating [MD](#).

5. For each of the [S-parameters](#), we calculate the covariance matrix by using numpy (a Python package), which utilizes the following formula [\[31\]](#):

$$\Sigma_{X_i X_j} = \text{cov}[X_i, X_j] = E[(X_i - E[X_i])(X_j - E[X_j])] \quad (4.2)$$

where,

X is a multidimensional column vector of random variables. $X = (X_1, \dots, X_n)^T$

Σ is the covariance matrix.

(i, j) is the i^{th} row and j^{th} column entry in the covariance matrix Σ .

$E[X]$ is the expectation operator on random variable X , resulting in the expected value (mean).

6. The number of observations made in the experiment significantly outnumbers the dimensionality of the data. Therefore, the inverse of the covariance matrix is non-singular and is calculated using the numpy's inverse method.
7. Finally, the complete dataset, including the Type-2 targets, are plugged into Equation [4.1](#) to obtain the [MD](#) for each observation.
8. This process is repeated for each port combination.

For the [GRR](#) analysis, the same process is followed, with one additional factor of an operator. Two operators are considered, and the data from both is considered in constructing the Mahalanobis space.

4.3 Observations and Results

After applying the distance metric, we obtain the measurements for each of the [S-parameters](#) as a scalar value. With this data, we apply [ANOVA](#). To explain and demonstrate the steps taken in the analysis, we limit the presentation to port combination (6, 14) (VCC and GND of the target, respectively).

In this section, we present the analysis of [GRR](#) and Assembly Variability studies on the Z80 data.

4.3.1 Gage R&R

We consider the data captured according to the experimental configuration 1111 in the experimental array for the [GRR](#) analysis, where each digit corresponds to the factor level of A, B, C , and D , respectively.

To calculate the reproducibility and repeatability of the experimental setup, we apply the [ANOVA](#) analysis as follows. The Operator is represented by O and the Target is represented by P (Part).

- Calculate the Sum of Squares using the following formulae:

$$SS_T = \sum_{i=1}^n \sum_{j=1}^k \sum_{m=1}^r (X_{ijm} - \bar{X})^2 \quad (4.3)$$

$$SS_P = \sum_{i=1}^n kr (\bar{X}_{i..} - \bar{X})^2 \quad (4.4)$$

$$SS_O = \sum_{j=1}^k nr (\bar{X}_{.j.} - \bar{X})^2 \quad (4.5)$$

$$SS_e = SS_W = \sum_{i=1}^n \sum_{j=1}^k \sum_{m=1}^r (X_{ijm} - \bar{X}_{ij})^2 \quad (4.6)$$

$$SS_{O \times P} = SS_T - SS_O - SS_P - SS_e \quad (4.7)$$

- Calculate the mean squared values for the Operator, Target and interaction.

$$MS_* = SS_*/f_e \quad (4.8)$$

- Calculate the F-ratio with the mean squared error and the p-value as per the F-distribution CDF.

$$F_* = MS_*/MS_e \quad (4.9)$$

$$p_* = 1 - F_*(F_*, f_*, f_e) \quad (4.10)$$

- Assign a significance to the p-values.

p-value	Significance
> 0.1	
> 0.05	.
> 0.01	*
> 0.001	**
others	***

Table 4.3: Two-Way ANOVA Table: Operational Variability (S00)

Source	DF	SS	MS	F	p	
Target	19	37.04	1.95	5.80	8.98e-09	***
Operator	1	12.77	12.77	38.03	2.68e-08	***
Target * Operator	19	6.38	0.34	1.99	0.0182	*
Error/Residuals	80	13.52	0.17			
Total	119	69.72				

Table 4.4: Two-Way ANOVA Table: Operational Variability (S01)

Source	DF	SS	MS	F	p	
Target	19	14.30	0.75	4.14	3.66e-06	***
Operator	1	0.46	0.46	2.53	0.115	
Target * Operator	19	3.45	0.18	0.76	0.75	
Error/Residuals	80	19.26	0.24			
Total	119	37.47				

Table 4.5: Two-Way ANOVA Table: Operational Variability (S10)

Source	DF	SS	MS	F	p	
Target	19	15.68	0.83	3.16	1.67e-04	***
Operator	1	0.82	0.82	3.13	0.0806	.
Target * Operator	19	4.96	0.26	0.71	0.79	
Error/Residuals	80	29.38	0.37			
Total	119	50.83				

Table 4.6: Two-Way ANOVA Table: Operational Variability (S11)

Source	DF	SS	MS	F	p	
Target	19	9.42	0.50	0.60	0.89	
Operator	1	3.53	3.53	4.28	0.0417	*
Target * Operator	19	15.68	0.83	3.42	6.16e-05	***
Error/Residuals	80	19.33	0.24			
Total	119	47.96				

From ANOVA Tables 4.3, 4.4, 4.5 and 4.6 we can conclude that the variability due to the target is of high statistical significance. On the other hand, the operator factor is only $DOF = 1$, and the significance is dissimilar in the ANOVA results. Therefore, the effect of the Operator is inconclusive.

Next, we tabulate the percent contribution due to each factor and combine the reproducibility and repeatability to the value for GRR.

- Calculate the Equipment Variation, Interaction, Operator Variation, Repeatability, Reproducibility and the GRR using the formulae from Section 3.3.2.

$$\begin{aligned} \text{Estimate of Variance: } & \text{Equipment} = \tau^2, \quad \text{Operator} = \omega^2 \\ & \text{Interaction} = \gamma^2, \quad \text{Target} = \sigma^2 \\ & \text{GRR} = \tau^2 + \gamma^2 + \omega^2 \end{aligned} \quad (4.11)$$

- Calculate the percentage contribution and percentage variation using formulae in Section 3.3.2.

Table 4.7: Variance Components (S00)

Source	Variance Estimate	STD	% Total Variation (STD)	% Contribution (of Variance)
Total Gage R&R	$\tau^2 + \gamma^2 + \omega^2 = 0.432$	0.657	78.505	61.63
Repeatability (EV)	$\tau^2 = 0.169$	0.411	49.114	24.122
Reproducibility (OV)	$\omega^2 = 0.207$	0.455	54.383	29.575
Operator*Target	$\gamma^2 = 0.056$	0.236	28.166	7.933
Target	$\sigma^2 = 0.269$	0.519	61.943	38.37
Total	$\tau^2 + \gamma^2 + \omega^2 + \sigma^2 = 0.701$	0.837		

Table 4.8: Variance Components (S01)

Source	Variance Estimate	STD	% Total Variation (STD)	% Contribution (of Variance)
Total Gage R&R	$\tau^2 + \gamma^2 + \omega^2 = 0.245$	0.495	84.888	72.059
Repeatability (EV)	$\tau^2 = 0.241$	0.491	84.08	70.694
Reproducibility (OV)	$\omega^2 = 0.005$	0.068	11.684	1.365
Operator*Target	$\gamma^2 = 0$	0	0	0
Target	$\sigma^2 = 0.095$	0.308	52.859	27.941
Total	$\tau^2 + \gamma^2 + \omega^2 + \sigma^2 = 0.34$	0.583		

Table 4.9: Variance Components (S10)

Source	Variance Estimate	STD	% Total Variation (STD)	% Contribution (of Variance)
Total Gage R&R	$\tau^2 + \gamma^2 + \omega^2 = 0.376$	0.614	89.446	80.007
Repeatability (EV)	$\tau^2 = 0.367$	0.606	88.339	78.038
Reproducibility (OV)	$\omega^2 = 0.009$	0.096	14.031	1.969
Operator*Target	$\gamma^2 = 0$	0	0	0
Target	$\sigma^2 = 0.094$	0.307	44.714	19.993
Total	$\tau^2 + \gamma^2 + \omega^2 + \sigma^2 = 0.471$	0.686		

Table 4.10: Variance Components (S11)

Source	Variance Estimate	STD	% Total Variation (STD)	% Contribution (of Variance)
Total Gage R&R	$\tau^2 + \gamma^2 + \omega^2 = 0.481$	0.694	100	100
Repeatability (EV)	$\tau^2 = 0.242$	0.492	70.858	50.209
Reproducibility (OV)	$\omega^2 = 0.045$	0.212	30.627	9.38
Operator*Target	$\gamma^2 = 0.194$	0.441	63.57	40.411
Target	$\sigma^2 = 0$	0	0	0
Total	$\tau^2 + \gamma^2 + \omega^2 + \sigma^2 = 0.481$	0.694		

As discussed in Section 3.3.2, equipment variance is equivalent to the unaccounted variance or the error/residual variance in the ANOVA results. From the ANOVA Tables 4.3, 4.4, 4.5 and 4.6 and Variance Composition Tables 4.7, 4.8, 4.9 and 4.10 we can infer the Repeatability (Equipment Variance) is large compared to reproducibility (Operator Variance). It can be explained by the following possible reasons as suggested in [24]:

- Gage may need to be redesigned to be more repeatable.
- The target receptacle needs to be improved.
- There is excessive within-target variation.

The key takeaway is that the total variance due to the gage accounts for most of the variance. As suggested in [24], if the GRR is above 30%, it is a sign to consider improving the measurement system. On the other hand, the acceptability of a measurement system is entirely dependent on the use case of the gage.

Acceptability of SilGeo

To discuss the acceptability of measurements made by [SilGeo](#) in the use case of detecting counterfeits, we look at Type-2 targets. Table 4.11 and Table 4.12 contain the summary statistics for nominal devices and devices of Type-2, respectively, captured using 1111 experiment configuration. Comparing the two, we can observe that the counterfeit targets are at least one order of magnitude away from the mean of nominal targets. Although the variability due to the [GRR](#) is high, as the MSA Reference Manual guides [24] the acceptability of a gage depends completely on the use case. Therefore, in this instance, the use case being the sifting of known counterfeit Z80 ICs, the use of [SilGeo](#) is acceptable.

Table 4.11: Summary Statistics (Nominal Targets)

Metric	d_S00	d_S01	d_S10	d_S11
Mean	79.46	39.57	38.52	83.08
Standard Deviation	0.99	0.75	0.82	0.88
Minimum	75.66	37.78	36.38	80.99
25th Percentile	78.73	39.05	38.04	82.50
Median (50th Percentile)	79.48	39.57	38.52	83.05
75th Percentile	80.20	40.02	38.99	83.65
Maximum	82.26	42.48	41.52	85.62
Count	360	360	360	360

Table 4.12: Summary Statistics (Counterfeit Targets)

Metric	d_S00	d_S01	d_S10	d_S11
Mean	8296.50	937.84	977.34	1354.97
Standard Deviation	441.92	12.89	15.11	39.26
Minimum	5887.43	913.66	952.74	1272.31
25th Percentile	8083.87	929.78	966.85	1348.04
Median (50th Percentile)	8420.78	935.27	973.62	1357.43
75th Percentile	8543.00	944.54	984.46	1371.56
Maximum	8850.10	995.43	1032.90	1442.56
Count	180	180	180	180

If the variability due to [GRR](#) would have been less than 30%, then [SilGeo](#) could have been used for detecting differences in the targets within the type.

4.3.2 Variability Analysis

We consider the data captured as per the experimental array to analyse the variability due to assembly. The experimental array features a balanced variation in four factors: the Adapter, multiplexer, cable, and the [VNA](#) and their interactions. For each configuration (row in the experiment array), we captured data for every target device three times with re-connection and three measurements for each of those connections, resulting in a total of 9 runs. This part aims to analyse the variability due to the assembly factors. Therefore, we consider the data only from one operator (Operator 2) and average the measurements of all the targets to obtain a single scalar measurement value for each of the [S-parameters](#) for analysis.

$$N = 144, \quad r = 9$$

where,

N is the total number of observations in the data

r is the repetitions made for each target, including re-connections and repeat measurements.

The total number of experiments is $16 = N/r$.

After this step, we apply the [ANOVA](#) technique and obtain the following:

- Sum of all observations and Correction Factor.

$$T = \sum_{i=1}^{16} \sum_{j=1}^r (X_{ij}), \quad CF = T^2/N \quad (4.12)$$

- Sum of Squares corresponding to each column in the experimental array.

$$SS_{\text{Col}} = \left(\sum_{i=1}^{n \text{ levels}} \frac{\left(\sum_{j=1}^r (X_{ij}) \right)^2}{(r * (\text{number of } X \text{ with level } i))} \right) - CF \quad (4.13)$$

- Consolidate the lower [DOF](#) columns into corresponding higher [DOF](#) factors following the rules of decomposition of variance and degrees of freedom.

$$SS_A = SS_{A1} + SS_{A2} + SS_{A3} \quad (4.14)$$

$$f_A = f_{A1} + f_{A2} + f_{A3} \quad (4.15)$$

$$SS_{AB} = SS_{A1B} + SS_{A2B} + SS_{A3B} \quad (4.16)$$

$$SS_{AC} = SS_{A1C} + SS_{A2C} + SS_{A3C} \quad (4.17)$$

- Use the sum of squares for each un-aliased column assigned to a factor and calculate the error sum of squares.

$$SS_e = SS_T - SS_{Bet} \quad (4.18)$$

$$SS_{Bet} = SS_A + SS_B + SS_C + SS_{AB} + SS_{AC} + SS_{BC} + SS_{CD} \quad (4.19)$$

$$f_e = f_T - f_{Bet} \quad (4.20)$$

- Calculate the mean squared contribution of each factor and their interactions, along with the percent contribution.

$$MS_{Col} = SS_{Col}/f_{Col} \quad (4.21)$$

$$P_{Col} = 100 * (SS_{Col}/SS_T) \quad (4.22)$$

- Calculate the F-ratio with the mean squared error and the p-value as per the F-distribution CDF.

$$F_{Col} = MS_{Col}/MS_e \quad (4.23)$$

$$p_{Col} = 1 - F_{CDF}(F_{ratio}, f_{Col}, f_e) \quad (4.24)$$

- Assign a significance to the p-values.

p-value	Significance
> 0.1	
> 0.05	.
> 0.01	*
> 0.001	**
others	***

The resulting [ANOVA](#) tables for [S-parameters](#) S00, S01, S10, and S11 are shown in Tables [4.13](#), [4.14](#), [4.15](#) and [4.16](#) respectively.

Table 4.13: ANOVA Table: Assembly Variability (S00)

Source	DF	SS	MS	Percent	F	p	
A	3	4.74	1.58	22.49 %	31.42	2.44e-15	***
A1	1	0.44	0.44	2.07 %	8.69	0.0038	**
A2	1	0.04	0.04	0.21 %	0.87	0.353	
A3 = BD	1	4.26	4.26	20.21 %	84.70	7.77e-16	***
B	1	0.32	0.32	1.53 %	6.42	0.0125	*
C	1	3.05	3.05	14.46 %	60.60	1.93e-12	***
AB	3	1.02	0.34	4.85 %	6.77	2.80e-04	***
A1B	1	0.70	0.70	3.33 %	13.95	2.80e-04	***
A2B	1	0.31	0.31	1.45 %	6.09	0.0149	*
A3B = D	1	0.01	0.01	0.07 %	0.28	0.596	
AC	3	3.19	1.06	15.11 %	21.11	3.33e-11	***
A1C	1	0.25	0.25	1.20 %	5.05	0.0264	*
A2C	1	2.91	2.91	13.81 %	57.87	5.00e-12	***
A3C	1	0.02	0.02	0.10 %	0.42	0.519	
BC	1	0.92	0.92	4.36 %	18.29	3.65e-05	***
CD	1	1.30	1.30	6.16 %	25.83	1.27e-06	***
Error/Residuals	130	6.54	0.05	31.02 %			
Total	143	21.09					

Table 4.14: ANOVA Table: Assembly Variability (S01)

Source	DF	SS	MS	Percent	F	p	
A	3	10.20	3.39	29.40 %	37.70	1.11e-16	***
A1	1	10.10	10.10	29.30 %	112.00	1.11e-16	***
A2	1	0.0572	0.0572	0.166 %	0.637	0.426	
A3 = BD	1	3.86e-05	3.86e-05	0.112 %	0.430	0.983	
B	1	0.922	0.922	2.67 %	10.3	0.0017	**
C	1	3.82	3.82	11.1 %	42.6	1.39e-09	***
AB	3	0.818	0.273	2.37 %	3.04	0.0315	*
A1B	1	0.0515	0.0515	0.149 %	0.574	0.450	
A2B	1	0.0612	0.0612	0.177 %	0.681	0.411	
A3B = D	1	0.705	0.705	2.04 %	7.86	0.00584	**
AC	3	4.65	1.55	13.1 %	17.3	1.70e-09	***
A1C	1	1.02	1.02	2.96 %	11.4	9.88e-04	***
A2C	1	3.00	3.00	8.69 %	33.4	5.25e-08	***
A3C	1	0.630	0.630	1.83 %	7.02	0.00907	**
BC	1	1.07	1.07	3.11 %	12.0	7.37e-04	***
CD	1	1.40	1.40	4.05 %	15.6	1.27e-04	***
Error/Residuals	130	11.7	0.0898	33.8 %			
Total	143	34.5					

Table 4.15: ANOVA Table: Assembly Variability (S10)

Source	DF	SS	MS	Percent	F	p	
A	3	17.9	5.98	15.1 %	53.8	1.11e-16	***
A1	1	15.7	15.7	13.2 %	141	1.11e-16	***
A2	1	1.26	1.26	1.06 %	11.3	0.00102	**
A3 = BD	1	0.993	0.993	0.837 %	8.94	0.00334	**
B	1	0.0659	0.0659	0.0555 %	0.593	0.443	
C	1	68.7	68.7	58.0 %	619	1.11e-16	***
AB	3	3.41	1.14	2.87 %	10.2	4.33e-06	***
A1B	1	1.68	1.68	1.42 %	15.1	1.58e-04	***
A2B	1	1.67	1.67	1.41 %	15.0	1.67e-04	***
A3B = D	1	0.0535	0.0535	0.0451 %	0.482	0.489	
AC	3	9.82	3.27	8.28 %	29.5	1.31e-14	***
A1C	1	5.90e-05	5.90e-05	4.97e-05 %	5.31e-04	0.982	
A2C	1	7.16	7.16	6.04 %	64.5	5.12e-13	***
A3C	1	2.66	2.66	2.25 %	24.0	2.82e-06	***
BC	1	0.0551	0.0551	0.0465 %	0.496	0.482	
CD	1	4.15	4.15	3.50 %	37.4	1.06e-08	***
Error/Residuals	130	14.4	0.111	12.2 %			
Total	143	119					

Table 4.16: ANOVA Table: Assembly Variability (S11)

Source	DF	SS	MS	Percent	F	p	
A	3	5.61	1.87	13.2 %	21.06	3.49e-11	***
A1	1	3.27	3.27	7.67 %	36.80	1.33e-08	***
A2	1	1.02	1.02	2.39 %	11.49	9.27e-04	***
A3 = BD	1	1.32	1.32	3.10 %	14.90	1.78e-04	***
B	1	4.80	4.80	11.3 %	54.10	1.91e-11	***
C	1	8.39	8.39	19.7 %	94.50	1.11e-16	***
AB	3	1.23	0.409	2.88 %	4.61	0.00425	**
A1B	1	0.123	0.123	0.288 %	1.38	0.241	
A2B	1	0.285	0.285	0.669 %	3.21	0.0754	.
A3B = D	1	0.819	0.819	1.92 %	9.22	0.00289	**
AC	3	9.81	3.27	23.0 %	36.84	1.11e-16	***
A1C	1	0.296	0.296	0.694 %	3.33	0.0703	.
A2C	1	8.42	8.42	19.7 %	94.80	1.11e-16	***
A3C	1	1.10	1.10	2.58 %	12.39	5.96e-04	***
BC	1	1.23	1.23	2.90 %	13.90	2.86e-04	***
CD	1	0.00186	0.00186	0.00437 %	0.02	0.885	
Error/Residuals	130	11.5	0.0888	27.1 %			
Total	143	42.6					

After the analysis of data presented in Tables 4.13-4.16, we can immediately notice that the main factors A (Adapter), B (Multiplexer) and C(VNA) are statistically significant. On the other hand, the statistical significance of interactions is inconclusive.

We now move to discussing the contribution of the factors to the variability. The error represents about 30% of the variability on average. Moreover, even though the factor B has a statistically significant contribution, the percentage contribution is about 3 – 4% on average, lower as compared to other factors. The statistical significance of factor D is inconclusive at first sight, as it is confounded with one interaction of factors A and B . The significance of factor D can be ignored, as it can be observed that it shows significance only when the factor B shows statistical significance.

Now that we have discussed the effect of the main factors, the effect of interactions can be explained. As we can see in Figure 4.7, the magnitude of change due to variation in factors A , and C is higher than the change due to factors B and D , which corroborates with the higher significance of interactions AC (heredity principle, Section 3.4.2). Similarly, the interactions AB and BC are statistically significant, but the overall contribution is small, similar to the contribution of factor B .

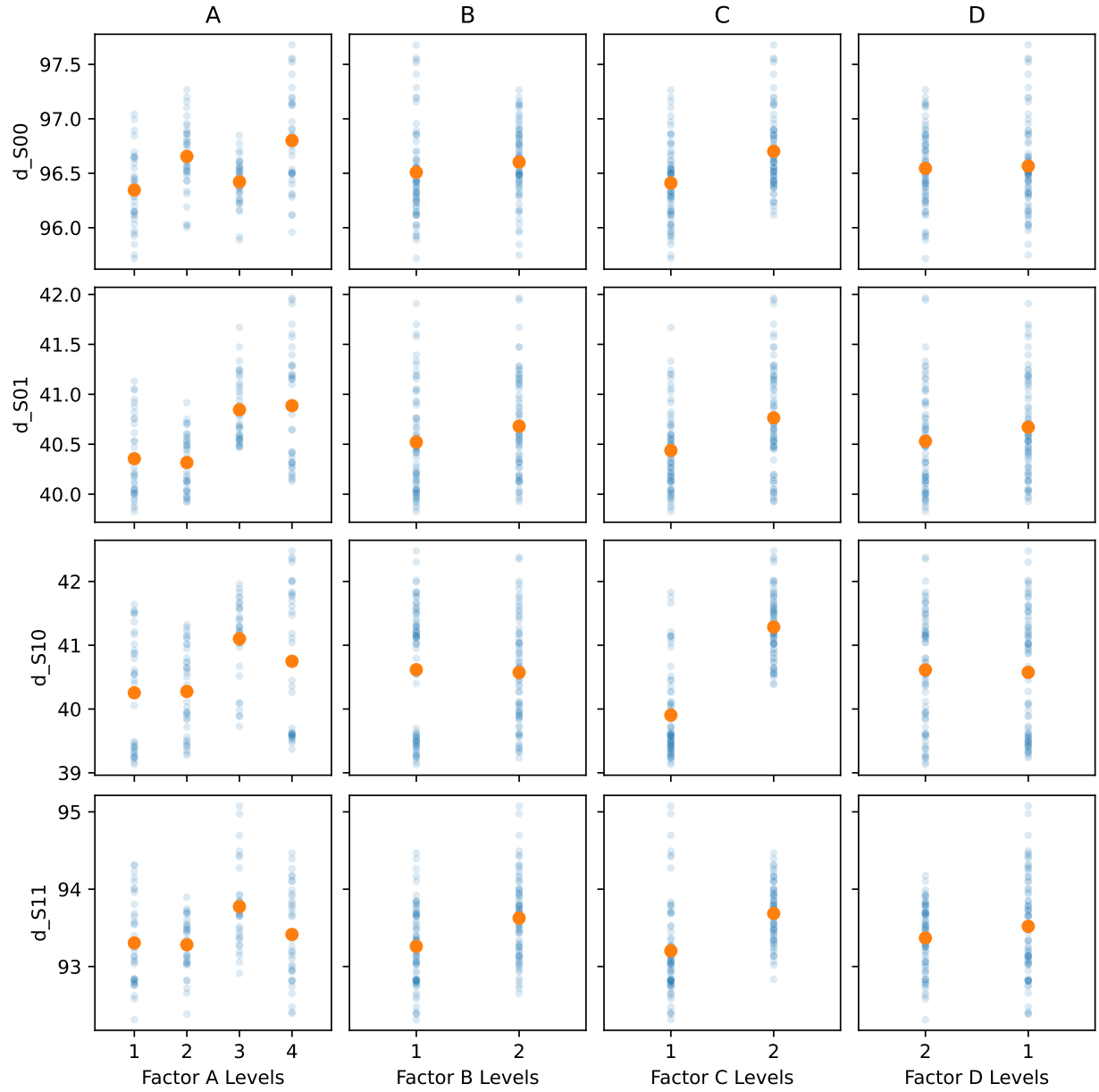


Figure 4.7: Main Effects Plot of SilGeo Assembly

Note: The orange points represent the mean of measurements, and the plot presents the effects of *qualitative* factors. The y-axes d_S00, d_S01, d_S10 and d_S11 denote the distance metric value of the S-parameters S_{00} , S_{01} , S_{10} , and S_{11} respectively.

Assembly Variability

The [ANOVA](#) analysis gives us an initial perspective on the following:

- The majority of variability can be attributed to random error.
- Factors such as the [VNA](#) (Measurement Instrument), the Adapter Board and their interaction have significantly higher contributions to variability.
- The Multiplexer board and the Cable are relatively less variable in their contribution; hence, they can be removed from future studies.

Type-2 Targets (Counterfeits)

Another takeaway from the experiment and the analysis is comparing the data with the captured values corresponding to 10 known counterfeit targets. Table [4.17](#) shows the summary statistics for the known legitimate targets. Comparing it with Table [4.18](#), we can see that the counterfeit targets are at least one order of magnitude away from the mean of nominal targets.

Table 4.17: Summary Statistics (Nominal Targets)

Metric	d_S00	d_S01	d_S10	d_S11
Mean	96.56	40.60	40.60	93.44
Standard Deviation	0.38	0.49	0.91	0.55
Minimum	95.72	39.83	39.13	92.31
25th Percentile	96.31	40.19	39.62	93.05
Median (50th Percentile)	96.52	40.54	40.66	93.44
75th Percentile	96.79	40.92	41.29	93.79
Maximum	97.68	41.97	42.48	95.08
Count	144	144	144	144

Table 4.18: Summary Statistics (Counterfeit Targets)

Metric	d_S00	d_S01	d_S10	d_S11
Mean	8454.40	788.06	772.14	866.64
Standard Deviation	2070.47	72.91	83.72	32.88
Minimum	6102.67	675.55	647.14	811.24
25th Percentile	7070.64	729.88	693.52	845.88
Median (50th Percentile)	7858.56	779.55	763.80	865.53
75th Percentile	8862.95	831.46	827.43	879.56
Maximum	13693.46	975.91	985.40	982.16
Count	144	144	144	144

Chapter 5

Conclusion

We now summarise our findings and list the key takeaways (TAs) as follows:

- TA1** To distinguish between two targets, the measurement **GRR** must be smaller than the variability between the two targets. For the Z80 targets, **SilGeo**'s measured **GRR** was sufficiently small to distinguish between instances of the nominal (Type-1) and counterfeit (Type-2) targets. Therefore, **SilGeo** is an acceptable measurement system for distinguishing the nominal targets from counterfeit targets in this Z80 case study.
- TA2** **SilGeo**'s **GRR** was too large to distinguish between two instances within the same target type, for example, distinguishing between two Z80 ICs in the nominal class.
- TA3** The magnitude of the effect due to the operator and its interaction with the target device is inconclusive.
- TA4** The main factors contributing to assembly variability are the **VNA** and the Adapter board.
- TA5** The majority of variability in the assembly can be attributed to experimental/random error resulting from uncontrolled sources of variation.

As a concluding remark, we can say that **SilGeo** is acceptable for detecting counterfeit Z80 ICs that we saw in this case study. However, this statement cannot be generalised to include unknown counterfeits in Z80 or other target categories.

But, the fact is that we performed this case study using real parts procured from trusted and untrusted vendors verified using X-ray imaging. Additionally, [SilGeo](#) will successfully detect counterfeit targets if their intrinsic variation is due to them being counterfeit is larger than the equipment variability.

Furthermore, we observe that most variability in assembly is due to random, uncontrollable, noisy factors. [SilGeo](#) has significant room for improvement in the robustness of design and repeatability aspects.

5.1 Recommendations

Based on the results of the variability analysis of the data, in this section, we make recommendations to minimize the effect of the sources of variability. Figure 5.1 shows the effect of remedial actions on the sources of variability in [SilGeo](#). We classify the approach to counter the sources of variability as follows:

Assembly Calibration

The MSA Reference Manual [24] defines calibration as a process involving a set of operations under specific conditions to establish a relation between the measurement system and a reference standard. As discussed in this thesis, [SilGeo](#) as a hardware assurance system does not operate with a standard known reference compared to a conventional measurement system. Hence, a calibration mechanism in [SilGeo](#) must address each main controllable source of variability in the system's assembly individually.

Performing assembly calibration contributes to alleviating the effect of high [GRR](#) described in [TA2](#) and minimizing the variability in components [TA4](#).

Operational Refinement

Improvement in the operation of [SilGeo](#) to capture hardware fingerprints and deliver hardware assurance can significantly improve the reproducibility of results. Additionally, we recommend having a well-defined test procedure, care in handling and proper operator training. This can curb the variability of existing and future effects due to factors such as wear and tear (Ageing) on certain parts of the measurement system.

Choosing suitable operators may also reduce the variability due to specific human factors such as aptitude, positive attitude, and experience.

Another improvement is to remove operator intervention in the measurement process. However, this option requires significant upfront investment in equipment in addition to research and development.

Operational refinement contributes towards mitigating the drawbacks described in [TA2](#) and [TA5](#).

Design Modifications

Following are the possible design improvements that we recommend for [SilGeo](#) on the basis of the results obtained in the study.

- DM1** Adopt design verification methods such as [RF](#) design simulation before manufacturing.
- DM2** Design or use [RF-capable](#) and impedance-matched receptacles to connect with the target devices.
- DM3** Minimizing the use of mechanically actuating parts for electrical contacts in the design.
- DM4** To counter [EMI](#), a noisy factor, we recommend having shielding to the exposed areas of the test setup.
- DM5** Employing guiding fixtures for target insertion into the receptacle will help in the reproducibility of results.
- DM6** General increase in stiffness and robustness in mechanical design to avoid variations due to vibration and accidental impacts.

Making design improvements has a significantly more positive effect to the performance of [SilGeo](#). Improvements [DM2](#), [DM3](#), [DM4](#) and [DM6](#) will improve the overall variability and minimize [TA5](#) and reduce assembly variability [TA4](#). Adding a guiding fixture as mentioned in [DM5](#) will minimize the variability in operation and therefore contribute to mitigating the high [GRR TA2](#).

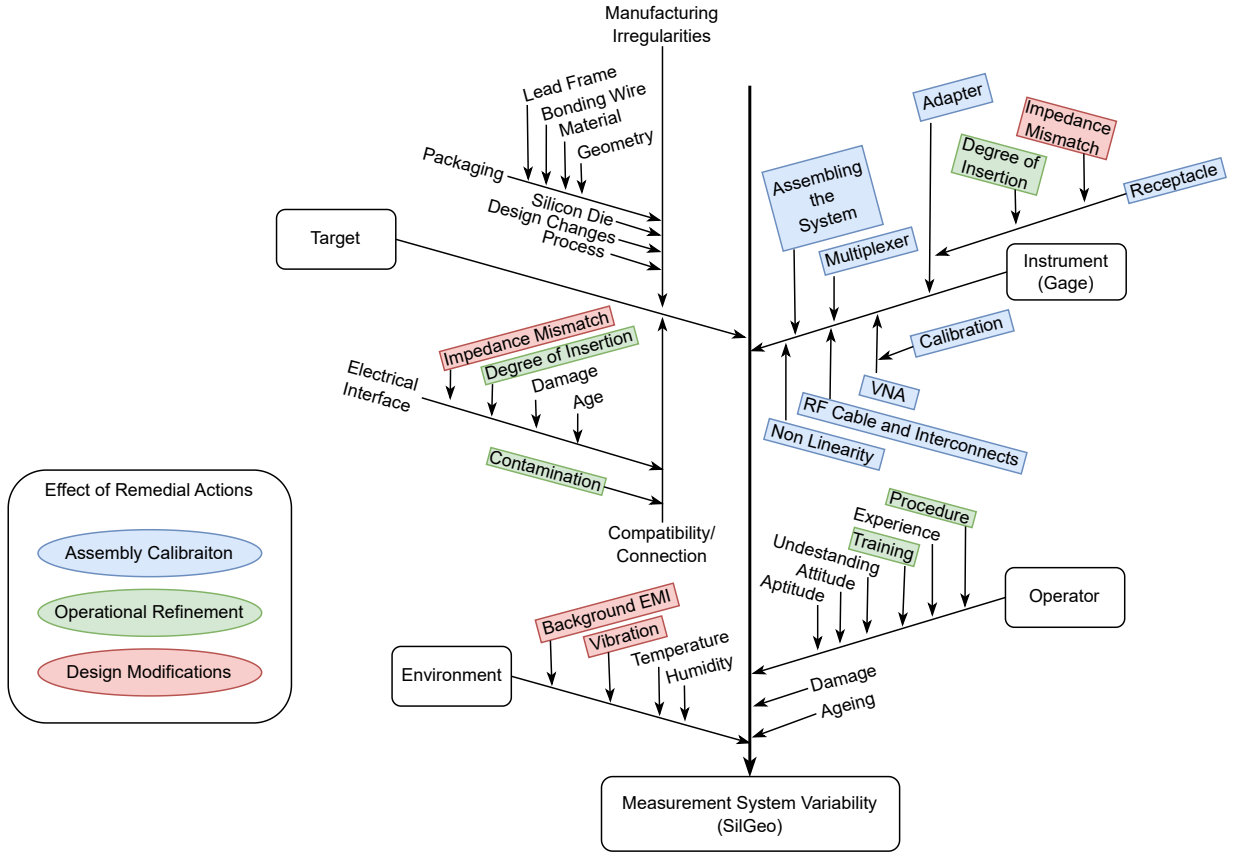


Figure 5.1: Potential Impact of Remedial Actions on Sources of Variability in [SilGeo](#)

5.2 Future Work

This section discusses the next possible steps to reduce the observed variability in [SilGeo](#). The repeatability and reproducibility of the measurement system must be improved to increase precision and accuracy so that differences within the target category are detectable. This goal would first and foremost involve increasing the robustness and noise tolerance of the hardware design.

An important step for [SilGeo](#) is developing an assembly calibration procedure. This can be achieved by calibrating the VNA with the cables separate from the assembly and then measuring the [S-parameters](#) of the remaining assembly components. This process of performing individual characterization of the components in the assembly is called de-embedding [32]. Ideally, the de-embedding process should be applied to all components

except the target device. But in the case of [SilGeo](#), there is no S-parameter data for receptacles sourced from third-party manufacturers. Hence, de-embedding can be done till the multiplexer board.

Another application of the variability results is the [Machine Learning \(ML\)](#) model's transferability between assemblies. As discussed, [SilGeo](#) employs a threshold-based novelty detection [ML](#) model designed by Schmidt [33] to classify targets as nominal or counterfeit on a case-by-case basis. When trained on data captured from one [SilGeo](#) assembly, this [ML](#) model is not necessarily transferable; in other words, the model is locked to the hardware assembly. To alleviate this limitation, the variability results from the study can be used to relax the model threshold. This approach is feasible because counterfeit targets present at a significant distance compared to the nominal targets. The main challenge is translating the distance metric used in the variability analysis into the metric used by the [ML](#) model.

Further, the variability study on [SilGeo](#) is only a partial evaluation of [SilGeo](#) and one step towards standardising the technique for industry. [SilGeo](#) remains to be evaluated on long-term performance and stability. Therefore, the evaluation of ageing effects, temperature, and miscellaneous effects on the measurement is yet to be seen and is a definitive next step from this study.

Finally, the results from the variability study only apply to the Z80 type targets and adapters. Therefore, each new adapter type would have to be evaluated for variability. Fortunately, the study gives us insights into the factor effects and interactions; therefore, for further evaluations, the size of the experimental array can be significantly smaller (two to four experiments).

References

- [1] Scott Fortmann-Roe. Understanding the Bias-Variance Tradeoff. <https://scott.fortmann-roe.com/docs/BiasVariance.html>.
- [2] Gen'ichi Taguchi, Subir. Chowdhury, Yulin. Wu, Shin. Taguchi, and Hiroshi Yano. *Taguchi's Quality Engineering Handbook*. John Wiley & Sons, Hoboken, N.J, 2005.
- [3] Keysight. P9370A Keysight Streamline USB Vector Network Analyzer, 4.5 GHz, Feb 2021. <https://www.keysight.com/us/en/product/P9370A/streamline-usb-vector-network-analyzer-4-5-ghz.html>.
- [4] Chien-Fu Wu and Michael S. Hamada. *Experiments : Planning, Analysis, and Optimization*. Wiley series in probability and statistics. John Wiley & Sons, Inc., Hoboken, NJ, third edition. edition, 2021.
- [5] Tapesh Joshi, Saad Azam, Daniel Juarez-Robles, and Judith A. Jeevarajan. Safety and Quality Issues of Counterfeit Lithium-Ion Cells. *ACS Energy Letters*, 8(6):2831–2839, 2023.
- [6] Asurion Recalls Counterfeit BlackBerry branded Batteries Due to Burn and Fire-Hazards. *U.S. Consumer Product Safety Commission*, Aug 2010. <https://www.cpsc.gov/Recalls/2010/Asurion-Recalls-Counterfeit-BlackBerry-branded-Batteries-Due-to-Burn-and-Fire-Hazards>.
- [7] Underwriters Laboratories LLC. Determining the authenticity of cells & batteries in the global market. *Emerging Technology News*, Sep 2021. <https://etn.news/energy-storage/determining-authenticity-of-cells-batteries-in-the-global-market>.
- [8] US Attorney's Office District of Connecticut. New york man admits supplying falsely remarked computer chips used in u.s. military helicopters, Jul

2015. <https://www.justice.gov/usao-ct/pr/new-york-man-admits-supplying-falsely-remarked-computer-chips-used-us-military>.
- [9] Zhiye Liu. Fake Samsung 980 Pro SSDs Are Spreading Around. *Tom's Hardware*, March 2023. <https://www.tomshardware.com/news/fake-samsung-980-pro>.
 - [10] Carlos Moreno, Sebastian Fischmeister, and Philippe Vibien. A method and apparatus for detection of counterfeit parts, compromised or tampered components or devices, tampered systems such as local communication networks, and for secure identification of components. Patent No.: US20220341990A1.
 - [11] Philippe Vibien. SilGeo: A Method for the Detection of Counterfeit, Compromised, or Tampered Electronic Devices. Master's thesis, University of Waterloo, August 2022.
 - [12] Carlos Moreno, Philippe Vibien, Srijan Pabbi, Lukas Schmidt, Goksen Guler, and Sebastian Fischmeister. Microcircuit Defect and Counterfeit Detection via Silicon Geometry. In *2023 34th Annual SEMI Advanced Semiconductor Manufacturing Conference (ASMC)*, pages 1–6, 2023.
 - [13] Franklin Kuo. *Network Analysis and Synthesis, 2nd ed.* John Wiley & Sons, 2006.
 - [14] Michael Steer. *Fundamentals of Microwave and RF Design*. NC State University, Sept 2019.
 - [15] Leif Bjørnø, Thomas H. Neighbors, and David Bradley. *Applied underwater Acoustics*. Elsevier, Amsterdam, Netherlands, 2017 - 2017.
 - [16] Angelos Mallios, Eduard Vidal, Ricard Campos, and Marc Carreras. Underwater Caves Sonar Dataset. *The International Journal of Robotics Research*, 36(12):1247–1251, 2017.
 - [17] D.E. Weston. Applied Underwater Acoustics: by D. G. Tucker and B. K. Gazey. The Commonwealth and International Library. Pergamon Press. 244 + xvi pp. 27s. 6d. 1966. *Deep-sea research and oceanographic abstracts*, 14(1):140–142, 1967.
 - [18] How the Eyes Work | National Eye Institute, April 2022. <https://www.nei.nih.gov/learn-about-eye-health/healthy-vision/how-eyes-work>.
 - [19] Robert S. Witte and John S. Witte. *Statistics*. J. Wiley & Sons, Hoboken, NJ, 9th ed. edition, 2010.

- [20] David C. Howell. *Statistical Methods for Psychology*. Wadsworth Cengage Learning, Belmont, CA, 8th ed. edition, 2013.
- [21] Douglas C Montgomery. *Design and Analysis of Experiments*. John Wiley & Sons Inc, 8th ed. edition, 2013.
- [22] Charles Zaiontz. Two Factor ANOVA with Replication. <https://real-statistics.com/two-way-anova/two-factor-anova-with-replication/>.
- [23] Jiju. Antony. *Design of Experiments for Engineers and Scientists*. Butterworth-Heinemann, Oxford, 2003.
- [24] General Motors Corporation (AIAG) Chrysler Corporation, Ford Motor Company. *Measurement Systems Analysis Reference Manual, Fourth Edition*. AIAG, 2010.
- [25] Mark Allen. Durivage. *Practical Attribute and Variable Measurement Systems Analysis (MSA) : a Guide for Conducting Gage R & R Studies and Test Method Validations*. ASQ Quality Press, Milwaukee, Wisconsin, 2015 - 2015.
- [26] Ranjit K. Roy. *A Primer on the Taguchi Method*. Society of Manufacturing Engineers, Dearborn, Mich, 2nd ed. edition, 2010.
- [27] Raghu N Kacker, Eric S Lagergren, and James J Filliben. Taguchi's Orthogonal Arrays are Classical Designs of Experiments. *J. Res. Natl. Inst. Stand. Technol.*, 96(5):577–591, September 1991.
- [28] Hand-Flex Interconnect, 0.141" center diameter, 18.0 GHz | 141-3SMR+ | Mini-Circuits. <https://www.minicircuits.com/WebStore/dashboard.html?model=141-3SMR%2B>.
- [29] Touchstone Format - Keysight Knowledge Center. <https://edadocs.software.keysight.com/display/genesys2010/Touchstone+Format>.
- [30] R. De Maesschalck, D. Jouan-Rimbaud, and D.L. Massart. The Mahalanobis Distance. *Chemometrics and Intelligent Laboratory Systems*, 50(1):1–18, 2000.
- [31] Kun Il. Park. *Fundamentals of Probability and Stochastic Processes with Applications to Communications*. Springer International Publishing, Cham, 1st ed. 2018. edition, 2018.

- [32] A.M. Mangan, S.P. Voinigescu, Ming-Ta Yang, and M. Tazlauanu. De-embedding Transmission Line Measurements for Accurate Modeling of IC Designs. *IEEE Transactions on Electron Devices*, 53(2):235–241, 2006.
- [33] Lukas Schmidt. Novelty Detection for SilGeo Hardware Assurance. Master’s thesis, University of Waterloo, 2023.
- [34] Damir Akhoundov. 2022 Annual report. *ERAI BLOG*, Jan 2023. https://www.eraai.com/eraai_blog/3181/2022_annual_report.
- [35] Judy Jeevarajan. What makes counterfeit batteries unsafe? *UL Research Institutes*, Feb 2023. <https://ul.org/news/what-makes-counterfeit-batteries-unsafe>.
- [36] Alana Semuels. When your Amazon Purchase Explodes, Apr 2019. <https://www.theatlantic.com/technology/archive/2019/04/lithium-ion-batteries-amazon-are-exploding/587005/>.
- [37] Ruth M. Mickey, Olive Jean. Dunn, and Virginia Clark. *Applied Statistics : Analysis of Variance and Regression*. Wiley series in probability and statistics. Wiley-Interscience, Hoboken, N.J, 3rd ed. edition, 2004.
- [38] Pretesh Biswas. Analysis of Variance - ANOVA, Nov 2020. <https://preteshbiswas.com/2020/11/07/analysis-of-variance-anova/>.
- [39] The Ultimate Guide to Cause and Effect Diagrams, May 2018. <https://www.juran.com/blog/the-ultimate-guide-to-cause-and-effect-diagrams/>.

Glossary

checksum A checksum is like a unique fingerprint of a file that can be used to verify whether two files are identical. Each time you run a checksum, a number string is created for each file. Even if one byte of data has been altered or corrupted, that string will change. [8](#)

deflagration Deflagration is subsonic combustion in which a pre-mixed flame propagates through an explosive or a mixture of fuel and oxidizer. In other words, an explosion where the flame speed is lower than the speed of sound. [2](#)

Die A die, in the context of integrated circuits, is a small block of semiconducting material on which a given functional circuit is fabricated. [61](#)

DRAM Dynamic random-access memory (dynamic RAM or DRAM). [2](#)

hashing Hashing is the process of transforming any given key or a string of characters into another value. This is usually represented by a shorter, fixed-length value or key that represents and makes it easier to find or employ the original string. [8](#)

hydrographic surveying Hydrographic surveying is an important civil engineering service that determines the physical features of an underwater area. [9](#)

Line-Replaceable Unit (LRU) The LRU is a component which is designed to be removed and replaced by line maintenance operator on aircraft (or any other manufactured device) with time constraints (typically with a limit of 20 mins). [10](#), [96](#)

multiplexing Multiplexing, or muxing, is a technique of sending multiple signals or streams of information over a communications link at the same time in the form of a single, complex signal. [13](#), [14](#)

oceanography Oceanography is the study of the physical, chemical, and biological features of the ocean. [9](#)

receptacle A receptacle is an electrical housing/connector into which the plug of an electrical device or the electrical device may be inserted. [10](#)

RF Radio Frequency: RF part of the electromagnetic spectrum is generally defined with frequencies anywhere in the range of 3 kHz to 300 GHz. [7](#), [9](#), [10](#), [13](#), [15](#), [28](#), [30–32](#), [35](#), [55](#), [63](#), [88](#)

RF-capable Capability to conduct electrical signals in RF spectrum. [13](#), [14](#), [30](#), [63](#), [88](#)

Shop-Replaceable Unit (SRU) The SRU is part of a [Line-Replaceable Unit \(LRU\)](#) and is designed to be removed and replaced in the shop. [10](#)