MACHINING ACCURACY ENHANCEMENT USING MACHINE TOOL ERROR COMPENSATION AND METROLOGY

BY

HUA-WEI KO

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Mechanical Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2019

Urbana, Illinois

Doctoral Committee:

Professor Shiv G. Kapoor, Chair
Professor Placid M. Ferreira
Professor T. Kesh Kesavadas
Assistant Professor Chenhui Shao

ABSTRACT

This dissertation aims to enhance machining accuracy by machine tool error reduction and workpiece metrology. The error characteristics are studied by building a quasi-static error model. Perturbed forward kinematic model is used for modeling a 5-axis Computer Numerical Control (CNC) machine with one redundant linear axis. It is found that the 1st order volumetric error model of the 5-axis machine is attributed to 32 error parameter groups. To identify the model by estimating these parameter groups using the least-squares fitting, errors at 290 quasi-randomly generated measurement points over the machine's workspace are measured using a laser tracker. The identified error model explains 90% of the mean error of the training data sets. However, the measurements using the laser tracker take about 90 minutes, which may cause the identified error parameters to be inaccurate due to the slow varying and transient natures of thermal errors.

To shorten the measurement time, an experimental design approach, which suggests the optimal observation locations such that the corresponding robustness of identification is maximized, is applied to design the optimal error observers. Since the observers must be uniformly distributed over the workspace for gaining redundancy, the constrained K-optimal designs are used to select 80 K-optimal observers for the 5-axis machine. Six measurement cycles using 80 observers are done at machine's different thermal states within a 400-minute experiment. Six error models are trained with consistent performances and are found to be comparable to the one trained by 290 quasi-random observations. This shows the feasibility of using smaller but more strategical-chosen point-set in data-driven error models. More importantly, the growth on mean nominal (119.1 to 181.9 microns) and modeled error (26.3 to 33.9 microns) suggest the necessity of thermal error tracking for enhancing the machining accuracy.

A point-set based metrology is also developed to compensate the inaccuracies introduced by workpiece and fixtures and enhance machining accuracy. The machinability of all planar features is examined by virtually comparing the scanned data with the nominal machining planes, which are also known as virtual gages. The virtual gaging problem is modeled as a constrained linear program. The optimal solution to the problem can compensate the displacement introduced by workpiece and fixtures and hence guarantee a conforming finished part. To transfer point-set data into mathematical constraints, algorithms that align, segment, downsize and filter the point-set data are exploited. The concept of virtual gage analysis is demonstrated using experimental data for a simple raw casting. However, for the case where the casting is defective, and some virtual gages are not feasible, the corresponding linear program was found to have no solution. By introducing slack variables to the original linear programming problem, the extended problem has been solved. The extended model is validated for the data obtained for another casting. Further, the analysis predicts the machining allowances on all functional features.

Cylindrical surface and its tolerance verification play important role in machining process. Although there exist many approaches that can fit the maximum, minimum and minimum zone cylinders, the cylinder fitting problems can be even simplified. The proposed methodology seeks to reduce the number of parameters used in cylinder fitting model by using the projection model that considers the degenerated tolerance specifications of the projected 2-D point-set. Also, to avoid the problem of local optimum by introducing the optimal direction of projection such that the 2-D point projected onto this direction has optimal tolerance specifications (maximum, minimum and minimum zone circles), global optimum solver such as Particle Swarm Optimization (PSO) is used. The proposed simplified method shows consistent results compared with the results from literature.

ACKNOWLEDGEMENTS

I wish to express my greatest appreciation to my co-advisors, Professor Kapoor and Professor Ferreira for their guidance, support and encouragement during these four years. They have continually challenged me to broaden my horizon. Also, I gratefully thank Professor Kesavadas and Professor Shao for their feedback and their consent to serve in my dissertation committee. My sincere thanks to Yujie, Nien, Rohit and Richard of Caterpillar Inc. for their inputs and support on experiments. I would like to acknowledge the financial support from Department of the Army through the Digital Manufacturing and Design Innovation Institute under project DMDII 14-07-05: Integrated Manufacturing Variation Management.

I would like to thank Professor Huang, Professor Chen and Professor Lu of National Taiwan University, who motivated me to pursue a doctorate. I also appreciate Professor Wu and Professor Fu of National Chiao Tung University for being my references.

My labmates, Jorge, Ricardo and Soham and Patrick, Adam, Jeremy and Krishna of Missouri University of Science and Technology deserve the deepest appreciation for all the assistance. Huge thanks to all my friends in University of Illinois at Urbana Champaign, especially my roommates and best friends, CM, Shannon, Jack, Angela, PC and PL. The time hanging out with you will be my most valuable memory.

Special thanks to the old friends, Eric, Hsiun, Yen, Wayne, Ted and Billy for being my "intercontinental consultants". I enjoy the productive discussions with you and acknowledge all the helpful suggestion. Your contributions to this dissertation are beyond imagination.

The sincere appreciation to my parents for always motivating me to do my best and pursue my dreams is beyond description. Your love, encouragement, and confidence in me help me through the ups and downs of life.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1. INTRODUCTION

1.1 BACKGROUND AND MOTIVATION

A Computer Numerical Control (CNC) machine tool uses a computer-controlled servo system to interpret and execute pre-programmed sequences of motion commands. They were developed to machine parts with complex shape in a precise manner [16]. The required tasks for producing a part using CNC machines include:

(a) Offline tasks, which start with a geometric model, i.e., by 2D or 3D Computer Aided Design (CAD) model and generate a RS-274 part program (often referred to a g-code program) for controlling the NC machine to produce it,

(b) Online tasks, composed of reading and interpreting the part program generated by the offline tasks and the subsequent generation of motion of the tool relative to the workpiece so that the part is properly machined, and

(c) Post-processing tasks that include inspection of the finished part, either on the machine or with a Coordinate Measurement Machine (CMM). The measured dimensions are compared with the CAD model of the part and corrections are made, if necessary. The compensation of the measured error can be achieved by modifying tool offsets and performing additional machining operations.

CNC machine tools are very flexible and versatile. They are capable of machining a variety of parts with different geometries by changing the part programs and using different cutting tools. In contrast to manual machines where motion is manually controlled by hand wheels and the workpiece is intermittently measured by external devices to achieve the desired accuracy, CNC machines must rely on the accuracy and stability of their mechanical frame and their motion control system for accuracy. This introduces uncertainty in the positioning of the tool relative to the workpiece and, hence, the accuracy of the parts machined, because of quasi-static flexural and thermal deformations of a machine's frame as well as dynamic errors (for example, due to machining forces) in its positioning system. Periodic maintenance and calibration for the machine tool can alleviate some of this uncertainty. However, the errors that accrue from the above-mentioned effects can be addressed by modeling the machine tool error and employing the workpiece metrology.

In general, a machine tool designer should consider (i) the overall static stiffness against flexural strain caused by self-weight of moving parts and the workpiece or the assembly error, (ii) thermal stiffness against thermally-induced deformation, and (iii) the dynamic performance (e.g., sensitivity to periodic forcing and resonances). A well-designed machine produces less machine tool error, defined by "the difference between the actual response of a machine to a command issued according to the accepted protocol of that machine's operation and the response to that command anticipated by that protocol" [1]. Errors can be reduced by avoidance or elimination. The former seeks to remove the error sources and thus avoid the errors being produced. For example, the thermally-induced error can be avoided at the design stage of the machine tool by distributing the internal heat sources uniformly [2]–[5] or by operating the machine in an environment with the temperature well controlled [6]. However, the nature of the quasi-static errors are experimentally observed to be slowly time varying, which makes them not only difficult to be identified and compensated but also dominant in machining inaccuracy [7]–[10]. In fact, quasi-static errors are inevitable even for well-designed machine tools since the designers are not able to consider all the operating conditions [5]. Thus, the strategy of error elimination, which aims to predict and compensate the errors without removing the error sources, is crucial for machining error reduction. However, the current modeling and compensation approaches are difficult to be implemented in production environments because they are either time-consuming or not robust in parameter identification procedures. These disadvantages also cause the existing approaches susceptible to thermal variation of the machine tool and not capable of tracking the thermal errors of the machine over time.

In order to secure the quality of machining process, metrologist examines and compares the finished part with the nominal geometry to perform the compensation [11]. The examination performed by the metrologist is based on the dimension and tolerance given by the print, which are defined following the national standard, ASME Y14.5 the Geometric Dimensioning and Tolerancing (GD&T) standards [12]. However, traditional hard gage-based metrology method takes measurements at only few points on the part and may not truly representative. Therefore, the results cannot be used to fully determine the dimensioning and toleranceing requirements. A comprehensive metrology method such as point-cloud-based metrology that uses the entire surface profile of workpiece is preferable. However, the difficulties of implementing point-cloud based-metrology include:

(a) not having sufficient computational power and well-developed algorithms to process and manipulate a large point-set of size of millions points.

(b) lack of probe-based metrology instruments to provide condensed and representative measurements of workpiece surface.

Laser scanning technology allows a metrologist to depict the surface profile in a more efficient way using point-cloud data set, which is more condensed and representative for the surface. Further, many powerful and robust open-source point-set manipulation algorithms are available online. Quick and accurate optimization solvers that not only deal with linear programs but also complex nonlinear programs are also available. The advances in instrumentation as well as the computational power have opened an opportunity to overcome the difficulties involved in developing point-cloud based metrology.

1.2 RESEARCH OBJECTIVES, SCOPE, AND TASKS

This dissertation aims to generate a system by which a manufacturer, in an automated fashion, can compensate for machine tool workspace errors induced due to part, fixture, tooling, or machine tool errors, specifically of 5-axis machine tool. The specific focus will be on thermal error tracking, construction of virtual gages for displacement error compensation and cylindrical surface's tolerance verification.

The scope of this dissertation is limited to applying the newer metrology instruments (e.g. laser tracker and laser scanner), optimization techniques (e.g. design of experiments, least-square fitting, linear/nonlinear programming and particle swarm optimization) and computational geometry (e.g. nearest and farthest Voronoi diagrams) to make machine tool calibration and point-set based metrology more efficient and accurate. The proposed research objectives will be achieved through specific research tasks, with associated sub-tasks, described as follows.

1.2.1 TASK 1: MAPPING THE WORKSPACE ERRORS OF A MACHINE TOOL

This research task focuses on understanding the characteristics of the quasi-static errors of machine tools by building a kinematic error model for a 5-axis machine tool and designing a quick measuring cycle with optimized error model observers.

- Sub-Task 1: Quasi-Static Error Modeling Using Laser Tracker

A modeling approach introduced by Ferreira and Liu [7] and extended by Kiridena [8], [13], [14] is proposed to be used for the quasi-static errors of a 5-axis machine tool with one redundant axis. By introducing errors to the ideal joints and shape transforms of the kinematics of the machine, an error model will be developed. First order error characteristics will be used to parameterize the introduced errors. To identify these error parameters, the volumetric error components at randomly chosen points will be measured with a laser tracker. The unknown modeling parameters can be obtained by the least-squares estimation, and the volumetric error model of a 5-axis machine will be built.

- Sub-Task 2: Quasi-Static Error Observer Design for Thermal Error Tracking

The parameter identification procedure for identifying the parameters of a volumetric error model of a large and complex machine tool usually requires a large number of randomly-chosen observations of volumetric error components in its workspace [15], which makes the thermal variations of the volumetric error model difficult to track. Therefore, several optimal designs, including A, D and K-optimal designs will be applied to reduce the number of observations and hence reduce measurement time in the error parameter identification procedure. The feasibility of using a smaller but strategical-chosen observation set to track thermal errors will be checked by periodically repeating the measurement cycle.

1.2.2 TASK 2: POINT-SET BASED WORKPIECE METROLOGY

The goal of this research task addresses the problem of characterizing the raw casting for the purpose of (a) deciding on the acceptability of the part based on its material condition, (b) make adjustments to the machining reference coordinate system to allow the part to be successfully machined, and (c) showing the surfaces with defects if any such surface is not machinable.

- Sub-Task 1: Data Pre-Processing Algorithm

A laser scanner is used to collect the point-set data that represents the physical location and shape of the workpiece. Depending on the scanning strategy, the collected data usually have a large number of points. In this sub-task, a pre-processing algorithm that (a) roughly aligns the raw data set with the solid model of the workpiece for data sampling and referencing, (b) breaks the original data set into several data sets based on features of the solid model, and (c) removes the points, which are redundant or obviously contain a lot of noise (e.g., outliers and points locate by the edges) will be developed.

- Sub-Task 2: Construction of the Virtual Gage

In order to check if the dimensional defects of the part can be compensated by shifting and rotating the machining coordinate system, virtual gages will be placed as imaginary boundaries based on the prints and the design of the final casting [16]–[18]. The geometrical shapes formed by the imaginary boundaries will be transformed into algebraic expression as constraints of inequality, which will be used later in the optimization procedure.

- Sub-Task 3: Displacement Compensation Using Virtual Gage

The goals of this research task are: (a) checking if the workpiece has positive machining allowances for all surfaces to be properly machined; (b) estimating the displacement offsets if the workpiece does have enough material to be machined, and (c) experimentally verifying the proposed approach using a misplaced part with locating errors. Form tolerances will be defined by min-max algorithms [9], [19], [20] and linear programming problems, which seek to identify a rigid body transformation such that all constraints of inequalities generated in previous sub-task are satisfied, and all functioning features are considered equally significant. If such a rigid body transformation exists, then all surfaces to be machined will have a positive machining allowance. On the other hand, if the algorithm does not yield a solution and a rigid body transformation, it shows the workpiece has insufficient material or bad placement that is impossible to be compensated by changing the machining reference coordinate system alone. In this case, the workpiece could be rejected for the next level of machining process or efforts can be made to detect the defects.

- Sub-Task 4: Defective Feature Detection

The algorithm developed in sub-task 3 only determines if the part is machinable but is not capable of finding defective features, which cause the part cannot be machined properly. This is important for being economical since rejecting a raw casting could be costly in a more complicated machining case using complicated part because the defects might be manually compensated by adding material. In this task, the algorithm proposed in sub-task-3 is extended such that the defective features with negative machining allowances not having sufficient material to be machined can be detected.

- Sub-Task 5: Tolerance Verification for Cylindrical Surfaces

Sub-tasks 1-4 are developed for checking planar surfaces' material conditions. In addition to the planar surfaces, cylindrical surfaces are the most common features in machining process. This sub-task seeks to develop a method to verify the material conditions of a point-set, which represent a cylindrical surface. The material conditions of a cylinder are quantified by minimum and maximum radii. The cylindricity error of a discrete point-cloud is also verified. The developed numerical algorithms are based on particle swarm optimization (PSO), which searches the regional extremum of a complex surface. The computational geometry-based methods are built based on nearest and fartherst Voronoi diagrams to geometrically determine the smallest, largest and minimum zone cylinders.

1.3 THESIS OUTLINE

This thesis is organized as follows. The existing relevant research for machining accuracy enhancement is reviewed in Chapter 2. The quasi-static error is one of the most important issues in machine tool accuracy, and dominating sources of quasi-static error is thermal error. A variety of modeling and tracking techniques are reviewed and compared. The variation and displacement of the workpiece can be modeled, measured and quantified by point-set metrology. Tolerancing design and verification are two prospects of GD&T specification and are commonly used in point-set metrology. Many researchers have contributed their knowledge on these fields of study. The tolerance verification approaches are reviewed next, which include the planar surfaces and the cylindrical surfaces. Finally, the gaps in knowledge leading to the work of this dissertation are outlined.

In Chapter 3, a methodology of modeling a 5-axis machine with a redundant axis is proposed based on existing joint transformation models. The ideal and perturbed models for three types of joint transformation are used. Besides, the error model for the rotary joint is proposed to be modeled using Fourier sine series. With all perturbed joint transformation, one is able to obtain the perturbed forward kinematic chain and build the error model. Although the model contains nonlinear terms in different joints and links, it can still be linearized by eliminating higher order terms in the model. The model can thus be identified experimentally using a laser tracker by taking quasi-random measurements in the machine's workspace. After the error parameters are identified, the error model is obtained and tested using another previously unseen training set.

The concept of error model observer is introduced in Chapter 4. Since the time spent on taking measurements is too lengthy to capture the thermal error of the machine even with quick measurement instrument such as a laser tracker, one must obtain the optimal locations that represent the most informative observations. This problem can be formulated using optimization with different criteria according to A, D and K-optimal designs. The observer design problems can thus be solved using optimization solvers. The thermal errors can thus be tracked by periodically updating the error model using the data-driven approach. Finally, the thermal behavior of each axis during the experiment is analyzed and correlated with the temperature variation.

In Chapter 5, the point-set based metrology is presented. Section 5.1 provides background knowledge, followed by the definition and the mathematical model of planar virtual gage given in Section 5.2. Section 5.3 elaborates the experimental setup for the validation. Two example problems are solved and used to demonstrate the feasibility of virtual gage analysis in Sections 5.4 and 5.5.

The tolerance validation problem is extended for cylindrical surface, which is discussed in Chapter 6. Initially, a 2-D circle fitting problem is considered. The Voronoi diagram-based approaches are reviewed and tested, while the numerical method based on particle swarm optimization (PSO) shows similar results in three different circle fitting problems in Section 6.2. The 3-D cylinder fitting problem is simplified by casting projection onto the x-y plane and degenerate to 2-D circle fitting problem in Section 6.3. The modeled problems are difficult to solve since it contains nonlinearity that describes by axis orientation. Therefore, PSO is used to solve the simplified nonlinear optimization problem with only 2 degrees-of-freedom. In Section 6.4, measurement data sets available in the literature are used as example problems, and the results are compared with the results reported in literature.

The contributions and conclusions drawn for this thesis are given in Chapter 7. Several future works for the continuation of this research are also recommended in this chapter.

CHAPTER 2. LITERATURE REVIEW

In this chapter, a literature survey on research relevant to this thesis is reviewed. Sections 2.1 and 2.2 present the work related to machine tool calibration, specifically focus on the identification of quasi-static errors and the thermal errors. To improve the accuracy and efficiency on error calibration procedure, design of experiments for optimal observer design is introduced in Section 2.3. Section 2.4 summarizes the standard of workpiece metrology given by ASME Y14.5 research works related to GD&T verifications using traditional measurements. Research works using point-set to verify GD&T specifications of planar and cylindrical surfaces are given in Sections 2.5 and 2.6, respectively. Finally, the gaps in knowledge leading to the work of this thesis are described in Section 2.7.

2.1 QUASI-STATIC ERROR MODELING OF MACHINE TOOL

About 70 percent of the inaccuracy of a machine tool is caused by quasi-static errors. As their name suggest, the quasi-static errors are slowly varying errors. The sources of the quasi-static error include assembly errors, flexural errors (due to self-weight of moving parts and the work piece), and thermal deformations (due to heat generation at the spindle, drives, guideways and cutting tools as well as ambient temperature variations, all of which gradually generate the geometric inaccuracies in the underlying kinematic structure of the machine [14], [21]). Compared with dynamic errors (e.g., servo-tracking errors; dynamic response to cutting forces), quasi-static errors vary slowly during the operation of the machine. Due to being constrained, small thermal changes cause structural members of the machine to undergo deformation that, in turn, are magnified by the Abbe effect [3], [22], [23]. Therefore, thermal errors can, depending on the mode of operation and the level of control of the factory environment, become the dominant component of quasi-static errors, especially for larger machines with variable operation cycles [24]–[26].

Quasi-static machine tool errors, because of their large imprint on workpiece inaccuracy, and their slow variation, are good candidates for compensation. Vast body of literature exists on characterization and measurement of the quasi-static errors. In the following two sections, research works relevant to this thesis are presented.

A number of researchers have reported the approaches.

2.1.1 MODELING APPROACHES

Denavit-Hartenberg (DH) parameters are commonly used for describing the relation between reference frames attached to the links of a spatial kinematic chain [27]. A generalized predictive error model, considering combinations of polynomials and functions of nominal positions and temperature was proposed by Donmez et al. [28]. Ferreira and Liu [7] applied rigid body transformations with small error parameters to develop a linear volumetric error in workspace and used least-squares to estimate them. The shape transformation for general inaccurate link is modeled by:

$$T_{shape} = \begin{bmatrix} 1 & -\alpha & \beta & a + \Delta a \\ \alpha & 1 & -\gamma & b + \Delta b \\ -\beta & \gamma & 1 & c + \Delta c \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{2.1}$$

where α, β and γ are the angular rotations characterizing the rotary form errors of the link and Δa, Δb and Δc are the components of its dimensional errors along the three axial directions.

The transformation for an inaccurate prismatic joint with angular and positioning errors are modeled as linear functions of position along the joint, and the straightness errors are depicted as quadratic functions,

$$\Phi_x = \begin{bmatrix} 1 & -x\frac{d\alpha}{dx} & x\frac{d\beta}{dx} & x(1+\delta x) \\ x\frac{d\alpha}{dx} & 1 & -x\frac{d\gamma}{dx} & \frac{x^2}{2}\frac{d\alpha}{dx} \\ -x\frac{d\beta}{dx} & x\frac{d\gamma}{dx} & 1 & -\frac{x^2}{2}\frac{d\beta}{dx} \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad (2.2)$$

where x is the position along the joint, $\frac{d\alpha}{dx}$, $\frac{d\beta}{dx}$ and $\frac{d\gamma}{dx}$ are the rates of change of the angular errors with respect to position along joint and δx is a rate of accumulation of positioning error.

For machines with rotary joint, Kiridena and Ferreira [8] used perturbations to Jacobian matrix to develop error maps of different 5-axis machine configurations. On the other hand, inverse kinematics is also a common approach used to identify error components [29], [30]. To model the thermally-induced error, which causes the deformation of the structure of the machine to vary with time, a strategy using finite element analysis (FEA) coupled with temperature field measurements by thermocouples was proposed by Creighton et al. [22]. Veldhuis and Elbestawi [23] also proposed a thermal error compensation strategy based on neural network for five-axis machine, which eliminates significant error sources.

For indirect error identification, Sheth and Uicker [31] modeled the kinematic structure of the machine with shape and joint transformations. For CNC machines, one deals primarily with prismatic joints [6] and rotary joints [32], while the shape transformation primarily involves translations across links [33]. A model for the errors of the machine is constructed by introducing small perturbations (errors) into the parameters of the shape and joint transformations and an expression for the volumetric error vector is obtained as the difference in the spatial location of the tool/spindle (relative to the workpiece table) produced by the perturbed and ideal kinematics of the machine. The equation expresses the volumetric error in terms of the yet-to-be-determined unknown perturbations of the links and joints of the machine. When observations/measurements of the volumetric errors in the machine's workspace are made, optimization techniques such as least-squares fits are used to estimate these perturbations [7], [13], [34]. For example, a 3-axis machine's volumetric error can be modeled using 24 error parameters [7]:

$$\begin{bmatrix} e_{x_1} \\ e_{x_2} \\ e_{x_3} \end{bmatrix} = \begin{bmatrix} 0 & f_1 & f_2 & 0 & f_3 & 0 & f_4 & f_5 & f_6 & f_7 \\ g_1 & 0 & g_2 & g_3 & 0 & g_4 & g_5 & g_6 & g_7 & g_8 \\ h1 & h_2 & 0 & h_3 & h_4 & h_5 & h_6 & h_7 & h_8 & h_9 \end{bmatrix} \vec{x}, \tag{2.3}$$

where $\left[e_{x_1}\ e_{x_2}\ e_{x_3}\right]^T$ is the modeled error, $f_1{\sim}f_7$, $g_1{\sim}g_8$ and $h_1{\sim}h_9$ are 24 undetermined error parameters and $\vec{x} = [x_1^2\ x_2^2\ x_3^2\ x_1x_2\ x_2x_3\ x_3x_1\ x_1\ x_2\ x_3\ 1]^T$ is a vector determined by observations and machine's forward kinematics.

The error component along x_1 direction has linear relationship with error parameters, $f_1{\sim}f_7$ and can be estimated by n observations in the workspace ($n \geq 7$),

$$\vec{e}_{x_1} = \begin{bmatrix} e_{x_{11}} \\ \vdots \\ e_{x_{1n}} \end{bmatrix} = \begin{bmatrix} x_{21}^2 & x_{31}^2 & x_{21}x_{31} & x_{11} & x_{21} & x_{31} & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{2n}^2 & x_{3n}^2 & x_{2n}x_{3n} & x_{1n} & x_{2n} & x_{3n} & 1 \end{bmatrix} \begin{bmatrix} f_1 \\ \vdots \\ f_7 \end{bmatrix} = X\vec{f}, \tag{2.4}$$

where X is the design matrix.

The best estimator of $\vec{f}$ is given by the least-squares estimation,

$$\vec{f} \cong (X^T X)^{-1} X^T \vec{e}_{x_1}. \tag{2.5}$$

To ensure the observations carry sufficient information to make accurate estimates of the values of the unknown parameters, a large number of observations, (quasi-) randomly distributed across the workspace, are required [15], [35], [36]. Figure 2.1 shows a quasi-random point-set in a 5-axis machine's workspace measured by a laser tracker [15]. The collected data is used to identify the parameters in the error model.

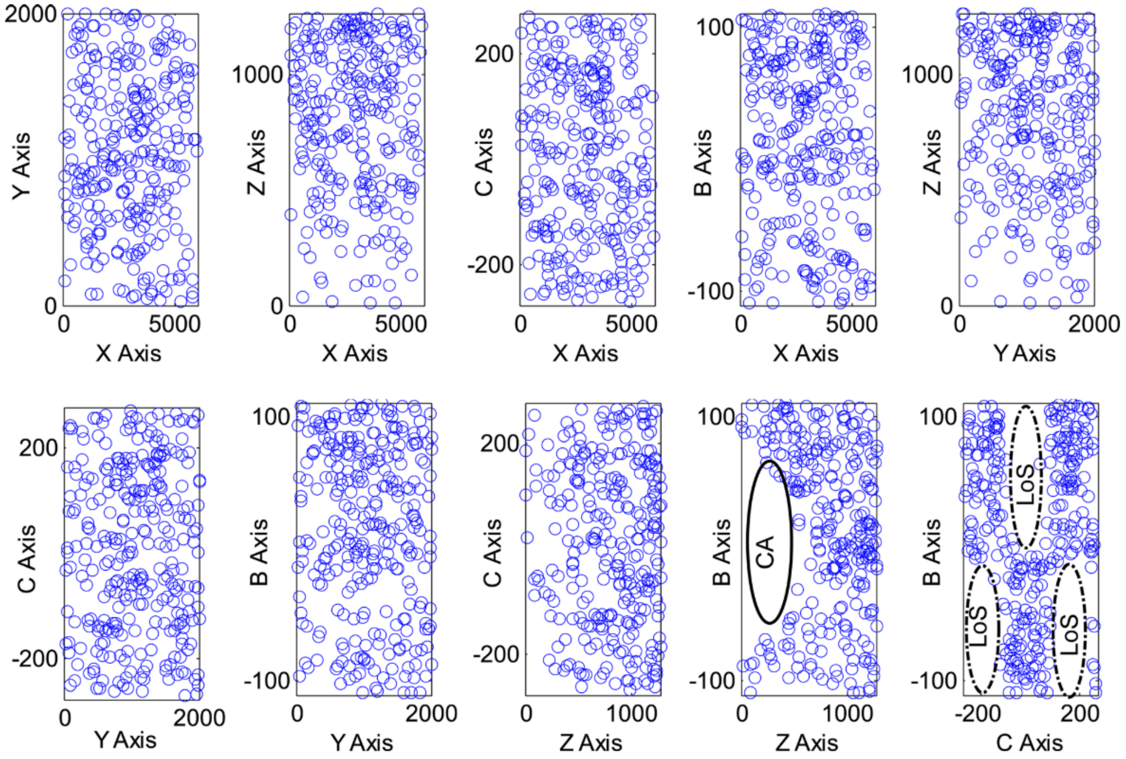

Figure 2.1: Observation point-set for error model identification of a 5-axis machine. Collision avoidance (CA) and line of sight (LOS) areas are not measured [15]

A large set of observation points enforced redundancy in the observations leads to long measurement times, even with automated devices like laser trackers. As a result, it limits the use of such a volumetric error calibration approaches to static, base-line machine tool calibrations, not addressing the changes that may occur as a result of thermal variations during the operation of the machine.

2.1.2 MEASUREMENT APPROACHES

The circular test and the use of a telescoping ball-bar, introduced by Bryan [37] in 1982 can make measurements in several planes and only requires, at its core, a short-range, high-resolution measuring device like a linear variable differential transformer (LVDT). The circular test with a telescoping ball-bar, when conducted in several planes at different locations in the machine is capable of exposing scale mismatch errors and squareness errors between the axes as well as squareness errors between the axes as shown in Figure 2.2. More importantly, this approach exposes backlash during axes reversals, dynamic contouring/gain-mismatch errors (when conducted at different speeds). Several other devices such as the laser ball-bar [38] or the grid encoder [39] are an attestation to the power of the circular test and its ability to reveal hidden characteristics of a machine. While it provides some values (backlash, scale errors) for compensation, the observations made by the ball-bar only contain the error component along the ball-bar direction. The test is not meant to provide a complete geometrical/kinematic calibration of the machine.

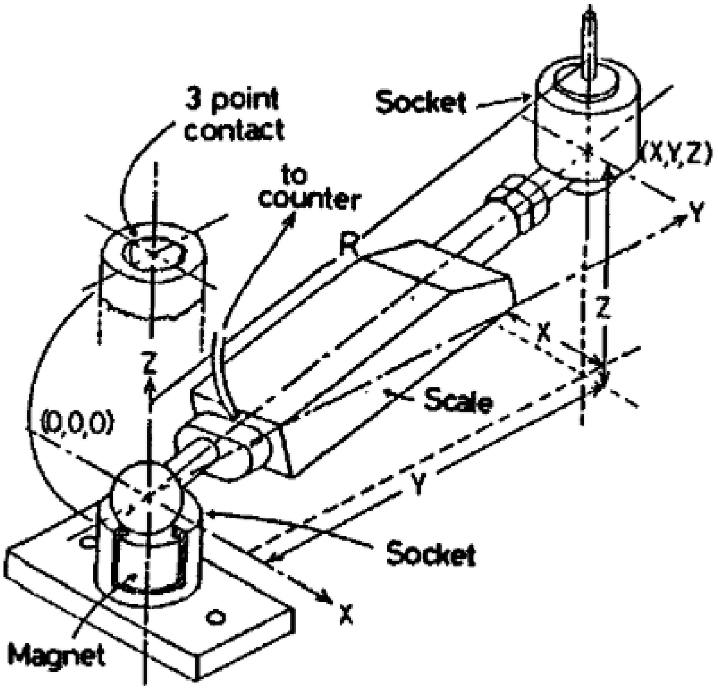

Figure 2.2: Setup of the telescoping magnetic ball bar system [40]

The ASME B5.54 [41] is a standard that establishes methodology for specifying and testing the performance of CNC machining centers. It provides a series of examinations for each axis (linear or rotary) of the machine. The calibration of the machine is built up axis-by-axis as shown in Figure 2.3. Built largely around laser interferometry measurements, each axis is calibrated while the other axes are fixed [25], [42], [43]. The tested axis is traversed along the entire extent of its (linear or angular) range with programmed periodic measurements taken by measurement system comparing the commanded to the measured position. Interferometric measurements can also evaluate relative changes in the angular and straightness errors at each measured position. While such beam-based measurements are well-suited for linear axes, laser interferometers are used along the axis of rotation to characterize (translation and angular) wandering of the center and axis of rotation. While capable of very high accuracy measurements, the use of the laser interferometer for B5.54 measurements makes it time-consuming (imposing length downtimes on productive and expensive machines) and difficult measurements requiring skill and expertise with the measurement equipment. While the B5.54 measurements provide sound axis calibration, they do not provide a full characterization of the interactions between axes of a machine. Therefore, they only provide a partial characterization of the volumetric errors of a machine and generally cannot be used to identify parameters of volumetric error models.

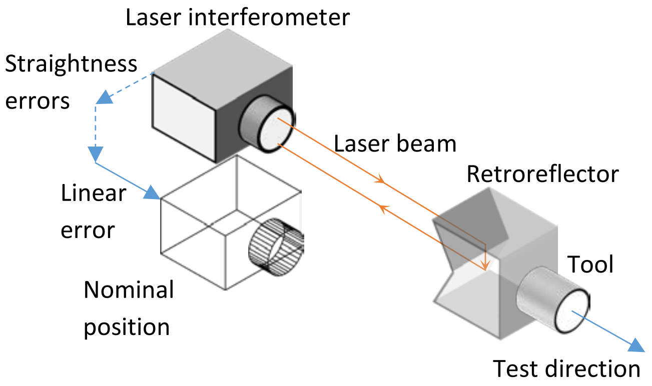

Figure 2.3: Schematic of calibrating machine tool using ASME B5.54 standard [41]

Several artifact-based methods exist for testing of machine tools. Bringmann and Küng [44], for example, designed a ball plate, an artifact that consists of an array of precision spheres and a measuring device consisting of 4 linear probes (LVDTs). As shown in Figure 2.4, the

NAS979 is an artifact that consists of a circle, diamond and square that must be cut by the machine exposing the machines linear and circular contouring capabilities, axes backlash, cornering (or lag) errors, squareness between axes and leadscrew pitch errors [45]. This is a test commonly used for acceptance testing of machine tools. Kiridena and Ferreira [10], [13] designed a grid of precision cubes and a measuring tool consisting of 3 orthogonal LVDTs for tracking the changes in the volumetric errors of a machine due to thermal effects. Several artifacts like calibrating spheres (use for ASME B89 calibration of CMMs) and rings are also use to implement the aforementioned circular tests on machine tools. Notwithstanding the convenience of having a calibration artifact, mechanical probing (measuring with a touch trigger probe) is slow and time-consuming. Further, artifacts can be expensive for absolute calibration as they must be made of low CTE (coefficient of thermal expansion) material and must be periodically calibrated themselves.

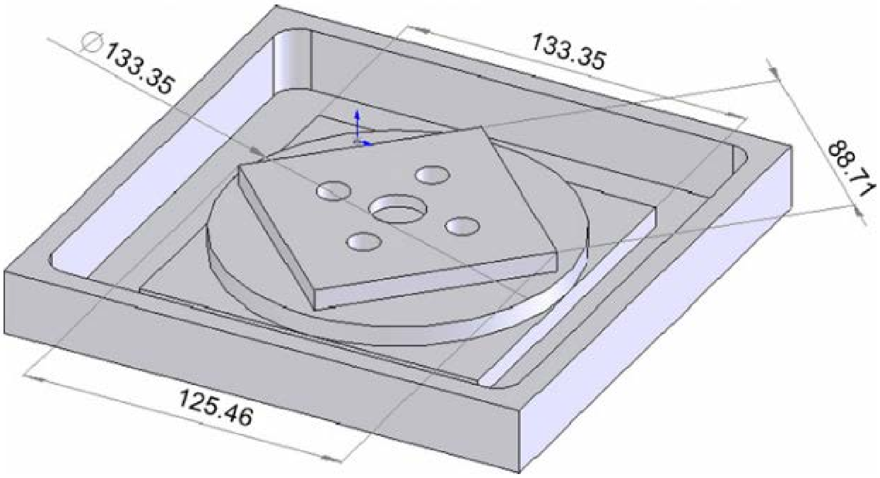

Figure 2.4: Testing artifact NAS979 [45]

On the other hand, the use of a versatile metrology instrument such as a laser tracker [3], [46]–[48] allows for a model with a large number of parameters to be identified, thus improving the effectiveness of the modeling procedure. The experiment setup for tracker-based error parameter identification procedure is shown in Figure 2.5. The ease and speed of making measurements with a tracker open the possibility of capturing the thermal drift of the machine by periodically rebuilding the error model based on sufficiently many measurements over the whole workspace taken in a short time interval. The laser tracker-based calibration provides an opportunity to revisit the quasi-static error modelling, which could be identified with a more convenient and robust approach. The reader is referred to Machine tool Metrology by G. Smith [49] for a complete description of different machine tool testing and calibration techniques.

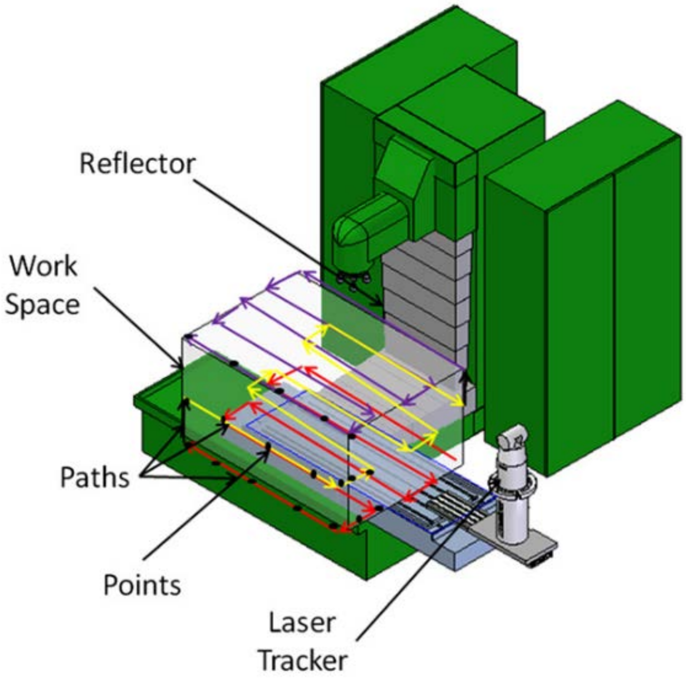

Figure 2.5: Schematic of tracker based calibration setup [46].

2.2 THERMAL ERROR MODELING

In the context of machine tool accuracy, quasistatic, or seemly static, errors account for about 70% of the observed machine errors and accrue from geometric/kinematic errors induced by manufacturing and assembly inaccuracy, flexural deformations and thermally-induced deformations, with the latter be one of the primary sources [50]. The thermal errors are considered to be the inaccuracy caused by the thermal elastic deformations of machine’s structural members [24], [50] due to changes in ambient conditions and heat input from internal sources such as spindle and axes motors, friction in the guideways, and the machining process. It is difficult to separate thermal effects from the other quasi-static error sources. Strategies to control their effects include avoidance of thermal errors. This can be achieved by operating the machine at its thermal steady state in a temperature-controlled environment but is difficult to justify in a production environment. A rich body of research exists for using thermal error models and compensating to deal with thermal errors. An adaptive learning model is proposed by Blaser et al. [51]. Finite element models (FEA) [22], [52], as shown in Figure 2.6, neural networks [53], and perturbation models [8], [10], [54] have also been proposed. Yan and Yang [55] and minimize the number of thermal sensors on a CNC turning center based on the synthetic grey correlation theory as shown in Figure 2.7. These models seek to correlate the machine’s thermal drifts with the temperature readings.

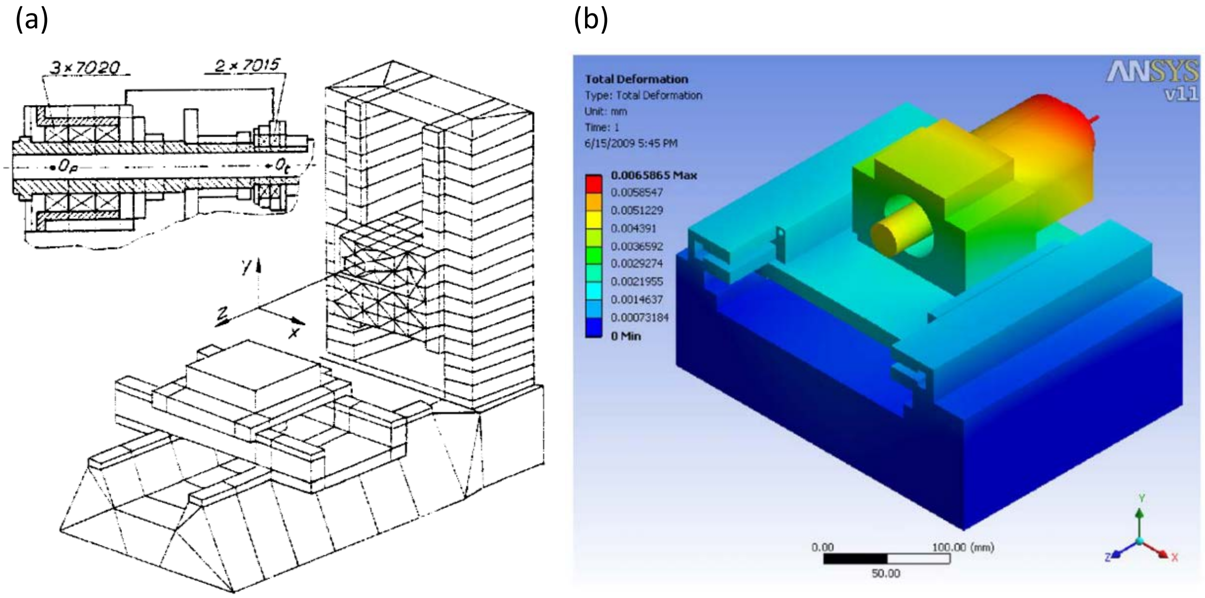

Figure 2.6:(a) Discretized geometrical structure for finite element method [52] ;(b) thermal deformation FEA simulation under thermal stresses [22]

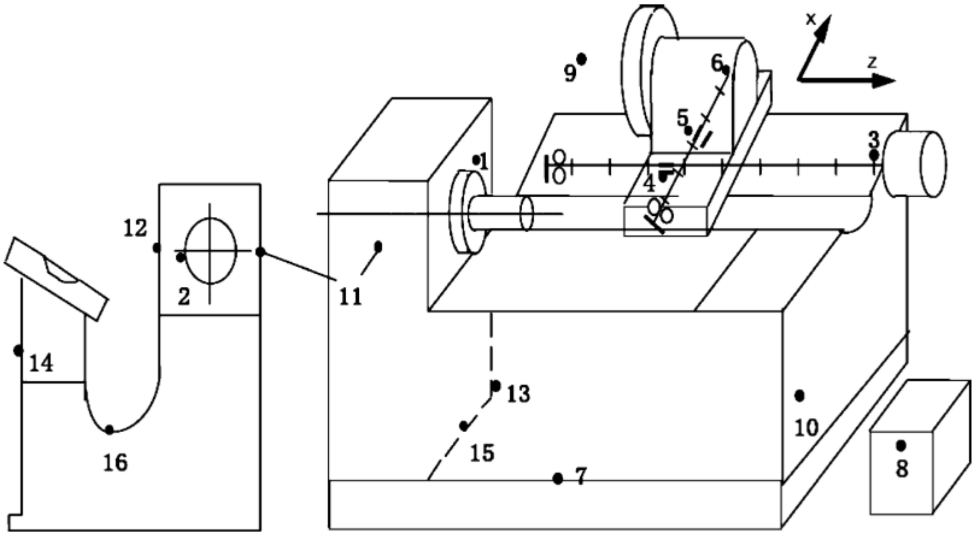

Figure 2.7: Temperature sensors placement of a turning center thermal error modeling [55]

While several approaches exist for measuring the errors of the workspace of a machine, they are time consuming (requiring extended machine down-times) and manual (typically requiring the time of both a machine operator and a metrologist). Because of the quasi-static nature of the dominant error sources, e.g., thermal and flexural deflections and an inability to efficiently update the error map at regular intervals, the effectiveness of error maps for compensating machine tool errors is often called into question. Also, the main difficulty with these modeling approaches is in characterizing the heat sources, thermal characteristics of joints and surfaces and, hence the thermal state of the machine's structural members under varying operating conditions. Data-driven approaches require large amounts of observations during the operation of the machine. Lengthy measurement processes are not possible because of the transient nature of a machine's thermal state. Therefore, even in the support of this strategy, the concept of optimal design of observations is important to identify most informative observations and shorten the length of the time interval for measurements. Some related works using the concept of optimal observer designs in the fields of sensor placement and machine tool calibration are reviewed in the following section.

2.3 OPTIMAL OBSERVER DESIGN

Design of experiment (DOE) is a discipline that addresses the question of designing observations to identify a robust model of a system. In the context of error parameter identification, the DOE question becomes one of determining where to locate points in a machine's workspace at which to measure the volumetric error components so that the error model's unknown parameters can be robustly estimated. Martínez and Bullo [36] addressed a similar problem in the determination of the best sensor location for a tracking control system, using the Fisher information matrix. A simulated annealing approach was proposed by Lin to handle the sensor placement problem by minimizing the maximum distance error in a sensor field satisfying given constraints [35]. Sensor placement algorithms that satisfy the entropy and mutual information criteria are described and demonstrated by Krause [56].

Consider a typical linear identification/design of experiments problem with n design points where a random process is considered:

$$\vec{e} = M(\vec{J}_1, \dots \vec{J}_n)\vec{p} + \vec{N}, \tag{2.6}$$

where $\vec{e} \in \mathbb{R}^n$ represents a vector of n observable values that is related to $\vec{p} \in \mathbb{R}^k$, a set of k unknown parameters is the vector consisting of all undetermined parameters, $p_1, \dots p_k$ (whose values are to be estimated) by the design matrix, $M(\vec{J}_1, \dots \vec{J}_n) \in \mathbb{R}^{n \times k}$, whose row vectors are functions of $\vec{J}_1, \dots \vec{J}_n$, sets of variables that can be independently controlled, $\vec{N} \in \mathbb{R}^n$ represents the observational noise vector with elements being random errors, normally distributed, with a mean of 0 and a variance of σ^2.

In many parameter identification/design of experiments situations, one has latitude in selecting the location of the observation/design points. Thus, the problem of selecting appropriate locations and number of design points in the space of $\vec{J}_i$ to get robust estimates of the parameter vector, $\vec{p}$ is the design-of-experiments problem or the problem of designing an observer for the model given in Equation (2.6).

The least-squares unbiased estimator of $\vec{p}$, $\hat{p}$ minimizes the sum of square errors, $\|\vec{e} - M\vec{p}\|_2$. $\hat{p}$ is also the best linear unbiased estimator (BLUE), which can be obtained by,

$$\hat{p} = (M^T M)^{-1} M^T \vec{e}. \tag{2.7}$$

As the Gauss-Markov theorem states, the variance associated with BLUE, given by the variance-covariance matrix is minimized for the design characterized by M,

$$Var[\hat{p}|M] = \sigma^2 (M^T M)^{-1}, \tag{2.8}$$

where $M^T M \in \mathbb{R}^{k \times k}$ is called the information matrix.

The variance-covariance matrix captures the uncertainty in the correlation between the elements of the estimator, $\hat{p}$. It must be noted that σ^2 is the variance of the residual, a property of the random process. So, the uncertainty in the values of the parameter vector and the predictions they make can be seen to be completely dependent on M. Because of the above considerations, a number of optimality criteria associated with different matrix norms of the design matrix, M have been proposed in both the design of experiments (DOE) and the design of observers (for continuous/on-line estimation and compensation). For a linear regression design problem with n observations and k unknown parameters to be determined ($n \geq k$), the optimal design represents the selection of n observations that carry the largest amount of information and the correspondingly, the estimator has the lowest variance.

D-optimality is the most commonly used criterion because the target function to be minimized is simpler than the other criteria [57]. A D-optimal design seeks to maximize information carried by the observations and quantified by the determinant of the information matrix. It does so by minimizing the volume of the confidence volumes or the uncertainty region around the estimator and its predictions [58], [59] with minimizing the objective function given below,

$$\min_{\vec{J}_1 \cdots \vec{J}_n} |(M^T M)^{-1}| = \min_{\vec{J}_1 \cdots \vec{J}_n} \prod_{i=1}^{k} \frac{1}{\lambda_i}, \tag{2.9}$$

where $\vec{J}_1 \ldots \vec{J}_n$ are n sets of controllable variables (in our case, the commanded axial positions) that control each row in the design matrix M, and λ_i is the i$^{\text{th}}$ eigenvalue of $M^T M$.

Similarly, an A-optimal design seeks to minimize the average variance of the estimations on the regression coefficients, and its objective is given by:

$$\min_{\vec{J}_1 \cdots \vec{J}_n} tr((M^T M)^{-1}) = \min_{\vec{J}_1 \cdots \vec{J}_n} \sum_{i=1}^{k} \frac{1}{\lambda_i}, \tag{2.10}$$

where $\vec{J}_1 \ldots \vec{J}_n$ are n sets of controllable variables (in our case, the commanded axial positions) that control each row in the design matrix M, $tr((M^T M)^{-1})$ is the trace of $(M^T M)^{-1}$ and λ_i is the i$^{\text{th}}$ eigenvalue of $M^T M$.

The K-optimality criterion which seeks to minimize the sensitivity of estimator to observation/measurement error does so by minimizing the condition number of the design matrix [60] denoted by $\kappa(M)$ which is always greater or equal to 1. It implies that the error in observation always corrupt the estimation. The condition number can be infinity if (and only if) M does not have full column rank. Consider an ordinary linear estimation system,

$$\vec{e} = M\vec{p}, \tag{2.11}$$

where $\vec{e}$ is the accurate observation and $\vec{p}$ is the correct estimation.

Now, introduce observational error, $\Delta\vec{e}$ (e.g. measurement noise and disturbance) to the system that causes errors in the estimations of parameters, $\Delta\vec{p}$ using least-squares fitting:

$$\vec{e} + \Delta\vec{e} = M(\vec{p} + \Delta\vec{p}), \tag{2.12}$$

where $\Delta\vec{e}$ is the errors in observation, $\Delta\vec{p}$ is the error in estimation due to $\Delta\vec{e}$.

Condition number of M, $\kappa(M)$ is defined by the worst-case relative error caused by the error in observations,

$$\frac{||\Delta\vec{e}||}{||\vec{e}||} \leq \kappa(M)\frac{||\Delta\vec{p}||}{||\vec{p}||}, \tag{2.13}$$

Also, $\kappa(M)$ can be defined as the ratio of largest and smallest singular values of M,

$$\kappa(M) = \frac{\sigma_{max}}{\sigma_{min}}, \tag{2.14}$$

where σ_{max} and σ_{min} are the largest and smallest singular values of M respectively.

By definition of condition number given above, the K-optimal design can be formulated to minimize the objective function:

$$\min_{\vec{j}_1 \cdots \vec{j}_n} \kappa(M) = \min_{\vec{j}_1 \cdots \vec{j}_n} \frac{\sigma_{max}}{\sigma_{min}}. \tag{2.15}$$

It must be noted that all eigenvalues of $M^T M$ are non-negative and real, and the singular values of $M^T M$ are obtained by taking square root of the eigenvalues of $M^T M$. Thus, K-optimal design can be written as the following eigenvalue design problem, which is similar to D- and A-optimal design problem formulations specified in Equations (4.3) and (4.4):

$$\min_{\vec{j}_1 \cdots \vec{j}_n} \kappa(M) = \min_{\vec{j}_1 \cdots \vec{j}_n} \frac{\lambda_{max}}{\lambda_{min}}, \tag{2.16}$$

where λ_{max} and λ_{min} are the largest and smallest eigenvalues of the information matrix, $M^T M$.

In the field of machine tool calibration, the design of the error observers involves the determination of where to locate of observations of volumetric error in the machine's workspace. Kiridena and Ferreira [10] proposed a greedy algorithm for selecting a sub-set of points on a calibration artifact as shown in Figure 2.8(a). The linear identification procedure using all 81 possible observations is given by,

$$\vec{e}_{81} \cong M\vec{p}, \tag{2.17}$$

where $\vec{e}_{81} \in \mathbb{R}^{81}$ is the vector of all observed error components, $M \in \mathbb{R}^{81\times 17}$ is the design matrix and $\vec{p} \in \mathbb{R}^{17}$ is the error parameter vector.

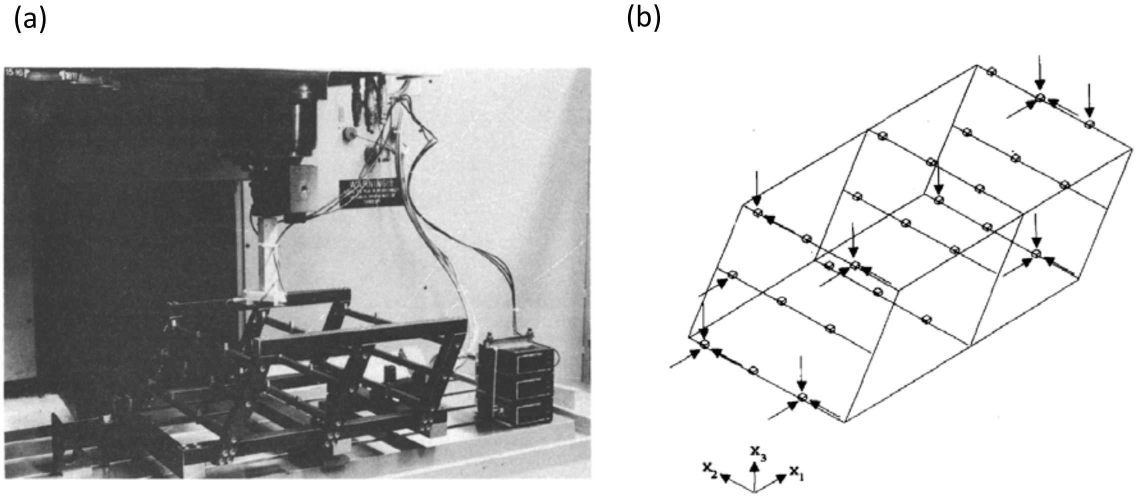

Figure 2.8: (a) Experimental setup of probe-based thermal error modeling (27 grids of measurements); (b) 17 optimized measure locations given by the greedy algorithm [10]

The proposed algorithm seeks to pick 17 most valuable observations (rows) over 81 possible ones to identify 17 error parameters. The 17 rows in M becomes B, a submatrix of M. Since the there are $\frac{81!}{17!\times(81-17)!}$ possibilities, method of exhaustion is not possible. The greedy algorithm seeks a local minimum of the condition number of the 17 by 17 submatrix of M. The optimized observation points are shown in Figure 2.8(b). The following pseudocode is used in the algorithm [10],

```
Given $M = \begin{bmatrix} B_{m\times m} \\ NB_{(n-m)\times m} \end{bmatrix}$, minprog=-$10^6$, maxproj=0
While (minproj<maxproj)
    minproj=$10^6$
    For i←1,n
```

```
        $P = [I - C^+C]$
        If ($mag(Pb_i) < minproj$)
            $minproj \leftarrow mag(Pb_i)$
            $P_m \leftarrow P$, $b_{min} \leftarrow b_i$
        endif
    endfor
    $maxproj = 0$
    For $i \leftarrow 1, m-n$
        If ($mag(P_m nb_i) > maxproj$)
            $maxproj \leftarrow mag(P_m nb_i)$
            $nb_{max} \leftarrow nb_i$
        endif
    endfor
    If ($maxproj > minproj$)
        $b_{min} \Leftrightarrow nb_{max}$
    endif
endwhile
```

In the greedy algorithm, C is a sub-matrix formed by removing one row, from b_i the basis. $C^+ = C^T(CC^T)^{-1}$ is the Moore-Penrose pseudoinverse of C, which is a rectangle matrix, P defines the null space of C and Pb_i is the projection of b_i into the null space of C. Two For loops first determine the best row to be removed from the basis and select one from the non-basis rows to replace it. Thus, the identification system of 17 optimized observations is given by,

$$\vec{e}_{17} \cong B\vec{p}, \tag{2.18}$$

where $\vec{e}_{17} \in \mathbb{R}^{17}$ is the vector of all observed error components, $B \in \mathbb{R}^{17\times17}$ is the square design matrix and $\vec{p} \in \mathbb{R}^{17}$ is the error parameter vector is now estimated by $\vec{p} \approx B^{-1}\vec{e}_{17}$.

2.4 GEOMETRIC DIMENSION AND TOLERANCE (GD&T)

Geometric Dimensioning and Tolerancing (GD&T) is widely used to describe the nominal designs and the maximum acceptable variations from the nominal ones. The conventions of geometric tolerances follow the American Society of Mechanical Engineer (ASME) standard Y14.5 [12] and related International Standards Organization (ISO) standards. There are several prospects to understand GD&T. First, the tolerances are directly affect the machining accuracy and hence the cost. Therefore, the designers control tolerances as designable variables that quantify the worst-case variability of the part. A rich body of research has been done from this point of view. X. Zhao et al. [61] proposed a model that supports integrated measurement processes by combining ASME Y14.5M-1994 [12] with Dimensional Measuring Interface Standard (DMIS) and Standard for the Exchange of Product model data (STEP). Turner and Wozny [62], [63] used numerical approaches including Monte Carlo method and linear programming to design the tolerance variables for part assemblies. To optimize the design of tolerance variables and model geometric tolerance, tolerance map (T-map) was proposed and utilized [64]–[66] as can be seen in Figure 2.9(a) and (b). Chen et al. [67] studied the advantages and disadvantages of four 3D tolerance analysis methods including T-Map, matrix model, unified Jacobian–Torsor model and direct linearization method (DLM). Menq et al. [68] presented an approach for aligning measurement data with the CAD model based on least-square fitting technique with the application of error comparative analysis. Marziale and Polini [69] compared two different tolerance modeling methods, vector loop (as shown in Figure 2.9(c)) and matrix.

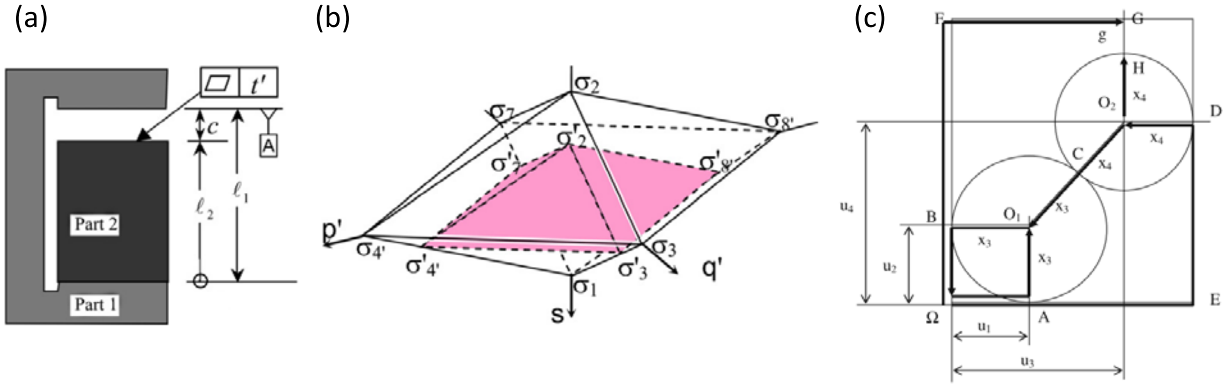

Figure 2.9: (a) Cross-section of an assembly with clearance c [64]; (b) tolerance zone analysis of (a) using T-Map [64]; (c) assembly variables and tolerances of vector loop model [69]

From the metrologist and manufacturer’s prospect, finished parts have to be examined and compared the finished part with the nominal geometry to perform the compensation to secure the quality of machining process [11]. When large castings are finished by machining processes, it becomes necessary to be highly adaptive. Large dimensions (here, of the order of 1000 to 2000 mm) cause the magnification of errors that accrue from shrinkage and warpage. Further, similar large dimensions of the structural and transmission elements of the machine tools used to machine such castings magnify the effects of small temperature changes and flexural deformations, giving rise to large quasi-static thermal and flexural errors in the work volume of the machine tool. The former can have magnitudes in the millimeter range and can quickly consume any allocated allowances [70], while the latter can grow to be several hundred microns and can easily exceed tolerances specified on critical machined features on the casting [71]. Furthermore, the placement of the castings on the fixture might be imperfect [3]. As a result, using only nominal, uncompensated NC programs and datum frames defined on machines and fixtures can lead to high levels of rejects and low yields. The castings tend to be of complex geometries and are not quickly and easily manufactured when out-of-tolerance conditions arise. Given the cost of producing such castings and the limited capacity to do so, it becomes necessary to adjust both, the datum frames and the machining programs to the geometry of each casting to achieve an acceptable, within-tolerance finished workpiece. Besides, incoming stock from casting and forging suppliers can vary to the point that standard machine tools cannot adequately respond the existing material condition in the as-programmed state. Due to the machine tool’s inability to dynamically respond to material stock variation, broken tooling, scrap parts, and severe delays occur in the entire value stream. The manufacturing community has attempted to solve this problem through part probing in the machine tool, programming sub-routines in the controller, or through manual adjustments made by the machinist. The approach yielded a sub-optimal process that requires significant human intervention and does not guarantee a conforming part.

2.5 TOLERANCE VERIFICATION FOR PLANAR SURFACES

The examination performed by the metrologist is based on the dimension and tolerance given by the print, which are defined following the national standard, ASME Y14.5 the Geometric Dimensioning and Tolerancing (GD&T) standards [12]. Tolerance verification can be done by evaluation of form errors, which requires measurements using coordinate measuring machines (CMMs). The probe-based CMMs are commonly used to verify the tolerance by taking discrete measurements on surface. However, the limited time on measuring limits the number of measurements, which has significant influence on the accuracy of the evaluation. A surface profile is within tolerance if the deviation at any point on the surface is within the specified bound. Cases may occur when sampled deviations are within bound, whereas non-sampled deviations are in fact out of tolerance [68]. Thus, researchers have proposed many measurement and sampling strategies for choosing the most information locations for evaluating geometric tolerance. A statistical analysis is used to determine the number of required points for surface profile measurement and tolerance specification [68]. Colosimo et al. [72] used a regression-based tolerance interval approach to optimize the sample locations. Summerhays et al. [73], [74] built a Chebyshev/Fourier model to select the most informative for estimating form errors on internal cylindrical surfaces. An adaptive sampling strategy based on surface patches' Gauss curvatures was proposed by Obeidat and Raman [75] to obtain the optimum number of measurements. Badar et al. [76] suggested that reducing the sample size and number of measurements using Tabu search and a hybrid search can maintain comparable accuracies on flatness evaluation. Carr and Ferreira proposed a methodology on tolerancing validation that transfers the form error problem into a linear programming problem (LPP) using point-set data [19], [20] as shown in Figure 2.10(a). The flatness verification algorithm for a set of data points P that represents the surface of a tolerance planar feature using linear program is given by [19], [77],

$$\min_{T}(d_{max} - d_{min}),$$

subjected to:

$$d_{min} \leq T\vec{p}_i \leq d_{max} \text{ for } i = 1\sim n, \qquad (2.19)$$

$$T_x^2 + T_y^2 + T_z^2 = 1,$$

where $\vec{p}_i = \left[p_{ix}\ p_{iy}\ p_{iz}\right]^T$ is the ith coordinates data point in point-set P, $T = \left[T_x\ T_y\ T_z\right]^T$ is the zone orientation direction vector, n is the number of points in P and d_{max} and d_{min} represent the farthest and closest distance from any point to the plane of the tolerance zone.

The general tolerance verification stated in Equation (2.19) is solved as a sequence of linear program as shown in Figure 2.10(b).

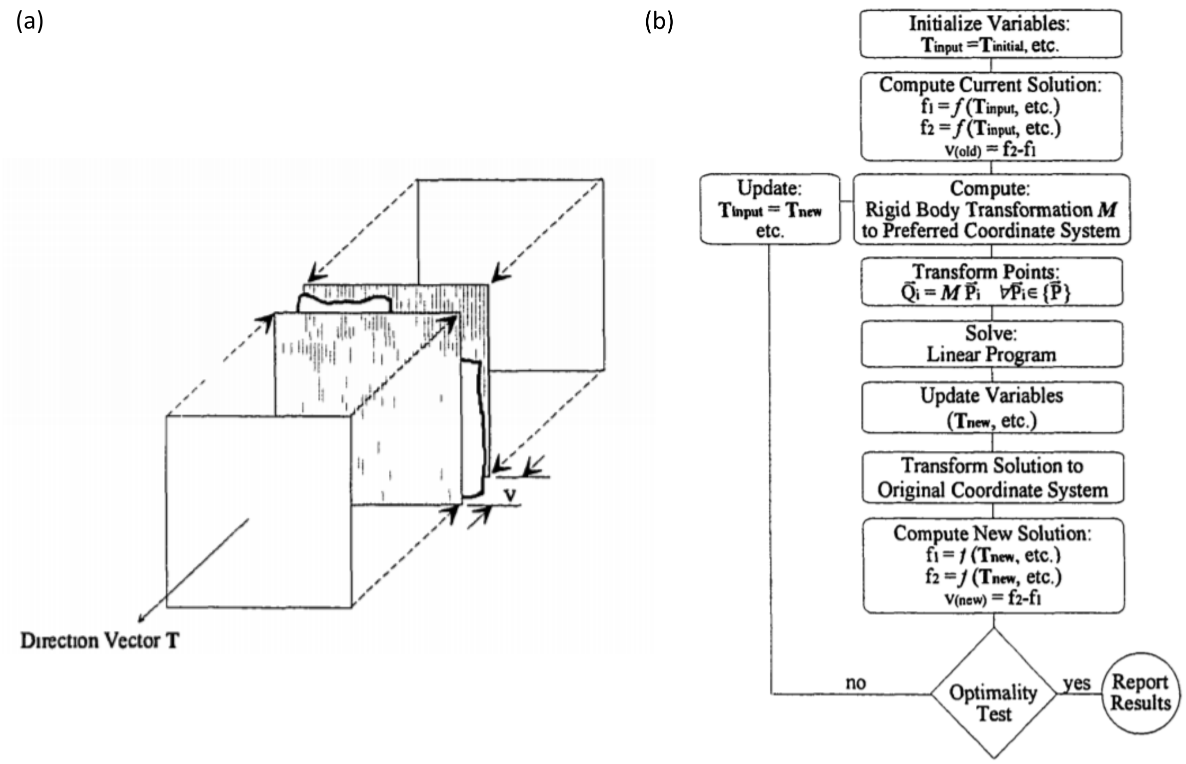

Figure 2.10: (a) Geometric model of flatness verification [77]; (b) Inspection flowchart of form and size tolerance specifications [77].

From the prospect of product quality, customers would prefer the machined surface to have shiny finish even though it is not a direct indicator of cleanliness per se [78]. Therefore, the virtual gage, defined by the boundary of a tolerance zone is used to check the tolerance model and the machinability of features virtually [16]–[18], [79]. The construction of the virtual gage provides a different point of view to geometric variation. If there exists intersection(s) between the real surface and the virtual gage, the part should be rejected because some surfaces will not satisfy the assigned GD&T design. It opens an opportunity to understand and compensate the machining errors caused by geometric variation and the fixture errors. To provide a pre-process, in-situ conformity test for the raw casting, a metrology instrument, which depicts the surface profile of the workpiece, is required. Using touch trigger probe is accurate but takes a long time to probe all the critical features on the workpiece. Alternatively, a laser scanner that combines controlled steering of laser beams with a laser rangefinder is able to depict the surface profile by taking a distance measurement of the surface shape of the scanned object [80]. By combining multiple surface models, which can be obtained from different scanning paths, a full 3D model of the object can thus be constructed [81]. Open computer vision and image processing source libraries providing basic point-cloud operations including surface reconstruction, model segmentation and point-set manipulation of large size of point-cloud are easy to access today [82].

2.6 TOLERANCE VERIFICATION FOR CYLINDRICAL SURFACES

Without a good algorithm, metrologist could not properly process a larger data sets collected by CMMs, which may lead to overestimation of the tolerance, reject the acceptable parts and increase the cost. Hence, a quick and accurate algorithm that processes the data set and analyzes the conformity of the workpiece is critically needed. Planar and cylindrical surfaces are two of the most common surfaces in machining, and there exists a rich body of research discussing how to verify the tolerance of cylindrical surfaces.

To begin with, a 2-D data set representing a circular part is considered. According to ASME Y14.5 [12], the most common tolerance specifications of a circular part include maximum and minimum radii and roundness error. The verifications of these three tolerance specifications can be done in numerical or computational geometry-based approaches. Numerical techniques including Monte Carlo, simplex and spiral search were tested in Murthy and Abdin’s work [83]. Wen et al. [84] proposed a genetic algorithm to verify circularity errors. Xiuming and Zhaoyao [85] found the lines for maximum circumscribed circle(MC) and minimum inscribed circle (MI) iteratively based on convex hull in polar coordinate to identify the control points that determine the roundness error. The computational geometry-based methods, on the other hand, seek to find the key geometries, minimum circumscribed circle (MC), maximum inscribed circle (MI) and minimum zone circle (MZ) in this case by geometrically finding the control points of them. Although the formulations of three problems are similar, the difficulties of fitting three types of circle are different. Finding MC of a given point-set is the easiest one since the optimization is convexly constrained. The most effective algorithm was proposed by Welzl [86]. Roy and Zhang [87], [88] proposed a computational geometric model based on nearest and farthest Voronoi diagrams to compute three pairs of concentric circles with the minimum radial separation under three cases (case 3+1, 1+3 and 2+2) as shown in Figure 2.11. The roundness error is defined by MZ, the smallest amount of separation among three pairs of concentric circles.

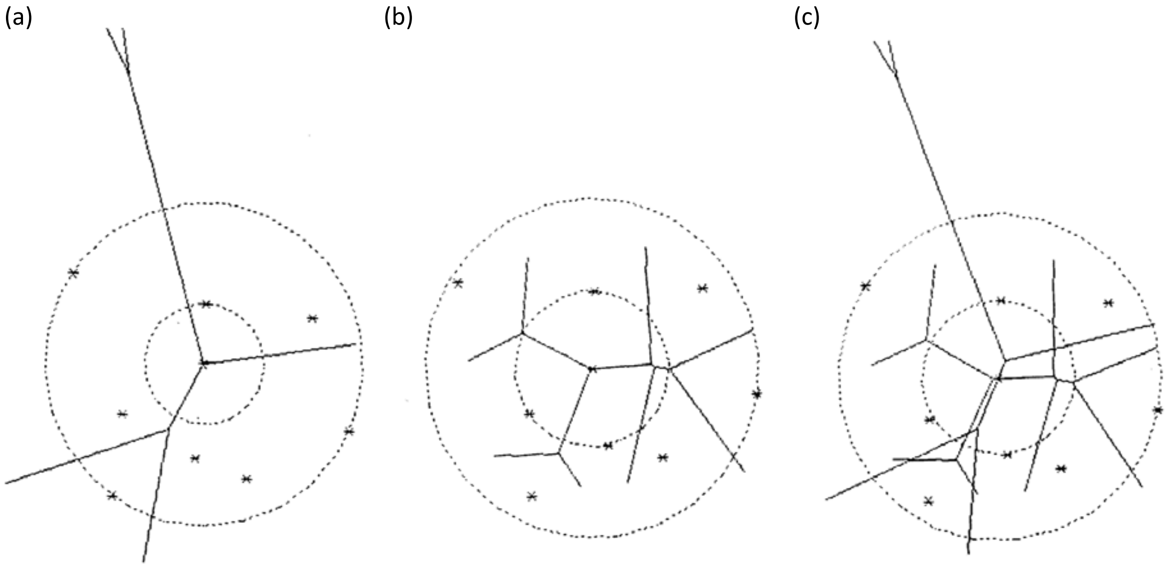

Figure 2.11: Three cases of roundness error: (a) case 3+1 by farthest Voronoi diagram; (b) Case 1+3 by nearest Voronoi diagram; (c) Case 2+2 by superimposing Voronoi diagrams [87]

Voronoi diagrams, by definition given by Okabe et al. [89], is "given a set of two or more but a finite number of distinct points in the Euclidean plane, we associate all locations in that space with the closest member(s) of the point-set with respect to the Euclidean distance. The result is a tessellation of the plane into a set of the regions associated with members of the point-set. We call this tessellation the planar ordinary Voronoi diagram generated by the point-set, and the regions constituting the Voronoi diagram ordinary Voronoi polygons."

Let $P = \{p_1, \dots, p_n\} \subset \mathbb{R}^2 (2 \leq n < \infty)$ be a finite point-set. The i$^{\text{th}}$ nearest Voronoi region (polygon for 2-D point-set), $V_N(p_i)$, associated with the i$^{\text{th}}$ point p_i in P is defined by [89],

$$V_N(p_i) = \left\{x \mid \|x - x_i\| \leq \left\|x - x_j\right\| \forall j \neq i\right\}. \tag{2.20}$$

The nearest Voronoi diagram is given by the set of all n Voronoi polygons,

$$V_N(P) = \{V_N(p_1), \dots, V_N(p_n)\}. \tag{2.21}$$

Similarly, the definition of the farthest Voronoi polygon, $V_F(p_i)$, associated with p_i is given by,

$$V_F(p_i) = \left\{p \middle| d(p, p_i) \geq d\left(p, p_j\right), p_j \in P \backslash \{p_i\}\right\}, \tag{2.22}$$

Or equivalently,

$$V_F(p_i) = \left\{p \middle| d(p, p_i) \geq \max_j \left\{d\left(p, p_j\right), p_j \in P \backslash \{p_i\}\right\}\right\}, \tag{2.23}$$

The farthest Voronoi diagram is given by the set of all n farthest Voronoi polygons,

$$V_F(P) = \{V_F(p_1), \dots, V_F(p_n)\}. \tag{2.24}$$

Figure 2.12 shows a nearest and a farthest Voronoi diagrams. The readers are referred to the book by Okabe et. al [89] for more detailed properties and applications of Voronoi diagrams. The concept of using Voronoi diagrams to define roundness error was extended by Kim et al. [90] using pixel maps for larger point-set. Besides, equi-angular diagrams using the inner hull were proposed by Samuel and Shunmugam [91] to find maximum inscribing and minimum zone circles. Liu et. al [92] applied intersecting chord method to develop sub-models, the 2+1 and 1+2 models and generalized the 3+1 and 1+3 models.

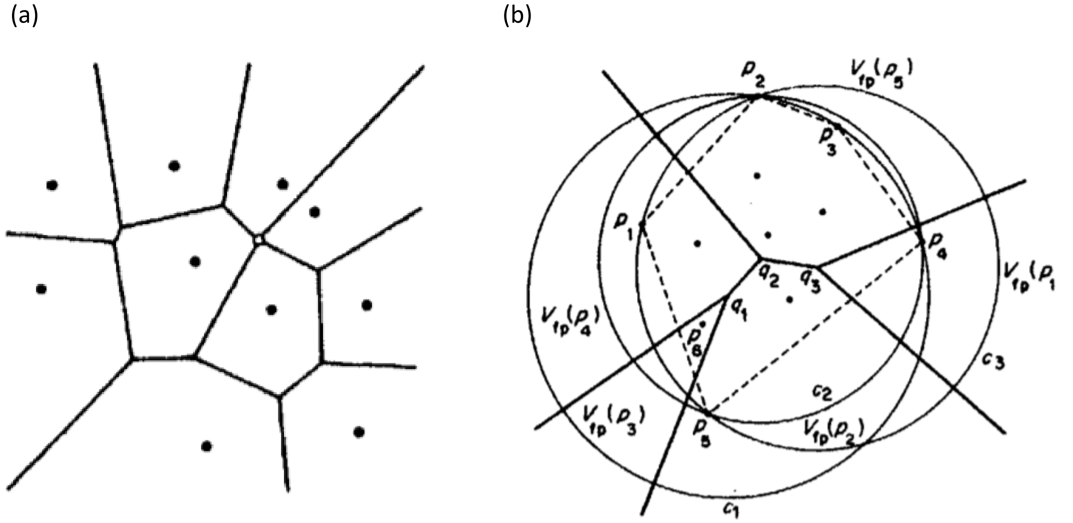

Figure 2.12: (a) A nearest Voronoi diagram; (b) a farthest Voronoi diagram [89]

In 3-D cases, minimum inscribed cylinder (MIC), maximum circumscribed cylinder (MCC) and minimum zone cylinder (MZC) play important roles in general tolerance (plus/minus) specification and form errors of part with cylindrical surfaces [77]. For example, the tolerance specification of a pins and holes is usually given by minimum, maximum acceptable radius and cylindricity, which can be verified by fitting the MIC, MCC and MZC respectively. As the matter of fact, finding the MIC, MCC and MZC of a point-set can all be modeled by optimization problems. However, these optimization problems cannot be solved easily since they have nonlinear rotation terms. Moreover, the non-differentiable objective functions due to discrete point-set as shown in Figure 2.16 make them impossible to be solved by traditional differentiation-based optimization techniques [93], [94]. Carr and Ferreira [19], [20] presented their approaches using sequential linear programming to remove the nonlinearity. The minimum circumscribed fit size tolerance of a point-set P is given by,

$$\min_{\vec{L},T} \max_{i}(d_i), \tag{2.25}$$

where $\vec{L}$ is locating point vector, T is axis direction vector and d_i is the distance between the ith point in P and the axis of cylinder as shown in Figure 2.13.

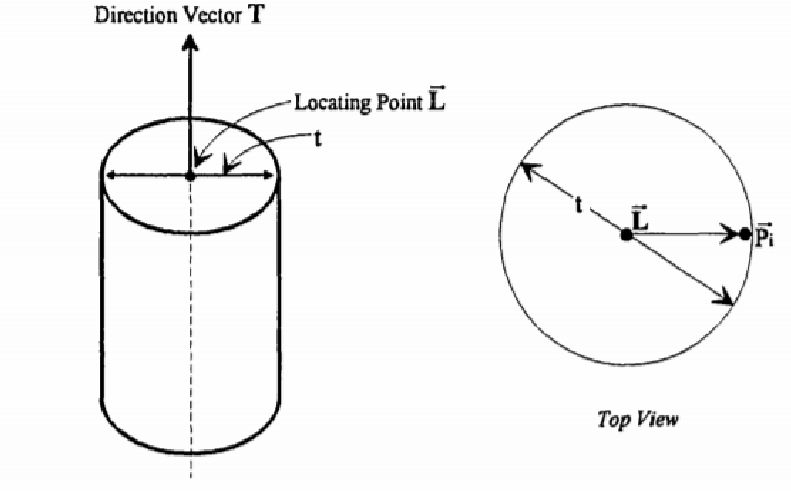

Figure 2.13: Cylinder tolerance zone: reference feature is the axis [77]

Equation (2.25) can also be written as the constrained minimization model,

$$\min_{\vec{L},T} d \ni d_i \leq d \forall i = 1 \sim n. \tag{2.26}$$

It can be seen that the optimization shown in Equation (2.26) is subjected to n linear constraints. Similarly, maximum inscribed fit size tolerance can be formulated by the unconstrained and constrained maximization models,

$$\max_{\vec{L},T} \min_{i} (d_i). \tag{2.27}$$

$$\max_{\vec{L},T} d \ni d \leq d_i \forall i = 1 \sim n. \tag{2.28}$$

Figure 2.14 schematically shows the geometric model of using coaxial cylinder to model cylindricity error. The form tolerance of cylinder can be written as,

$$\min_{\vec{L},T} \max_{i,j} (r_{max} - r_{min}), \tag{2.29}$$

where $r_{max} = \max_{i} d_i$ and $r_{min} = \max_{j} d_j$.

$$\min_{\vec{L},T}(r_{max} - r_{min}) \ni r_{min} \le d_i \le r_{max}. \tag{2.30}$$

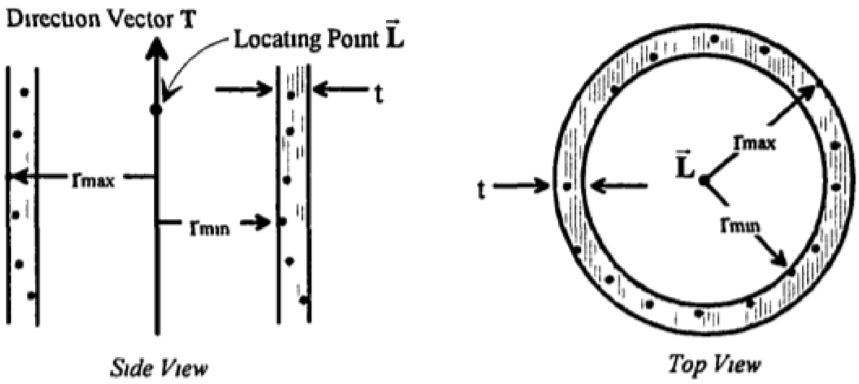

Figure 2.14: Two coaxial cylinders tolerance zone [77]

Equations (2.26), (2.28) and (2.30) are nonlinear programs since their objectives are distances from a given point to an arbitrary axis, specified by vectors $\vec{L}$ and T. Sinusoidal terms of elements in $\vec{L}$ and T make the programming problems difficult to be solved. But, these problems can be linearized and solved with linear programing solver by making the assumptions that all variables in $\vec{L}$ and T are sufficiently small. The orientation and location of the axis can thus be sequentially updated and finally converges to an optimum.

However, the sequential linearization approach relies on an initial guess close to an optimum [20] but still does not guarantee the convergence to global optimum. Cheraghi et al. [94] simplified the three cylinder fitting problems by projecting all data set onto the x-y plane. The MC, MI and MZ of the projected 2-D point-set were computed and updated by small rotation perturbations, and the flowchart is shown in Figure 2.15. The merit of this approach is that the translational variables representing the location of the cylinder axis no longer matter, hence the number of optimization variables can be decreased by two. However, the objective functions of fitting the MIC, MCC and MZC of a point-set have multiple local optimums within small region as shown in Figure 2.16, the stopping criteria of the sequential linear programming are easily to be achieved when the objective reaches a local optimum. Searching for the global optimum of the objective functions is thus a better strategy than solving the nonlinear optimizations sequentially. Hence, many bionics intelligent algorithms have been exploited in the field of tolerance verification. For example, generic algorithm (GA) was applied to verify tolerance including roundness, cylindricity and straightness [84], [95], [96]. Xianqing et al. [97] used geometry optimization searching algorithm (GOSA) to evaluate cylindricity iteratively. Wen et al. estimated sphericity error objective function using immune evolutionary algorithm (IEA) [98], and applied particle swarm optimization (PSO) to minimize conicity and cylindricity [99].

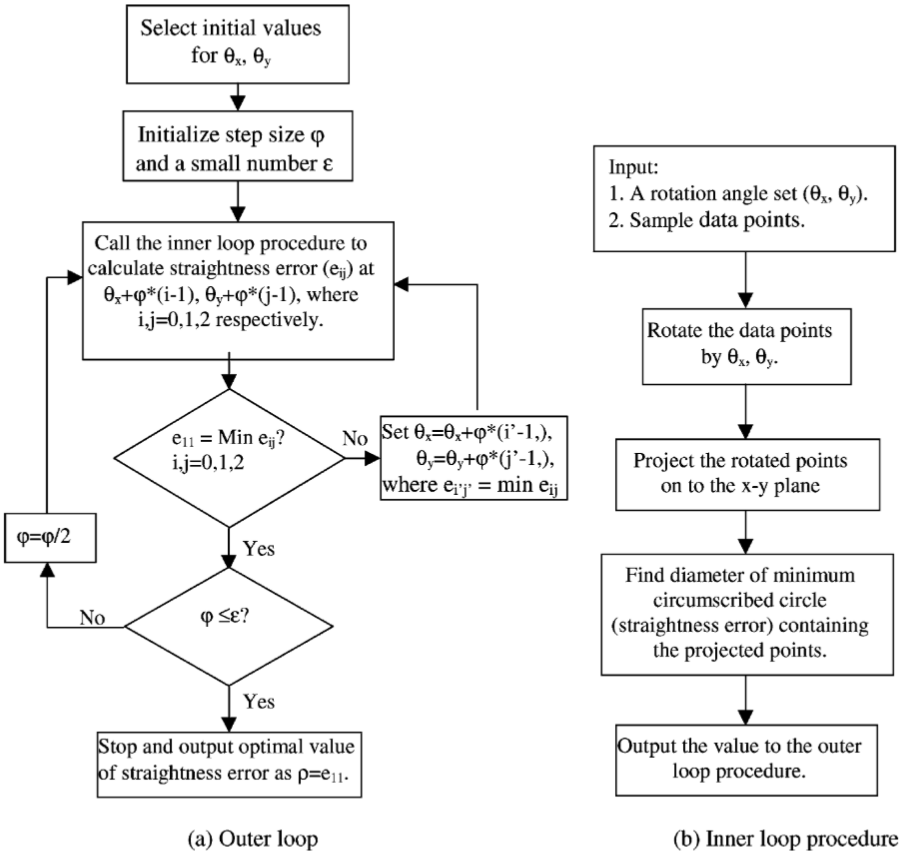

Figure 2.15: Cylindricity error evaluation flowchart using perturbation iteration [94]

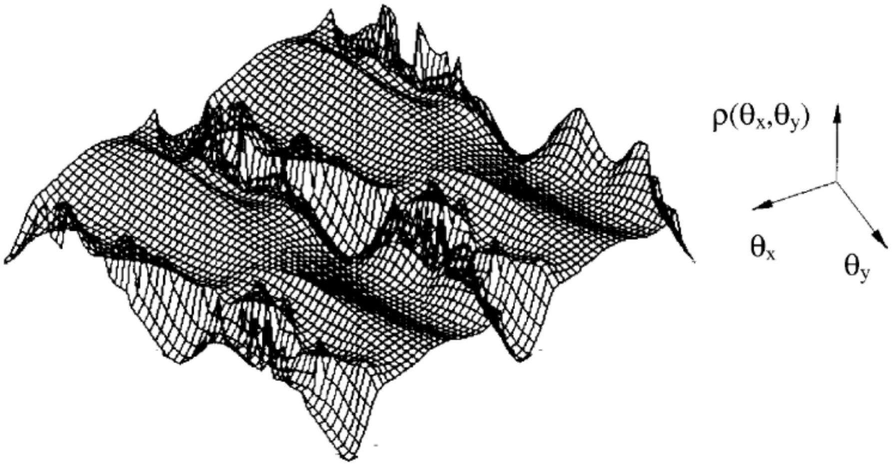

Figure 2.16: Non-differentiable objective function of minimizing cylindricity error [94]

2.7 GAPS IN LITERATURE

Error modeling of machine tool has been extensively studied since 1980’s. However, the error model of the rotary joint has not been modeled using Fourier sine series, which could be a good basis for modeling nonlinear but sinusoidal-like function. Further, literature reveals that researchers have not modeled a machine with a redundant axis since one might expect that a redundant axe would introduce parameters that would be confounded with each other. Also, a versatile metrology instrument such as a laser tracker, which allows a model with many parameters to be identified has not been used in conjunction with the modeling approach and the identification of the model. The ease and speed of making measurements with a tracker opens the possibility of capturing the thermal drift of the machine by periodically rebuilding the error model based on sufficiently many measurements over the whole workspace taken in a short time interval.

As the matter of fact, using laser track saves time spending on taking measurements for data-driven error modeling approach and makes data-driven approach more efficient. Nevertheless, a large (quasi-)randomly generated point-set that takes considerable time to be finished is still required for securing robust estimations. Due to the transient nature of thermal error, thermal error tracking becomes difficult using data-driven approach. Also, the periodical calibration procedure cannot be implemented in the production environment if a single calibration cycle is time consuming, which will greatly shorten the production time. Hence, choosing the most informative observations to shorten the length of time interval for measurements becomes a vital issue. The question of how to optimize the locations of observation for machine tool error modeling over the entire command space has not been fully addressed due to its complexity and dependency on higher computation speed. Therefore, a generalized optimization sequence based on DOE theories to optimize the error observation cycle with sufficiently many measurements in a shorter time interval is needed.

Even a rich body of research based on ASME Y14.5 has been published including traditional measurement and point-set based approaches, most of the research focus on the tolerance verification for single feature. The question of how to use point-set data of a complex part to verify the conformity remains unsolved. There is clearly a gap in current body of research regarding the problem that considers multiple complicated features simultaneously. Thus, robust algorithms including point-set manipulation and metrology for verifying multiple features' conformity are necessary.

In the problems of finding minimal, maximal radii and the roundness error of a given 2-D point-set, both geometry-based and numerical approaches can fit the circles. Maximum inscribed circle is not well-defined since the corresponding maximization problem can be unbounded. Hence, the definition should be given properly. In the 3-D case of cylinder fitting, some researchers have reported their works using different strategy to deal with the nonlinearity of axis orientation. The iterative approaches require an initial guess, which is sufficiently close to the global optimum, or the approaches are only able to provide local optima. Thus, the approaches of using intelligent searching algorithms to find the global optimum would be preferable. However, the efficiency and accuracy of such intelligent searching algorithms depend on the complexity of the objective function. There is still some room to improve by reducing the number of variables in the objective functions, so the efficiency and accuracy can be enhanced.

CHAPTER 3. MACHINE TOOL QUASI-STATIC ERROR MODELING

A comprehensive error modeling approach for machine tool is proposed in this chapter. In Section 3.1, three typical types of rigid body transformation used in machine tool error model are introduced. A 5-axis CNC machine with one redundant axis is used as example, and the linear error model is built. In Section 3.2, the error parameters in error model are identified experimentally using a laser tracker. The model is tested using another smaller but previously unseen point-set, while Section 3.3 presents a summary of the work done on quasi-static error modeling.

3.1 ERROR MODEL DERIVATION

A schematic of the 5-axis machine used in this study and its kinematic equivalent are shown in Figure 3.1. The travel of prismatic joints, X, Y, Z and W axis are 4 m, 2.5 m, 2.2 m and 800 mm respectively. The rotary joint, B axis allows the table to rotate about Y-axis by 360 degrees.

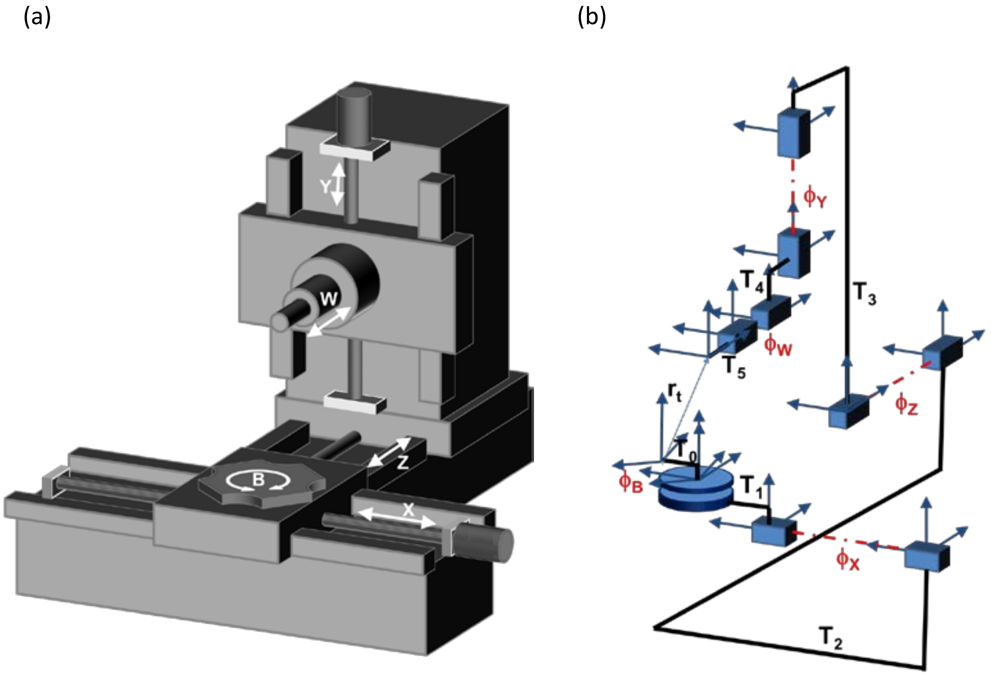

Figure 3.1: (a) Schematic of a 5-axis machine; (b) kinematic model showing the shape and joint transformation

The ideal kinematic of the machine from the table to the spindle can be expressed by the series of homogeneous transformation matrices (HTMs). This series consists of alternating joint and shape transformations. Joint transformations, denoted by $\Phi_B, \Phi_x, \Phi_z, \Phi_y, \Phi_w$, model the constraints and degrees of freedom of the transmission elements or joints of the machine, while shape transformations model the geometry and dimensions of the structural members that hold the joints. Thus, the ideal coordinate transformation that takes a point on the tool expressed in the spindle frame to a frame attached to the table is given by:

$$H = \Phi_B T_1 \Phi_x T_2 \Phi_z T_3 \Phi_y T_4 \Phi_w T_5. \tag{3.1}$$

To introduce rotational and translational errors into the shape and joint transformations, constant (not position dependent) components of errors are introduced into the shape transformations while the position dependent components are introduced into the joint transformations. There are three types of transformation for a machine tool.

3.1.1 SHAPE TRANSFORMATIONS (T_1~T_5)

If the small dimensional (translation) and deflection (angular) errors are introduced to T_i, an ideal shape transformation, the actual transformation, shown in Figure 3.2, becomes,

$$T_i' = \begin{bmatrix} 1 & -\alpha_i & \beta_i & x_i + \Delta x_i \\ \alpha_i & 1 & -\gamma_i & y_i + \Delta y_i \\ -\beta_i & \gamma_i & 1 & z_i + \Delta z_i \\ 0 & 0 & 0 & 1 \end{bmatrix}, i = 1{\sim}5, \tag{3.2}$$

where α_i, β_i and γ_i are small rotational errors about Z, Y and X directions, $[x_i \quad y_i \quad z_i]^T$ and $[\Delta x_i \quad \Delta y_i \quad \Delta z_i]^T$ are the constant shift vector and the small position errors vector in the workspace.

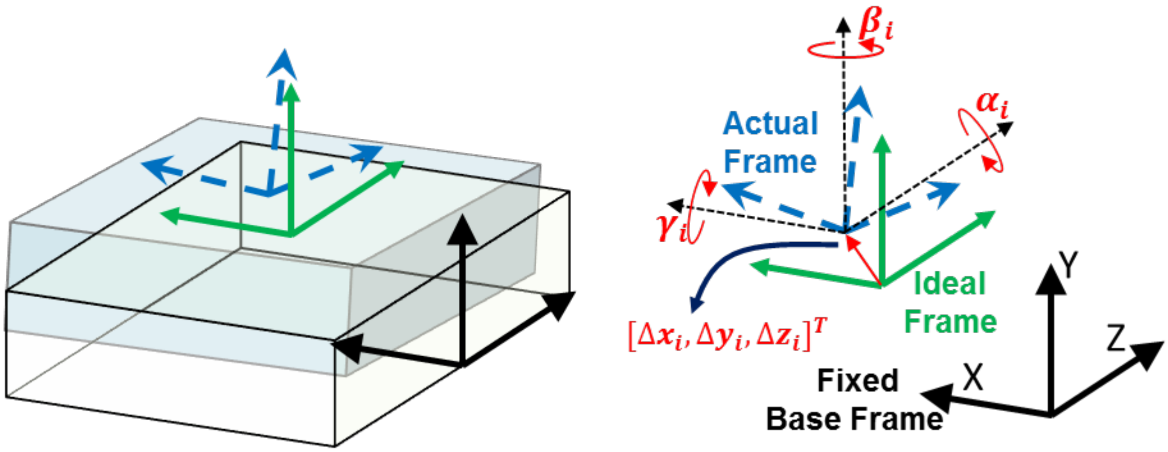

Figure 3.2: Ideal and actual shape transformations

3.1.2 PRISMATIC JOINT TRANSFORMATIONS (Φ_x, Φ_y, Φ_z and Φ_x)

An actual prismatic joint, in addition to producing the desired translation, will also produce error motions, including error in positioning along the joint, straightness errors and angular errors. As is evident for this transformation matrix, the error terms are functions of the joint displacement. A HTM of a prismatic joint, proposed by Ferreira and Liu [33] is given below:

$$\Phi_x = \begin{bmatrix} 1 & -x\dfrac{d\alpha}{dx} & x\dfrac{d\beta}{dx} & x(1+\delta x) \\ x\dfrac{d\alpha}{dx} & 1 & -x\dfrac{d\gamma}{dx} & \dfrac{x^2}{2}\dfrac{d\alpha}{dx} \\ -x\dfrac{d\beta}{dx} & x\dfrac{d\gamma}{dx} & 1 & -\dfrac{x^2}{2}\dfrac{d\beta}{dx} \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{3.3}$$

where x is the commanded joint position, δx is a rate of accumulation of positioning error and $\frac{d\alpha}{dx}$, $\frac{d\beta}{dx}$ and $\frac{d\gamma}{dx}$ are the rates of accumulation of angular errors (roll, pitch and yaw) as the joint moves along X-axis.

The linear variation of angular errors with displacement along the axis necessitates the addition of squared terms to the straightness error as suggested by Bryan [100]. One may add additional higher order terms to account for other effects. Figure 3.3 shows the relationships between the error terms and the fixed and moving coordinate frames for such a joint model. The actual transform introduces small angular motions and a positioning error to the ideal desired motion. Similarly, HTMs for inaccurate prismatic joints for Y, Z and W axes can also be derived.

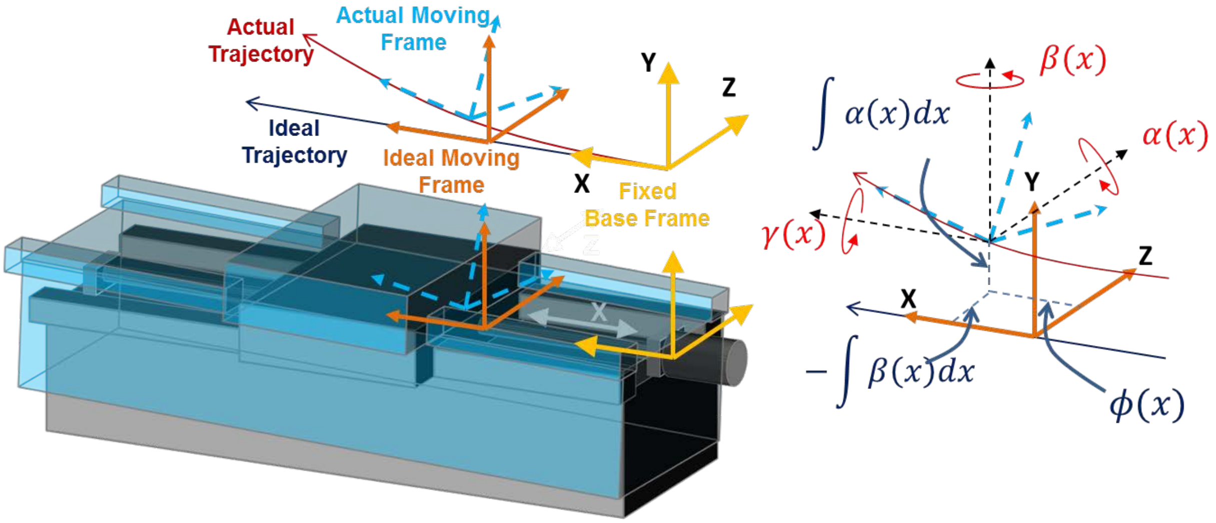

Figure 3.3: Ideal and actual prismatic joint transformations

3.1.3 ROTARY JOINT TRANSFORMATIONS (Φ_B)

The joint transformation of an ideal rotary joint (rotation about Y-axis) can be expressed as:

$$\Phi_B = \begin{bmatrix} \cos(B) & 0 & \sin(B) & 0 \\ 0 & 1 & 0 & 0 \\ -\sin(B) & 0 & \cos(B) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{3.4}$$

where B is command rotational displacement of the joint.

An actual joint introduces several error motions. First, when the joint is commanded to a position B, it may have positioning error, β. Further, the rotational errors may introduce tilts of α and γ, and the entire moving table may shift due to the accumulation of translation errors d_x, d_y and d_z. Thus, as depicted in Figure 3.4, the actual transform introduces small translational and rotational displacements to model the wandering and tilt of an actual joint. The joint transformation for an actual rotary joint (rotation about the y-axis), assuming small angles, the errors, is given by:

$$\Phi'_B = \begin{bmatrix} \cos(B) - \beta B'^{\,\sin(B)} & -\alpha B' & \sin(B) + \beta B' \cos(B) & B' d_x \\ \alpha B' & 1 & -\gamma B' & B' d_y \\ -\sin(B) - \beta B' \cos(B) & \gamma B' & \cos(B) - \beta B'^{\,\sin(B)} & B' d_z \\ 0 & 0 & 0 & 1 \end{bmatrix}, \qquad (3.5)$$

where B' is the error associated with the rotary command, which is modelled using Fourier sine series:

$$B' = \sum_{n=1}^{N} b_n \sin(\frac{nB}{2}). \qquad (3.6)$$

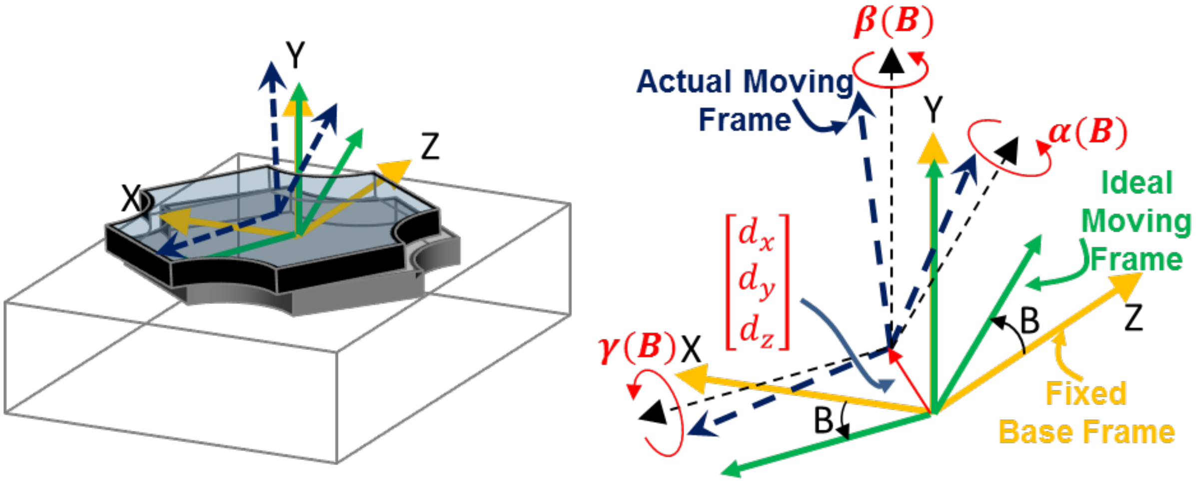

Figure 3.4: Ideal and actual rotary joint transformations

For a typical rotary joint as shown in Figure 3.5 and Figure 3.6, for purposes of consistency one would like to have the errors at $B = 0$ close to those at $B = 2\pi$ [32]. The error parameters are modelled by Fourier sine series instead of Taylor series to satisfy the consistency. For model efficiency and simplicity, only the first term in Equation (3.6) is used ($b_1 = 1$).

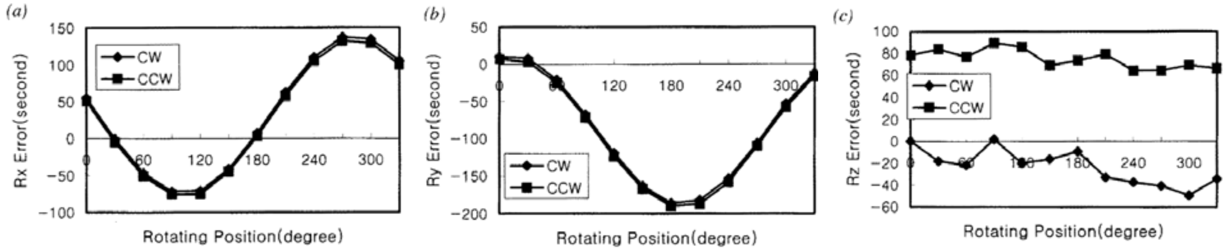

Figure 3.5: Errors of a rotary table: (a) about X-axis; (b) about Y-axis; (c) about Z-axis [32]

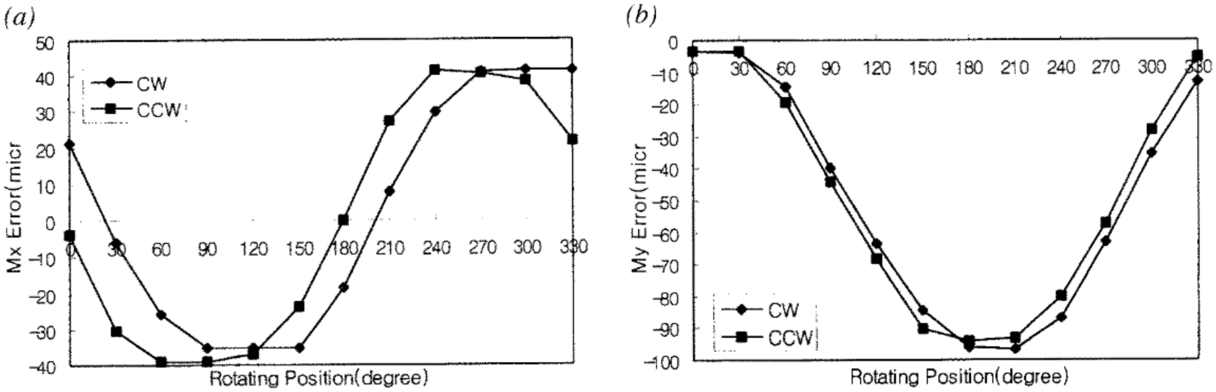

Figure 3.6: Errors of a rotary table: (a) x-directional translation error; (b) y-directional translation error [32]

3.1.4 LINEAR MODEL CONSTRUCTION

Now combining all the HTMs of inaccurate joints and structural members defined in Equation (3.2), (3.3) and (3.5), the actual coordinate transformation that takes a point on the tool expressed in the spindle frame to a frame attached,

$$R = \Phi'_B T_1{}' \Phi_x{}' T_2{}' \Phi_z{}' T_3{}' \Phi_y{}' T_4{}' \Phi_w{}' T_5{}'. \tag{3.7}$$

Eliminating second and higher-order terms of small errors, the first-order forward kinematic equation with errors for the machine can be written as:

$$R = H + \Delta H + O(2) \approx H + \Delta H, \tag{3.8}$$

where H is the ideal forward kinematics derived in Equation (3.1), determined by ideal machine joints and nominal dimensions of the structural members that hold them (i.e., ideal joint and shape transformations) are given, ΔH is the sum of ten first-order terms, and $O(2)$ represents all higher order terms.

Thus, the machine's volumetric error components can be defined as the difference between actual and ideal forward kinematics,

$$\begin{bmatrix} \vec{e} \\ 1 \end{bmatrix} = T_0(R - H) \begin{bmatrix} \vec{r}_t \\ 1 \end{bmatrix} \approx T_0 \Delta H \begin{bmatrix} \vec{r}_t \\ 1 \end{bmatrix}, \tag{3.9}$$

where $\vec{e}$ is the error vector $\vec{e} = [e_x \quad e_y \quad e_z]^T$, T_0 is one additional shape transformation added to obtain a convenient reference for measurement or programming, and $\vec{r}_t$ is the position of the target in the spindle frame as shown in Figure 3.1(b).

3.1.5 IDENTIFICATION OF ERROR MODEL PARAMETERS

The model, developed in Section 3.1.4, assembles all the error sources in the kinematic chain to obtain their influence on the volumetric error components of the machine. There are a total of 52 error sources or parameters (five shape transformations, each with six error parameters, four joint transformations for linear axes, each with four error parameters, and one for a rotary axis with six parameters) that are composed into an expression for the volumetric error components observed in the machine's workspace. To use this model for compensating the volumetric errors, it is necessary to obtain values for these parameters.

Estimation of the parameters in the error model is done by observing the volumetric errors of the machine at different points in its workspace with a laser tracker. The relationship between the measurement frame and the table frame (from the kinematic model of the machine) is captured by the homogeneous transformation matrix, T_0 as shown in Figure 3.7. However, to do so, the frame in which the laser tracker makes measurements, T_0 must be first estimated before the parameters of the error model can be obtained. This is done in the following two steps.

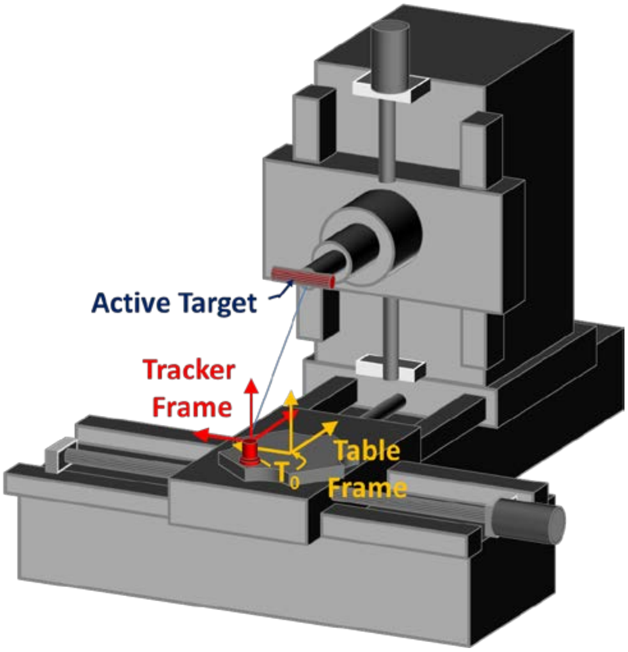

Figure 3.7: Depiction of measurement of volumetric error of the machine using a laser tracker

- Step 1: Find T_0, the best-fit measurement frame

Assume the machine to be ideal and identify the best values for T_0 to minimize the discrepancy between the laser tracker observations of position and the commanded position. Since T_0 is a rigid transformation, this step accounts any location and alignment errors between the machine and the laser tracker as shown in Figure 3.7. In addition, this step will also reduce the effects of any error sources that produce a rigid displacement of the entire machine's workspace. The residual errors that result from this process (of aligning measuring frame with the machines coordinates) are referred to as the nominal errors of the machine.

To identify T_0, assume ideal kinematics for the machine defined in Equation (3.1),

$$\vec{r}_{t0} = T_0 \Phi_B T_1 \Phi_x T_2 \Phi_z T_3 \Phi_y T_4 \Phi_w T_5 = T_0 H \vec{r}_t, \tag{3.10}$$

where T_0 has rigid body translations and small angle rotations as parameters to be identified,

$$T_0 = \begin{bmatrix} 1 & -\alpha_0 & \beta_0 & x_0 \\ \alpha_0 & 1 & -\gamma_0 & y_0 \\ -\beta_0 & \gamma_0 & 1 & z_0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{3.11}$$

Further, $\vec{r}_t$ is the position of the target in the spindle frame and $\vec{r}_{t0}$ is its image in the measurement frame. For different joint commands (or measurement points), the kinematic transmission of the machine, H will vary. For the i$^{\text{th}}$ measurement point, the error vector, $\vec{e}_i$ between the forward kinematic transmission and the measurement recorded by the laser tracker can be expressed as,

$$\vec{e}_i \cong T_0 H \vec{r}_t - \vec{q}_i, \tag{3.12}$$

where $\vec{q}_i$ is the measurement recorded by the tracker.

The best-fit homogeneous transformation, T_0 to the measurement frame can be obtained by minimizing the sum-of-squares of the discrepancy between the ideal machine's commanded positions and the measurements made by the tracker.

- Step 2: Identify the parameters of the error model from the nominal errors observed in the machine's workspace

The error sources in the kinematic chain of the machine cause the workspace of the machine to dilate/contract, shear and bend. These effects are encoded errors measured in the point-cloud of error measurements made by the laser tracker. In this step, least-square is used to identify the parameters. As mentioned earlier, there are 52 error sources/parameters in the error model. Further, for the derivation of the error model, these errors are assumed to be small. The model is linear in the set of parameters and can be expressed by an error parameter vector, $\vec{p}_{52}$ pre-multiplied by a coefficient matrix, M_{52}. Equation (3.12) is written as:

$$\vec{e}_i \cong M_{52,i} \vec{p}_{52}, \tag{3.13}$$

where $\vec{e}_i \in \mathbb{R}^3$ is the modeled error vector, $\vec{p}_{52}$ contains all 52 error parameters and $M_{52,i}$ is a matrix with three rows and 52 columns, where each term being a function of known machine constants and commanded positions. Each row then represents the coefficients of the linear combination that take the error sources to the X, Y and Z components of volumetric error at a point in the machine's workspace.

As would be expected, the influence of some parameters on the observed volumetric error components will be inseparable from each other by only change the commanded position of the tool. $\Delta x_1 \sim \Delta x_5$ are linear misalignments along X direction, which affect only the X component in the volumetric error. β_5 is the angular error about Y-direction, which only causes an Abbe error of $-r_t\beta_5$ along X-direction. $\Delta x_1 \sim \Delta x_5$ and β_5 contribute in exactly the same manner to volumetric error components at a point, irrespective of its location in the machine's workspace. They must therefore be identified as a group. Similarly, $\Delta z_1 \sim \Delta z_5$ are grouped, and $\Delta y_1 \sim \Delta y_5$ are grouped with $r_t\gamma_5$, Further, α_2 and α_3 are the Z-rotational errors of the two structural members that hold the Z-axis. They share the same leverage and cause identical effect of volumetric errors, and hence should be identified together. Similarly, β_3 and β_4 are grouped as well as γ_1 and γ_2. Also, other parameters such as α_4, α_5 and $\frac{d\alpha}{dw}$ must be removed because they have no influence on the volumetric error when the tool reference point lies along the axis of the spindle.

After all these redundant parameters are eliminated or grouped, the error model is derived,

$$\vec{e}_i \cong M_i\vec{p}, \tag{3.14}$$

where $M_i \in \mathbb{R}^{3\times 32}$ is a sub-matrix of $M_{52,i}$ and

$$\vec{p} =$$

$$[\alpha_1, \alpha_2 + \alpha_3, \beta_1, \beta_2, \beta_3 + \beta_4, \gamma_1 + \gamma_2, \gamma_3, \gamma_4, -\beta_5 r_t + \sum_{i=1}^{5} \Delta x_i, \gamma_5 r_t + \sum_{i=1}^{5} \Delta y_i, \sum_{i=1}^{5} \Delta z_i,$$

$\frac{d\alpha}{dx}, \frac{d\beta}{dx}, \frac{d\gamma}{dx}, \delta x, \frac{d\alpha}{dy}, \frac{d\beta}{dy}, \frac{d\gamma}{dy}, \delta y, \frac{d\alpha}{dz}, \frac{d\beta}{dz}, \frac{d\gamma}{dz}, \delta z, \frac{d\alpha}{dw}, \frac{d\beta}{dw}, \frac{d\gamma}{dw}, \delta w, \alpha, \beta, \gamma, d_x, d_y, d_z] \in \mathbb{R}^{32}$, where r_t is tool length.

The M_i matrix for the above set of parameters, constructed for a particular point in the machine's workspace, as previously mentioned, has elements made up of functions of the machine's constants and the axial commands that correspond to that point. Thus,

$$\begin{bmatrix} e_{x,i} \\ e_{y,i} \\ e_{z,i} \end{bmatrix} \cong \begin{bmatrix} M_{x,i} \\ M_{y,i} \\ M_{z,i} \end{bmatrix} \vec{p}, \tag{3.15}$$

where $M_{x,i}$, $M_{y,i}$ and $M_{z,i}$ correspond to the linear combinations of the error parameters that produce the i$^{\text{th}}$ measured set of error components $e_{x,i}$, $e_{y,i}$ and $e_{z,i}$.

Now consider an observation set consisting of errors observed at n points, under the assumption that the errors observed at the i$^{\text{th}}$ point is explained by the model:

$$\vec{e}_i = M_i \vec{p} + \vec{N}, \tag{3.16}$$

where $\vec{e}_i = [e_{x,i} \quad e_{y,i} \quad e_{z,i}]^T$contains the components of the errors observed at the i$^{\text{th}}$ point, M_i is the corresponding 3x32 relational matrix and $\vec{N} \in \mathbb{R}^3$ is the observation noise vector with elements drawn from the Gaussian distribution $N(0, \sigma)$, σ being the standard deviation of the observation noise, a system of 3n equations for estimating the parameters is built:

$$\vec{e} = M\vec{p}, \tag{3.17}$$

where $\vec{e} = [\vec{e}_1^T \ \dots \ \vec{e}_n^T]^T \in \mathbb{R}^{3n}$ vector containing the components of the measured error vectors in the point-set, $M = [M_1^T \ \dots \ M_n^T]^T \in \mathbb{R}^{3n \times 32}$ is the new coefficient matrix.

The least-squares estimate of $\vec{p}$ which minimizes the sum of squares of the discrepancy between the RHS and LHS of Equation (3.17) is given by:

$$\hat{p} = (M^T M)^{-1} M^T \vec{e}. \tag{3.18}$$

The estimate $\hat{p}$ minimizes the L_2-norm of the residuals, $\|\vec{e} - M\vec{p}\|_2 = (\vec{e} - M\vec{p})^T(\vec{e} - M\vec{p})$ and produces an unbiased estimate of $\vec{p}$ over the entire set of observations (and the workspace, if the point-set is a good representation of it).

3.2 EXPERIMENTAL VALIDATION

3.2.1 DATA COLLECTION

In order to identify the kinematic error model parameters, measurements of the machine tool are taken. These measurements are collected over the entire 3D space using a Laser Tracker and Active Target system (Figure 3.8) to ensure that all axis-dependent machine tool geometric errors are captured. The Laser Tracker used in this test is the API Radian which has a static measurement accuracy of +/- 10 μm or 5 ppm (2σ) according to the specifications provided by API. From this and the tracker's position on the machine, the largest measurement standard deviation (σ) value over the measured range was calculated to be 8.9 μm. In order to ensure that the Laser Tracker was thermally isolated from the machine tool, a plastic Isolation Block was placed between the Laser Tracker base and the machine tool.

Before measurements are taken, a measurement frame is identified. With the Laser Tracker attached to the machine tool bed and the Active Target attached to the machine tool spindle, as shown in Figure 3.8, the B-Axis is rotated with the other axes stationary in order to generate a circle of points. The normal vector of this circle is used as the vertical (Y-Axis) of the measurement frame. Next, the B-Axis is re-oriented to its 0° position and three points are measured as the machine moves along its X-Axis. The best fit line to these points is used as the X-Axis direction of the measurement frame. A right-handed frame is established from these two axes. This frame is then transformed into the negative Y-Axis direction by the Y-Axis encoder value of the machine tool to account for the Y position of the machine tool spindle during the measurement frame identification.

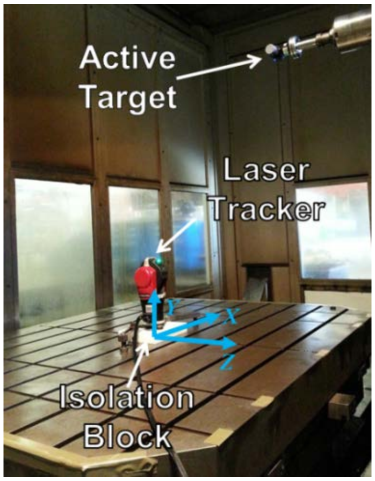

Figure 3.8: Machine tool work cell and table base frame

The machine tool repeatability, which establishes the maximum possible accuracy for a perfectly compensated machine tool, was calculated next. To determine the machine tool's repeatability, eight quasi-random points from the machine tool's working joint space were measured ten times each. Each cycle of the eight points was measured in a different randomized order to approximate arbitrary approach directions. The error of each measurement is given by,

$$e_{i,j} = \sqrt{\left(x_{i,j} - \bar{x}_i\right)^2 + \left(y_{i,j} - \bar{y}_i\right)^2 + \left(z_{i,j} - \bar{z}_i\right)^2}, \tag{3.19}$$

where $e_{i,j}$ is the error of the jth measurement of the ith point, $\left[x_{i,j}\ y_{i,j}\ z_{i,j}\right]$ is the jth measurement of the ith point, and $[\bar{x}_i\ \bar{y}_i\ \bar{z}_i]$ is the average measurement of the ith point.

From the measurements taken of the machine, the largest error was 0.0217 mm, which is used as the machine tool's repeatability. It should be noted that this repeatability value is only 2.4 times the measurement standard deviation meaning that a large portion of this value is may be due to the accuracy level of the laser tracker as opposed to the machine itself. Despite this fact, this repeatability still corresponds to the highest potential measured accuracy of the machine if it was perfectly compensated.

The measurement locations used for model identification and testing were selected next. For the identification set, 290 quasi-random points were selected throughout the machine tool's joint space, and an additional 50 quasi-random points were generated as a testing set. The number of points selected for identification and testing was selected through past experience with similar sized machine tools [15]. This number has the necessary richness to appropriately identify the geometric errors of the machine tool while minimizing the machine tool's down time. The joint ranges used to generate these points are shown below in Table 3.1, and the distributions of the points are shown below in Figure 3.9.

Table 3.1: Minimum and maximum commands used for modeling and testing

Axis	Minimum Command	Maximum Command
B	0°	360°
X	-1250 mm	1250 mm
Z	900 mm	2200 mm
Y	350 mm	2500 mm
W	-800 mm	-200 mm

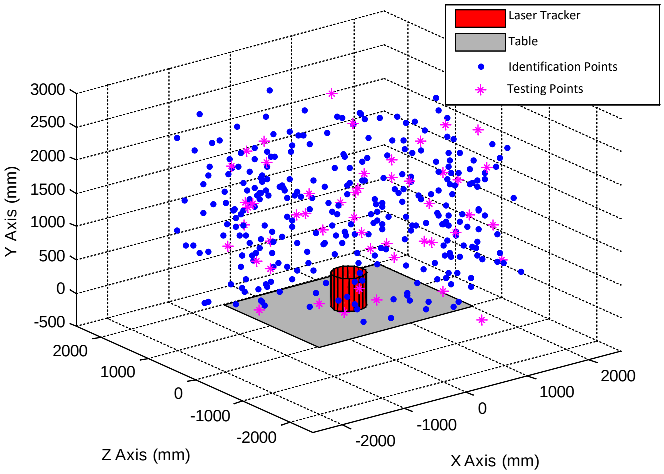

Figure 3.9: Positions of identification and testing points inside working envelop, given in machine tool base frame.

Using the Laser Tracker and Active Target system, the 290-point identification set was measured twice. In each measurement set a different length mount was used to attach the active target to the spindle as shown in Figure 3.8. Because the rotation of the spindle does not need to be modeled, these two mounts (Figure 3.10) allow for the spindle orientation to be determined for each point by finding the vector between the measurement sets.

Figure 3.10: Active Target machine tool spindle mounts

Because the same axis commands are used when taking both sets of identification points, it is possible to use the two sets of measurements to examine the potential existence of thermal drift in the measurement setup. For each point in the identification set, the distance between the two measurements of that point is ideally equal to the tool length difference of the two Active Target mounts (within machine tool repeatability). Therefore, if the distance between measurements is larger than the repeatability (0.0217 mm), then some shift must have occurred during the time that the system was measured. The distances between the measurements from each set (with the tool length offset removed) are shown in Figure 3.11. The distance between corresponding points ranges from -0.13 to 0.11 mm. Since this value is approximately six times the measured repeatability value, there is evidence that some sort of drift occurred during the measurement process. Furthermore, since the air temperature changed by 3.7°C during the measurement process, thermal effects is a likely source of some or all of this drift.

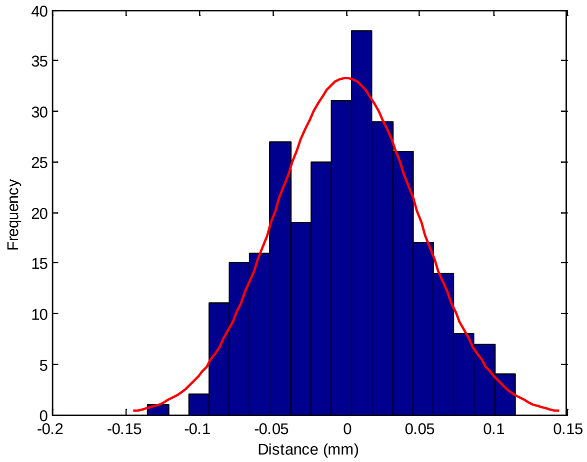

Figure 3.11: Distance between short tool and long tool measurements

3.2.2 BEST-FIT MEASUREMENT FRAME

The procedure described in Section 3.2.1 was used on the data collected in both the identification and testing data sets (described in Section 3.2.3). Table 3.2 shows the estimated errors between the nominal measurement frame and the machine's reference. Also shown in the table is the mean magnitude of the residual error vectors at the measurement points. For identification purpose, two sets of measurement were taken using different lengths of tool. After that, the identified parameters were used for modeling the testing sets.

Table 3.2: Best measuring frames of each measuring set

Set	Short Tool	Long Tool	Test 1	Test 2
x_0 (mm)	0.00845	0.0122	0.00715	0.0111
y_0 (mm)	0.351	0.304	0.293	0.298
z_0 (mm)	-0.0147	-0.0124	0.0095	0.0116
α_0 (rad)	-6.21E-06	3.68E-05	5.78E-05	6.68E-05
β_0 (rad)	-2.13E-05	-1.85E-05	-2.43E-05	-2.35E-05
γ_0 (rad)	1.05E-05	2.41E-05	3.52E-06	5.05E-06
Residual(mm)	0.4214	0.3175	0.2783	0.2745

3.2.3 ERROR PARAMETER IDENTIFICATION

The results of parameter identification are shown in Table 3.3. The data for the two different tools (short, 312.035mm, and long, 435.185mm) were analyzed separately to identify the error parameters of the two identification sets. From Table 3.3, the high correlation between the parameters identified in the two experiments is apparent. The deviations seen are due to the temperature changes between the two experiments and the uncertainty in the assembly of the target on the tool.

Table 3.3: Values identified for the parameters of the error model

Unit: mm			Unit: rad			Unit: rad/mm		
Parameter	Short	Long	Parameter	Short	Long	Parameter	Short	Long
$x_1+\ldots x_5-t_l\beta_5$	3.62E-02	5.92E-02	α_1	-2.36E-05	-1.96E-05	$d\alpha/dx$	1.55E-08	2.57E-08
$y_1+\ldots y_5+t_l\gamma_5$	-4.78E-02	-2.35E-02	$\alpha_2+\alpha_3$	-1.01E-05	-1.18E-05	$d\beta/dx$	1.13E-09	7.01E-09
$z_1+\ldots z_5$	-3.12E-01	-1.97E-01	β_1	-6.93E-05	-6.95E-05	$d\gamma/dx$	-1.04E-08	-1.16E-08
dx	1.13E-02	3.75E-03	β_2	4.16E-05	3.51E-05	$d\alpha/dy$	-1.26E-08	-1.80E-08
dy	2.62E-02	2.71E-02	$\beta_3+\beta_4$	5.83E-05	4.82E-05	$d\beta/dy$	2.57E-08	3.03E-08
dz	-1.28E-02	-7.25E-03	$\gamma_1+\gamma_2$	-1.85E-04	-1.77E-04	$d\gamma/dy$	-5.54E-09	-1.54E-09
Unit: dimensionless			γ_3	9.02E-05	8.19E-05	$d\alpha/dz$	-1.68E-08	-1.34E-08
Parameter	Short	Long	γ_4	-1.85E-04	-2.07E-04	$d\beta/dz$	-3.81E-08	-2.73E-08
δx	-1.23E-04	-1.12E-04	β	-9.57E-07	-9.72E-06	$d\gamma/dz$	2.97E-08	2.78E-08
δy	-1.10E-04	-1.09E-04	α	1.15E-06	6.78E-07	$d\beta/dw$	2.93E-08	1.80E-08
δz	-3.75E-05	-3.75E-05	γ	8.03E-06	4.84E-06	$d\gamma/dw$	-4.10E-07	-4.19E-07
δw	-1.06E-04	-1.06E-04						

Figure 3.12(a) shows the distributions of the residual errors. The statistical analysis of the results is shown in Table 3.4. Compared with the residual errors obtained from the frame alignment process, the error model reduces not only the mean but also the maximum (which characterizes the worst-case uncertainty of the machine/model) errors by 90% and 82% respectively.

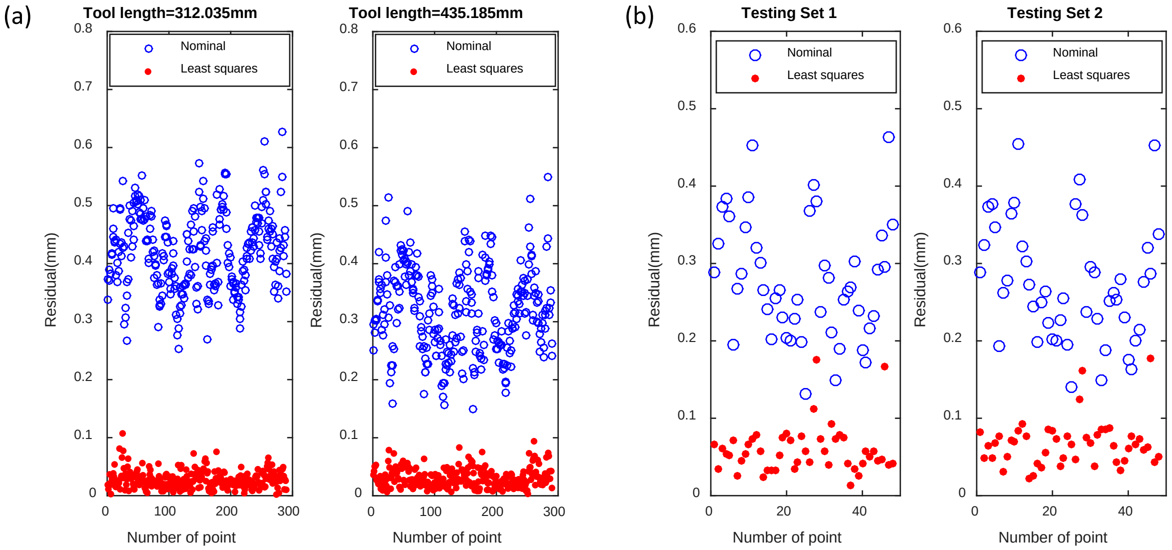

Figure 3.12: (a) The magnitudes of error residuals on two identification sets (290 points in each); (b) the magnitudes of error residuals on two testing sets (48 points in each).

Table 3.4: Model performance for two sets with two different tool lengths

Short tool	Mean Residual	% decrease	Max. Residual	% decrease
Nominal	0.4214 mm	N/A	0.6270 mm	N/A
Least squares	0.0277 mm	93.43%	0.1073 mm	82.88%
Long tool	Mean Residual	% decrease	Max. Residual	% decrease
Nominal	0.3175 mm	N/A	0.5492 mm	N/A
Least squares	0.0307 mm	90.34%	0.0941 mm	82.86%

3.2.4 ERROR MODEL TESTING

With the error model parameters obtained from the identification sets, the model's prediction capability are checked against two testing sets consisting of 48 previously-unseen data points, taken with the long tool. The results of this testing are shown in Table 3.5 and Figure 3.12(b). Compared with the nominal machine errors, the model can provide, approximately, a 75% reduction of average magnitude of errors vectors at the points in the data sets.

Table 3.5: Model performance for two testing data sets (Tool length=435.185mm)

Testing set 1	Mean Residual	% decrease	Max. Residual	% decrease
Nominal	0.2783 mm	N/A	0.4624 mm	N/A
Least squares	0.0590 mm	78.80%	0.1760 mm	61.94%
Testing set 2	Mean Residual	% decrease	Max. Residual	% decrease
Nominal	0.2745 mm	N/A	0.4546 mm	N/A
Least squares	0.0670 mm	75.59%	0.1767 mm	61.13%

3.3 SUMMARY

A kinematics model for a 5-axis machine tool with a redundant linear axis is developed in this chapter. This model introduced 52 parameters, linked to the error kinematics of the machine tool, which would need to be identified. Analysis of the model shows that only 32 of them have linearly independent effects on the volumetric errors in the workspace. A 2-step procedure for least-squares identification of the error model parameters from observations of the volumetric errors at points in the machine's workspace is also developed.

A laser tracker was used to make measurements at 290 randomly generated points in the machine's workspace. These measurements were repeated with tools of two different lengths characterizing the behavior of the machine with long and short tools. The error model parameters were estimated for these two different data sets. Despite some thermal drift on the machine between the experiments, the error model parameters estimated remained consistent in both magnitude and sign. Further, the model was able to reduce the errors at the observation points to about a third of their original values. The model was tested on two data sets of 48 observation points each. A similar model performance was observed. The proposed model has potential to be used for error prediction on commanded positions.

CHAPTER 4. ERROR OBSERVER DESIGN FOR MACHINE TOOL

In this chapter, the design and use of optimal error observer to track machine tool error is presented. The machine, modeling approach, and measurement techniques discussed in Chapter 3 are used to demonstrate the feasibility of using them to track machine's thermal error. It must be noted that the methodology of designing machine tool error observer is not limited to the error model developed in Chapter 3. Section 4.1 builds the mathematical model for optimal observer design of linear identification system. Section 4.2 describes the application of the optimal design theories in designing the thermal error observers for a 5-axis machine. Different design observer sets are proposed to identify the parameters in the volumetric error model. Section 4.3 describes the experimental setup to collect the measurement data, and Section 4.3.2~4.3.4 present results on the behavior of the model identified. Section 4.4 outlines the conclusions, drawn from this work.

4.1 OPTIMAL OBSERVER DESIGN FOR LINEAR SYSTEM

4.1.1 INTRODUCTION

As reviewed in Section 2.3, a linear identification problem with n design points is given by:

$$\vec{e} \cong M(\vec{J}_1, \ldots \vec{J}_n)\vec{p}, \tag{4.1}$$

where $\vec{e} \in \mathbb{R}^n$ represents a vector of n observable values that is related to $\vec{p} \in \mathbb{R}^k$, a set of k unknown parameters is the vector consisting of all undetermined parameters, $p_1, \ldots p_k$ (whose values are to be estimated) by the design matrix, $M(\vec{J}_1, \ldots \vec{J}_n) \in \mathbb{R}^{n\times k}$, whose row vectors are functions of $\vec{J}_1, \ldots \vec{J}_n$, sets of variables that can be independently controlled.

The best fit estimator of $\vec{p}$, $\hat{p}$ is given by least-squares fitting,

$$\hat{p} = (M^T M)^{-1} M^T \vec{e}. \tag{4.2}$$

The D-optimal design maximizes information by minimizing the volume of the confidence volumes or the uncertainty region around the estimator. The D-optimality is given by,

$$\min_{\vec{J}_1 \cdots \vec{J}_n} |(M^T M)^{-1}| = \min_{\vec{J}_1 \cdots \vec{J}_n} \prod_{i=1}^{k} \frac{1}{\lambda_i}, \tag{4.3}$$

where $\vec{J}_1 \ldots \vec{J}_n$ are n sets of controllable variables (in our case, the commanded axial positions) that control each row in the design matrix M, and λ_i is the ith eigenvalue of $M^T M$.

A-optimal design minimizes the average variance of the estimations on the regression coefficients, and its objective is given by:

$$\min_{\vec{J}_1 \cdots \vec{J}_n} tr((M^T M)^{-1}) = \min_{\vec{J}_1 \cdots \vec{J}_n} \sum_{i=1}^{k} \frac{1}{\lambda_i}, \tag{4.4}$$

where $\vec{J}_1 \ldots \vec{J}_n$ are n sets of controllable variables (in our case, the commanded axial positions) that control each row in the design matrix M, $tr((M^T M)^{-1})$ is the trace of $(M^T M)^{-1}$ and λ_i is the ith eigenvalue of $M^T M$.

The K-optimality criterion minimizes the sensitivity of estimator to observational error by minimizing the condition number of the design matrix,

$$\min_{\vec{J}_1 \cdots \vec{J}_n} \kappa(M) = \min_{\vec{J}_1 \cdots \vec{J}_n} \frac{\sigma_{max}}{\sigma_{min}} = \min_{\vec{J}_1 \cdots \vec{J}_n} \frac{\lambda_{max}}{\lambda_{min}}, \tag{4.5}$$

where σ_{max} and σ_{min} are the largest and smallest singular values of M, λ_{max} and λ_{min} are the largest and smallest eigenvalues of the information matrix, $M^T M$.

D, A and K-optimality criteria are all related to the eigenvalues of the information matrix [58], [59]. All three types of design problems deal with the maximization of information, quantified by surrogate functions of these eigenvalues. In Section 4.3, K-optimal design is selected to reject the measurement noise. However, the optimal design theories produce the best locations for observations to identify model parameters under the assumption that the form or degree (if it is a polynomial) of the underlying model of the linear system is known. In many situations, the functions used for machine tool error models are simplifications (typically with polynomials of axial displacements). Further, to keep the number of parameters manageable, they are assumed to be low-order polynomials. In such cases, there is always a possibility that neglected higher-order terms may be significant. Any observer design process must take steps to alleviate the deleterious effects of model inadequacy.

4.1.2 EXAMPLE PROBLEM

For example, if one tries to fit a straight-line model to a parabolic function, $y = x^2$ over the domain [0,1] with four observations. As can be seen in Figure 4.1, the modeling residuals of any line $y = p_1 x + p_2$ are not normally distributed but dependent on x because the linear model is inadequate. The identification system of p_1 and p_2 is given by:

$$\begin{bmatrix} y_1 \\ \vdots \\ y_4 \end{bmatrix} = \begin{bmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_4 & 1 \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} = M\vec{p}, \tag{4.6}$$

where $x_1, \dots x_4$ are the positions of the observations, $y_1, \dots y_4$ are the corresponding observations and M is the design matrix.

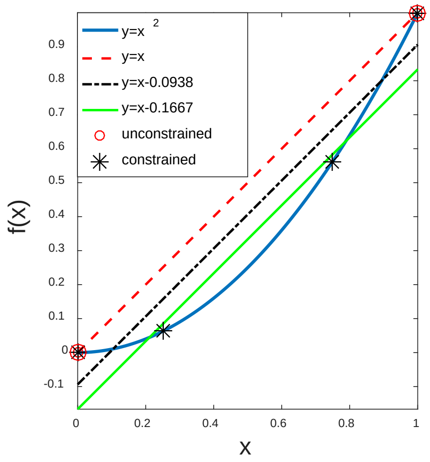

Figure 4.1: Quadratic function fitted by linear functions

The optimal design problem is given by,

$$\min_{\vec{J}_1 \cdots \vec{J}_n} f(M) \ni \vec{J}_i \in \Gamma(i = 1, \dots n), \tag{4.7}$$

where $f(M)$ could be D, A or K-optimization objectives defined in Equations (4.3), (4.4) and (4.5) and Γ is the design space.

In this example, $\Gamma = [0,1]$, $n = 4$ and $\vec{J}_i = x_i$ for $i = 1, \dots 4$. A, D and K-optimal designs all suggest that the best four observations for Equation (4.7) are x=0,0,1,1, and the line fitted by these observations is $y = x$. As shown in Table 4.1, the corresponding objectives, $tr((M^T M)^{-1})$, $|(M^T M)^{-1}|$, and $\kappa(M)$ of these four observations are minimized to be 1.5, 0.25 and 2.618. It's been observed in Figure 4.1 that the straight line defined by the end points only has good model performance at two ends. In fact, the best linear fitting of over that minimizes the sum of squared error is the green line in Figure 4.1. The observers produced by the optimal designs localize the observations at the boundaries of the design space, which causes the poor overall fitting performance.

Table 4.1: Optimal observers designed by A, D, K-optimal designs

Case	x_1^*	x_2^*	x_3^*	x_4^*	$tr((M^T M)^{-1})$	$\|(M^T M)^{-1}\|$	$\kappa(M)$
Unconstrained	0	0	1	1	1.5	0.25	2.618
Constrained	0	0.25	0.75	1	2.25	0.4	3.25

To avoid localized observation points, one can introduce constraints to the optimization procedure to distribute observations over the domain or design space. For example, the distribution can be one observation between 0 and 0.25, two between 0.25 and 0.75 and the last one between 0.75 and 1. A generalized constrained optimization problem is given by,

$$\min_{\vec{J}_1 \cdots \vec{J}_n} f(M) \ni \vec{J}_i \in \Gamma_i(i = 1, \dots n), \tag{4.8}$$

where $f(M)$ is the objective function to be minimized and Γ_i is the i^{th} constraint for the i^{th} set of variables.

In this case, $n = 4$, $\Gamma_1 = [0,0.25]$, $\Gamma_2 = \Gamma_3 = [0.25,0.75]$ and $\Gamma_4 = [0.75,1]$, and $\vec{j}_i = x_i$ for $i = 1, \dots 4$. As can be seen in Table 4.1, the solution to Equation (4.8) using A, D and K-optimality gives four different observers, x=0, 0.25, 0.75 and 1. The corresponding objectives, $tr((M^T M)^{-1})$, $|(M^T M)^{-1}|$, and $\kappa(M)$of these four observations are 2.25, 0.4 and 3.25, which are all larger than their unconstrained counterparts, 1.5, 0.25 and 2.618. Figure 4.1 shows the fitting result of the constrained optimization using the dashed black line. The line fitted by the observers of the constrained optimization is $y = x - 0.0938$, which is much closer to the best fitting line, $y = x - 0.1667$.

Thus, the judicious introduction of constraints to obtain distribution of the points balances the need to maximize the amount of information in the observer design with the need to guard against inadequacy of the proposed model. In this example, one might realize that the minimum number of observation points required for estimating the model parameters is two. By introducing redundancy (two additional observations) and constraining the locations of these extra points, one can provide the optimization procedure the flexibility to maximize the information content while, at the same time, ensure that all regions of the domain of the fit are represented. This strategy will be used in the next Section for the design of error observers for machine tools.

4.2 OBSERVER DESIGN FOR THE ERRORS OF A 5-AXIS MACHINE

The concepts discussed in Section 4.1 are tested on the machine and error model, which is built in the Section 3.1. The schematic of the 5-axis machine used in this study and its kinematic equivalent are shown in Figure 3.1. The machine has four prismatic joints, X, Y, Z and W axis with travels of 4m, 2.5m, 2.2m and 800mm respectively and a rotary joint, B axis that allows the table to rotate about the Y direction by 360 degrees. The Z and W axes are redundant axes.

To track the evolution of thermal errors, based on consultation by the users, it was decided that we design the observer so that the time for making measurements was limited to 25 minutes. Based on empirical experience, it is therefore decided to limit the number of measurement points for the observer to 80. The linear error model in Equation (3.14), as previously explained can be rewritten in the form of design matrix times the error parameter vector:

$$\vec{e}_i \cong M_i(\vec{J}_i)\vec{p}, \tag{4.9}$$

where $\vec{e}_i \in \mathbb{R}^3$ is the error observed at the ith measurement, $M_i \in \mathbb{R}^{3\times32}$ has elements that are functions of the ith commanded X, Y, Z, W and B axes positions, denoted by $\vec{J}_i$ and $\vec{p} \in \mathbb{R}^{32}$ is the error parameter vector.

With 80 observations, a system of 240 equations can be produced, given by:

$$\vec{e} \cong M(\vec{J}_1 \dots \vec{J}_{80})\vec{p} \tag{4.10}$$

where $M \in \mathbb{R}^{240\times32}$ is the design matrix controlled by 80 design points $\vec{J}_1 \dots \vec{J}_{80}$, $\vec{e} \in \mathbb{R}^{240}$ contains all components of the measured error vectors in the observer point-set.

The optimization problem that seeks to maximize the amount of information carried by 80 design points suggests the best set of axes command. Each design point is controlled independently by the commands of X, Y, Z, W and B axis. Therefore, the problem has 80x5 degrees of freedom subjected to the size command space defined by the limitation on each axis.

The A, D and K-optimal observers can be produced by solving the constrained optimization problem defined in Equations (4.7) and (4.8). This was encoded in a MATLAB program, using the generalized constrained optimization function FMINCON takes as input the definition of M in terms of axial positions of the machine and the constraints of the workspace. This function finds a local minimum, hence it was called several time with different randomly generated, feasible initial solutions. The objective functions converged to the similar values for all the cases. The positions of the measurements points for the three criteria in the work envelop of the machine are shown in Figure 4.2(a), (b) and (c). The values for the objective functions for three different criteria of optimal designs are listed in Table 4.2.

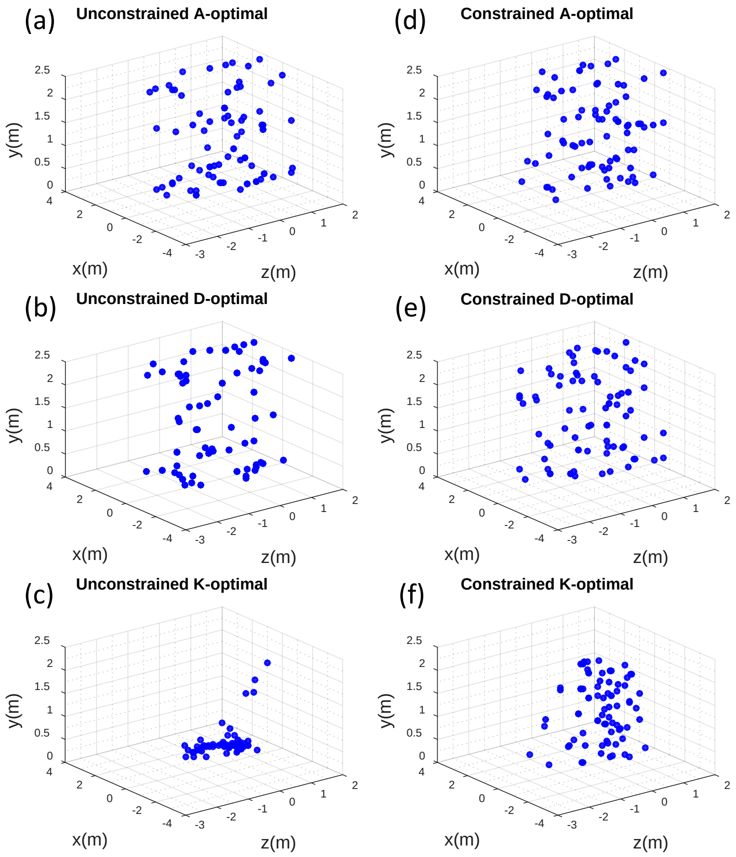

Figure 4.2: Designed observer sets given in machine tool base frame

Table 4.2: The unconstrained designs

Type	Points	$tr((M^TM)^{-1})$	$\|(M^TM)^{-1}\|$	$\kappa(M)$
A	80	99.7	2.27×10^{-39}	253.7
D	80	125.0	6.86×10^{-42}	337.5
K	80	183.2	5.04×10^{-15}	122.0
Random	290	80.9	5.95×10^{-46}	437.8

The locations of the measurement points for the A- and D-optimal observers are similar, and mostly located near the boundary of the workspace. This is because the objective of A and D-optimal designs maximizes the distance between observations in the domain to increase their influence on the estimates of the unknown parameters. However, the observations for the K-optimal observer are highly localized and located primarily near the bottom of the workspace. In the process of minimizing the influence of errors in observations, it also minimizes the influence of the observations. As mentioned in Section 4.1.2, the unconstrained locations for the optimal design of observers are expected to produce such localization. To obtain a more uniform location, the workspace of the machines is sliced into 4 zones along the y axis. Each slice is further decomposed into a central block and an annular space (having the same volume). Thus the workspace is broken up into 8 equal volumes as shown in Figure 4.3. Constraints formed by 8 volumes are then introduced into the optimization program to ensure that each of these 8 volumes contains 10 measurement points of the new “constrained” A, D and K-optimal observers. The results of the introduction of these distribution constraints is given in Figure 4.2(d), (e) and (f), and the new values of objective functions for the different criteria and designs are tabulated in Table 4.3.

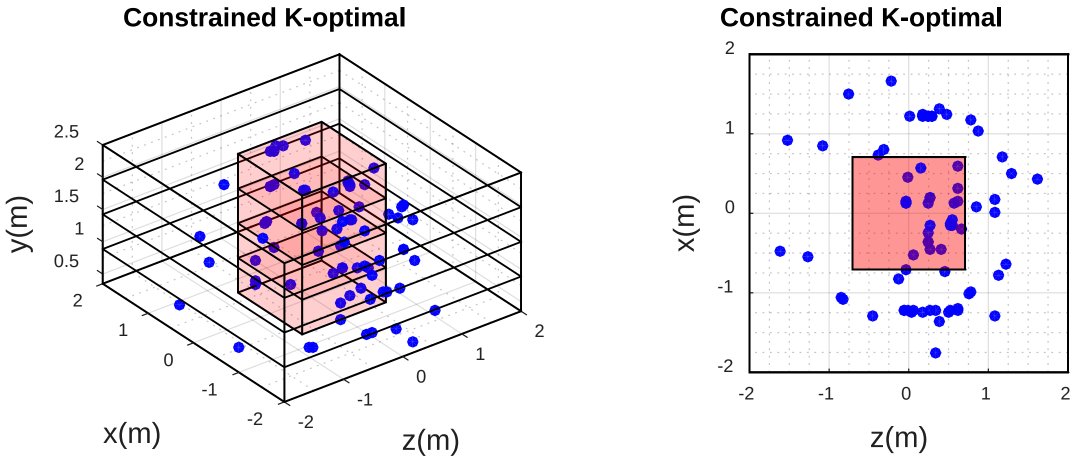

Figure 4.3: 8 constrain volumes and the constrained K-optimal observer set

Table 4.3: The constrained designs

Type	Points	$tr((M^T M)^{-1})$	$\lvert (M^T M)^{-1} \rvert$	$\kappa(M)$
A	80	100.5	1.33×10^{-38}	273.1
D	80	131.8	8.90×10^{-40}	330.1
K	80	108.3	3.74×10^{-29}	207.3
Random	290	80.9	5.95×10^{-46}	437.8

It can be seen in Figure 4.2 that the constraints successfully spread the measurement points over the whole workspace, but the price paid for introducing these constraints is also apparent in the value of the objective functions shown in Table 4.2 and Table 4.3.

4.3 EXPERIMENTAL VALIDATION

An experiment was designed to test the aforementioned 80-point, constrained K-optimal observer on the machine described in Section 4.2. A similar experiment with the unconstrained K-optimal observer was also performed. In this experiment, the objectives were to:

(1) Check how error models using the parameter estimates it produces compare with those using parameters estimated from the more traditional, measurement-intensive quasi-random point-sets.

(2) Determine its ability to track changes in these parameters as the thermal state of the machine changes.

(3) Assess improvements, if any, in the observer's performance brought about by the introduction of constraints to distribute the measurements in the workspace.

4.3.1 EXPERIMENT SETUP

Similar experiment setup was used as shown in Figure 3.8. An API Radian laser tracker with active target system was used to collect the data over the entire 3D space. The 80 points in the K-optimal observer were analyzed for reachability of the laser tracker. It was found that 4 points were not reachable. The value of objective function for the observer with the remaining 76 points increased from 207.3 to 222.0, which was not changed significantly. The machine was programmed to carry the active target of the tracker and dwell for a few seconds at these 76 measurement points. The tracker and machine were synchronized so that the tracker recorded the position of the target after the machine had settled at a measurement point. This measurement cycle was repeated at intervals of one hour. Six such measurements cycles were performed. The first one at the start of the experiment can identify the errors of the machine's initial state. In the four intervals between the first five measurement cycles, the spindle of the machine and the axes of the machine were exercised at roughly half their maximum speeds to heat up the machine. The machine could cool in the interval between the 5th and 6th measurement cycles. Figure 4.4 shows the schedule of the 8-hour experiment.

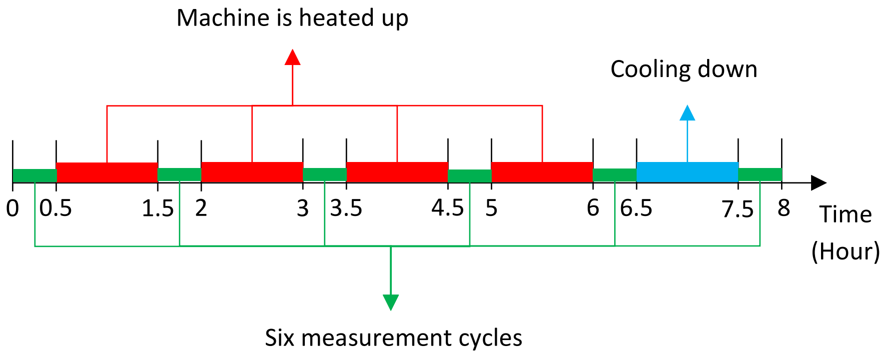

Figure 4.4: Measurement, heating and cooling cycles

Prior to the start of the experiment, the laser tracker is mounted at the center of the table on top of a thermal isolation block. A set of measurements are taken and processed to align the measurement frame to the machine's coordinate system as Section 3.1.5 and 3.2.2 show. Further, measurements are made to assess the repeatability of measurements of the laser tracker on the machine. This was found to be around 20 microns as discussed previously in Section 3.2.1.

Additionally, the machine was instrumented with 16 wireless thermal sensors to record temperatures at different positions of the machine structure. Four packet radio transceivers (shown in Figure 4.5), Adafruit Feather M0 RFM96 LoRa Radio (433 MHz) with 13 temperature sensors, TMP 36 were placed over the course of the experiment as shown in Figure 4.6, and the data acquisition system transmit temperature data to a computer-based server, which monitored and recorded the temperature readings in real time. Each of the 4 linear axes was instrumented with 3 temperature sensors (one on the drive and the others distributed around the length of the axes). One of the sensors on the Y-axis was used to monitor the spindle housing temperature. Temperatures were recorded at 1-minute intervals during the experiment.

Figure 4.5: Wireless transceiver unit

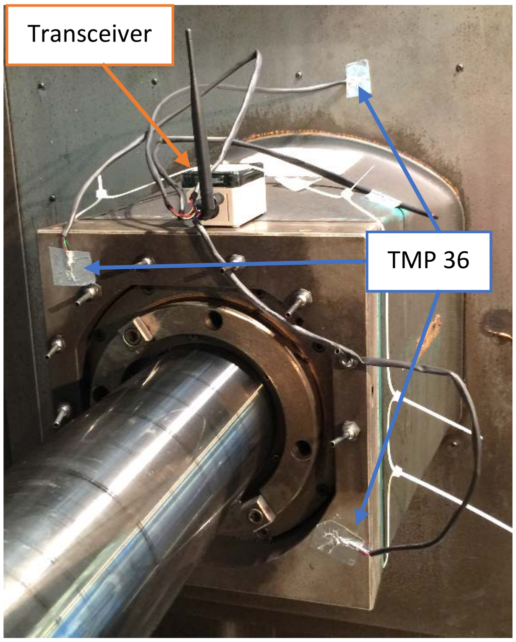

Figure 4.6: Temperature sensors setup on W-axis

The experiment was commenced in the morning and concluded late afternoon. The data recorded in each measurement cycle at the measurement points was fed into a MATLAB program and used to identify the parameters of the error model as demonstrated in Section 3.2.3 using least-squares fitting technique.

A similar experiment was conducted with the unconstrained K-optimal observer. In this case, the experiment was conducted without running the spindle between measurement cycles (the reason was to reduce the uncertainty in repeated mounting and dismounting the active target).

The following are some key points in the processing of the data obtained in each measurement cycle. For the first (initial) measurement cycle, misalignment between the measurement and movement frame, $T_{0,1}$ and error parameters, $\vec{p}_1$ are identified separately. In all subsequent cycles, the workspace drift is picked up by the constant terms of the error model. Thus, in the first measurement cycle, two minimization problems are solved as elaborated in Section 3.1.5:

1. Identify misalignment between the measurement and movement frame, $T_{0,1}$ in the first (initial) cycle,

$$\min_{T_{0,1}} \sum_{i=1}^{76} \left\| T_{0,1} H_i \vec{r}_t - \vec{q}_{i,1} \right\|^2, \tag{4.11}$$

where $T_{0,1} H_i \vec{r}_t$ is the ideal position of the i$^{\text{th}}$ measurement point predicted by the ideal forward kinematics and $\vec{q}_{i,1}$ is the actual position measured by the laser tracker at the i$^{\text{th}}$ measurement point.

2. Estimate the 32 error parameters, $\vec{p}_1$:

$$\min_{\vec{p}_1} \sum_{i=1}^{76} \left\| M_i \vec{p}_1 - (T_{0,1} H_i \vec{r}_t - \vec{q}_{i,1}) \right\|^2, \tag{4.12}$$

where $M_i \vec{p}_1$ is the modelled error, $T_{0,1} H_i \vec{r}_t - \vec{q}_{i,1}$ is the observed error.

For all subsequent (the 2nd to 6th) cycles, the misalignment between the measurement and movement frame $T_{0,j}$ is not updated. $T_{0,1}$ is used as the starting reference for the thermal drift of the machine and to give the growth in errors due to thermal effects. Therefore, the jth error parameter group denoted by $\vec{p}_j$ is identified using a single-step identification ($j = 2{\sim}6$),

$$\min_{\vec{p}_j} \sum_{i=1}^{76} \left\| M_i \vec{p}_j - (T_{0,1} H_i \vec{r}_t - \vec{q}_{i,j}) \right\|^2, j = 2{\sim}6. \tag{4.13}$$

The modelling residual at the ith observation of the jth cycle can be computed by,

$$RES_{i,j} = \left\| M_i \vec{p}_j - (T_{0,1} H_i \vec{r}_t - \vec{q}_{i,j}) \right\|_2, i = 1{\sim}76, j = 1{\sim}6. \tag{4.14}$$

4.3.2 MODEL PERFORMANCE

The statistics of the behaviour of the models identified for six different thermal states are shown in Table 4.4. The observed mean and max are the statistics of the errors observed after the tracker placement error are removed from the tracker readings. The mean and maximum residual are the statistics of the difference between the observed errors and those predicted by the identified models. One could expect these error statistics if the identified model was used for compensation. Figure 4.7 shows the average temperatures recorded by four wireless transmission systems on four linear axes.

Table 4.4: Model performances on different states (constrained K-optimal)

Machine State	Observed (μm)		Residual (μm)		% Decrease	
	Mean	Max	Mean	Max	Mean	Max
Initial	119.1	265.9	26.3	98.2	77.92%	63.07%
Heating	134.4	279.6	31.5	97.2	76.56%	65.24%
Heating	154.0	310.2	28.8	94.5	81.30%	69.54%
Heating	169.6	311.6	33.9	113.4	80.01%	63.61%
Heating	181.9	331.6	30.9	111.6	83.01%	66.34%
Cooling	157.3	309.4	29.8	101.7	81.06%	67.13%

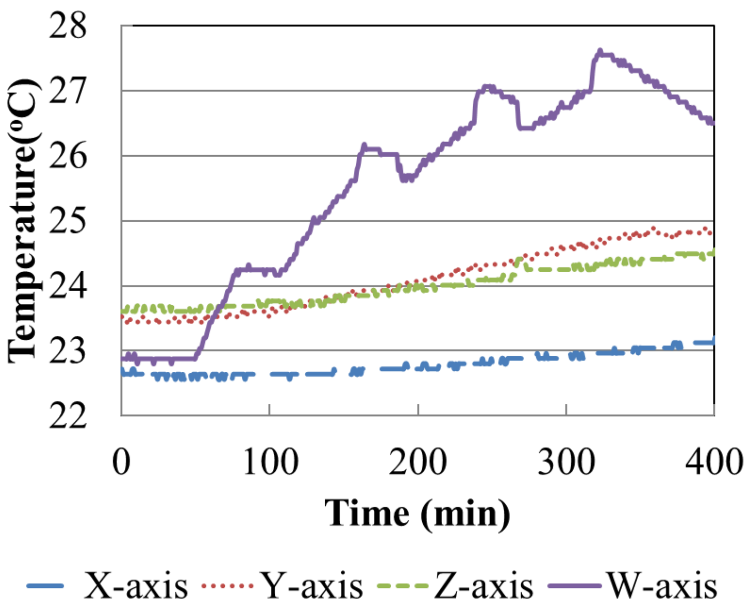

Figure 4.7: Temperature variations over the course of the experiment

The models identified by making measurements at points prescribed by the constrained K-optimal observer over 6 measurement cycles suggest highly repeatable performance for each thermal state. During the first data collecting cycle (initial state), the model provides 77% and 63% reductions in the mean and maximum magnitude of error, respectively. These percentages increased as the average magnitude of the machine's errors increased because the average magnitude of the residuals remained a relatively narrow (8 micron) band.

The observer performed well in terms of explaining the error. The repeatability of the positioning was around 20 microns (machine and laser tracker combined). Further, an additional uncertainty of about 5 microns was introduced because the active target had to be removed and remounted between measurement cycles (because the heating cycles required the spindle to be run). The average magnitude of the residuals at the observer measurement points over six experiments was close to 30 microns, which suggests that the identified models were capturing most of the systematic errors of the machine and adjusting the parameters appropriately to adjust to thermal changes of the machine.

These results are comparable to those reported in Section 3.2.3 on the same machine, using the same kinematic model but, instead using a quasi-randomly generated set of 290 observation points. In the aforementioned work, the average magnitude of the residuals was 27.7 microns as shown in Table 3.1. Thus, with the machine's thermal condition varying, additional uncertainty of removing and replacing the laser target in the spindle and a measurement set reduced by more than a third, the constrained K-optimal observer produced comparable performance. The feasibility of using a smaller and more strategically-chosen point-set to perform on-line thermal error tracking is thus demonstrated. The measurement cycle time for measurements for this reduced set of points is only 24 minutes. This suggests that, with a quick data collection strategy and a robust error model parameter estimation procedure, one might be able to track and compensate the thermal errors as they evolve by executing a process intermittent gaging and error updating strategy.

4.3.3 THERMAL ERROR TRACKING AND ANALYSIS

One can compare the performance of tracking approach to that of a static calibration approach that does not attempt to track and compensate thermal errors. In such a situation the machine is calibrated once (typically, a quarter or a month or, optimistically, at the beginning of a shift) and the results are used, without regards to the thermal state of the machine, for compensation of its errors during operation. Simulating an optimistic situation, where the machine is calibrated at the beginning of the shift and the results of the calibration are used through the entire shift, measurement frame discrepancy, $T_{0,1}$ and error parameters, $\vec{p}_1$ in the first cycle (initial state) are computed and used to calibrate the rest of five data sets. Thus, the residual at the i^{th} observation of the j^{th} cycle in this case is given by:

$$RES_{i,j} = \left\| M_i\vec{p}_1 - \left(T_{0,1}H_i\vec{r}_t - \vec{q}_{i,j}\right)\right\|_2, (j = 1, \dots 6). \tag{4.15}$$

The statistics of the residuals produced by this approach (or the performance of a static compensation approach) are given in Table 4.5. The average model residual grew from 26.3 to 155.1 microns in the five-hour heating up process and reduced to 120.7 after a one-hour cooling period. Over a 400-minute period of operation, the compensations estimated in the cold state of the machine, though producing some improvements in the error, are seen to become increasing ineffective. After 5 hours of heating, the compensations only produce a 15% reduction in error.

Table 4.5: Thermal drifts without updating the error parameters

Machine State	Observed (µm)		Residual (µm)		%Decrease	
	Mean	Max	Mean	Max	Mean	Max
Initial	119.1	265.9	26.3	98.2	77.92%	63.07%
Heating	134.4	279.6	77.1	122.6	42.63%	56.15%
Heating	154.0	310.2	115.4	180.0	25.06%	41.97%
Heating	169.6	311.6	138.6	208.8	18.28%	32.99%
Heating	181.9	331.6	155.1	235.0	14.73%	29.13%
Cooling	157.3	309.4	120.7	208.6	23.27%	32.58%

In many situations, instead of opting for a static calibration or attempting to update the entire parameter vector (to compensate for workspace drift and distortion), one may opt to probe a few points, estimate the drift and program in a shift of the programming origin based on these measurements. This situation is simulated by doing a full calibration in the first cycle, then use fixed error parameters, $\vec{p}_1$ but update workspace drift specified by $T_{0,j}$, which now has only three translational degrees-of-freedom. For the remaining five measurement cycles, the workspace drift, $T_{0,j}$ $\mathrm{T}_{0,\mathrm{j}}$is identified by:

$$\min_{T_{0,j}} \sum_{i=1}^{76} \left\| M_i\vec{p}_1 - (T_{0,j}H_i\vec{r}_t - \vec{q}_{i,j})\right\|^2, j = 2\sim 6, \tag{4.16}$$

where $T_{0,j}$ has only three translational variables,

$$T_{0,j} = \begin{bmatrix} 1 & 0 & 0 & x_{0,j} \\ 0 & 1 & 0 & y_{0,j} \\ 0 & 0 & 1 & z_{0,j} \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{4.17}$$

The same error parameters, $\vec{p}_1$ is used to predict the errors at the same observer locations. The prediction errors of the ith observation of the jth cycle is given by,

$$RES_{i,j} = \left\| M_i \vec{p}_1 - \left(T_{0,j} H_i \vec{r}_t - \vec{q}_{i,j} \right) \right\|_2, j = 2 \sim 6. \tag{4.18}$$

As shown in Table 4.6, updating $T_{0,j}$ is effective in controlling the inaccuracies caused by the thermal effect. The worst mean model residual occurs at the end of the heating period and is measured to be 98.1 microns, which is lower than that produced by using static calibrations. However, the model performance is seen to degrade severely when compared to the full identification of the model parameters. One can see that that using a static compensation and tracking only the drift of the workspace explain only 46.07% of the observed thermal error. These comparisons between full periodic parameter identification, partial (drift only) identification, and no identification not only illustrate the scale of the relative influence of thermal errors (workspace drift and distortion), but also demonstrates the need and importance of periodic updates to calibrations.

Table 4.6: Thermal drifts (only compensate the shift of the measurement frame)

Machine State	Observed (μm)		Residual (μm)		%Decrease	
	Mean	Max	Mean	Max	Mean	Max
Initial	119.1	265.9	26.3	98.2	77.92%	63.07%
Heating	134.4	279.6	68.7	184.4	48.88%	34.05%
Heating	154.0	310.2	73.3	203.0	52.40%	34.56%
Heating	169.6	311.6	88.6	243.1	47.76%	21.98%
Heating	181.9	331.6	98.1	255.5	46.07%	22.95%
Cooling	157.3	309.4	84.9	195.6	46.03%	36.78%

4.3.4 AXIAL BEHVIORS AGAINST THERMAL VARIATIONS

The thermal effects are observed to cause the average error of the machine to grow from 119.1 in the cold state to around 181.9 microns after 4 heating cycles. With compensation, the error model parameters (see Section 3.1 for more details) estimated by the single step identification process error, described in the Section 4.3.2, would hold the average error of the machine to around 30 microns. The error parameters were identified by the least-squares fitting. The variations of parameters related to X, Y, Z, and W axis at different machine's thermal states are shown in Figure 4.8.

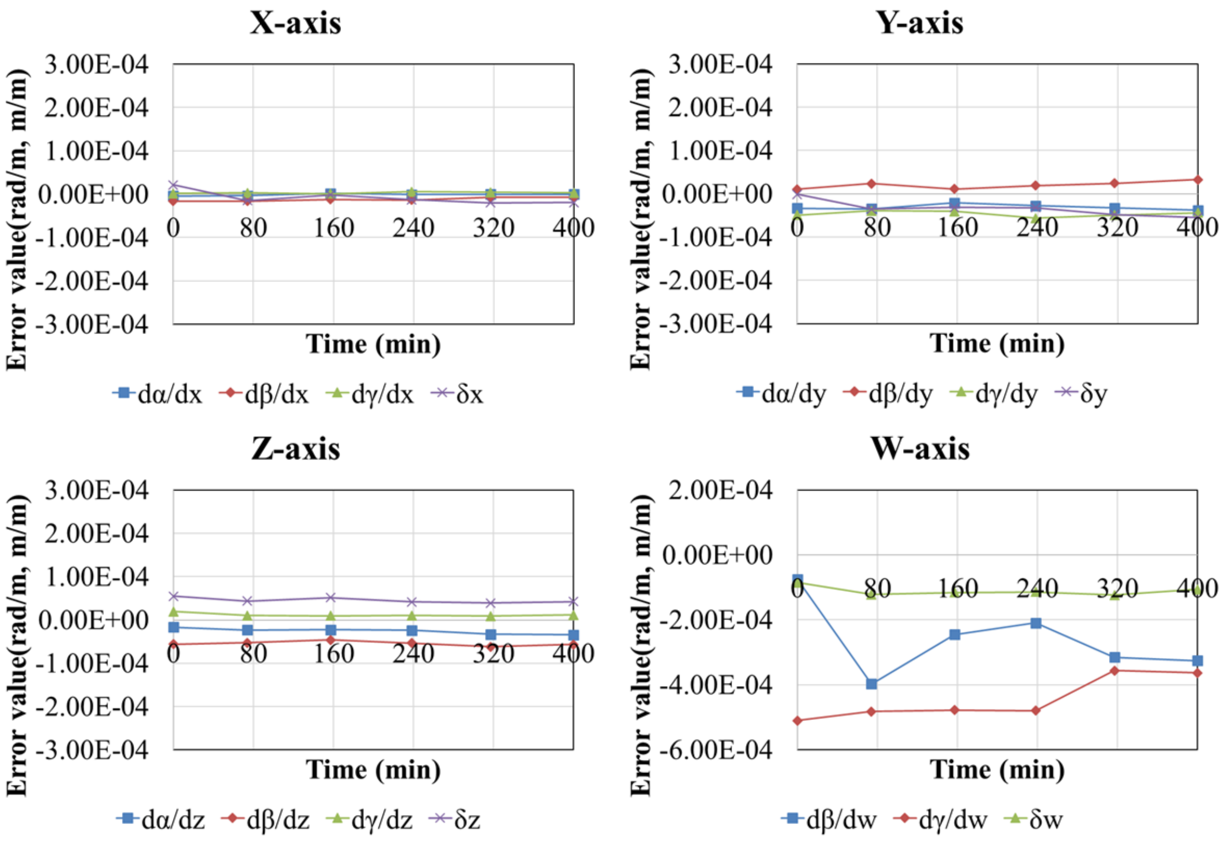

Figure 4.8: Variations of error parameters over the course of the experiment

Only parameters associated with the Y and W axes show significant changes during four heating cycles. The spindle was turned on during the heating process, and the readings of the wireless temperature sensors as depicted in Figure 4.7 show that the temperature readings of W and Y-axis increased by 4.5 and 3°C, respectively. The other parameters related to the X and Z-axis, on the other hand, had less than a 2°C rise in temperature during the 400-minute heating process. By studying the thermal behaviour of each axis, one can understand the characteristics of the machine, which could be used in error avoidance. For example, it is observed that the error parameters associated with W-axis are varying significantly during an operation. It shows that the machine's positioning error caused by distortion of W-axis could be more significant. Therefore, the positioning error could be avoided by replacing a W-axis movement with a Z-axis movement. Besides, it is observed that some parameters do not vary significantly over time and could be considered constants (e.g. parameters associated with X and Z axes). By making such assumptions, the number of undetermined error parameters can be reduced and thus reduce the needed number of observations.

The experiment on the unconstrained K-optimal observer design (i.e., observer obtained without measurement point distribution constraints by solving Equation (4.7)) was also conducted. The statistics of the results are shown in shown in Table 4.7. First, because the machine spindle was not run, between measurement cycles, the observed errors stayed constant. The temperatures during the experiment were observed to remain constant to within 2°C. While the parameters identified based on the measurements prescribed by the observer explain about 70% of the observed error, it can be seen that in terms of magnitude, the residuals are higher than those seen in the constrained observer (shown in Table 4.4) as well as in the quasi-random measurements (shown in Table 3.4).

Table 4.7: Model performances over time(unconstrained K-optimal)

Machine State	Observed (μm)		Residual (μm)		%Decrease	
	Mean	Max	Mean	Max	Mean	Max
Initial	176.0	460.9	47.9	114.5	72.78%	75.16%
Heating	180.4	449.4	49.2	114.8	72.73%	74.45%
Heating	177.2	444.1	48.9	110.8	72.40%	75.05%
Heating	176.7	453.7	48.4	119.5	72.61%	73.66%
Heating	172.9	443.5	47.6	110.8	72.47%	75.02%
Cooling	169.8	427.4	49.8	130.4	70.67%	69.49%
Cooling	163.8	417.6	46.5	105.8	71.61%	74.66%

4.4 SUMMARY

In this chapter, the idea of using error observers to track the machine tool errors is introduced. An optimal observer design identifies a set of locations in the machine’s workspace at which to make error measurements, so that the information contained in the set to estimate the parameters of a given error model is maximized. The approach can be used for any of the many proposed volumetric/quasistatic machine tool error models. The concept of applying the K-optimal design that minimizes the sensitivity of measurement errors on the parameter estimates has been proposed. The use of a single optimality criterion in the observer design leads to localization of the measurement points either near the center of the workspace or at its boundaries. To overcome the tendency, constraint volumes are used to uniformly distribute the observer points over the workspace. Redundancy (more measurement points than the minimum needed) is also introduced to guard against the effects of model inadequacy. Constrained and unconstrained observer designs based on the K-optimal design criterion have been generated for a 5-axis machine.

CHAPTER 5. POINT-SET BASED METROLOGY FOR PLANAR SURFACES

In Chapter 5, the concept of point-set based metrology combined with virtual gaging that not only verifies the finished specification of the manufactured part, but also adjust to the variability as it accrues between manufacturing steps of a part is proposed and verified experimentally. Section 5.1 introduces the problem, while mathematical model of virtual gages is proposed in Section 5.2 proposed. Section 5.3 shows the experimental work. The procedures of virtual gage analysis are proposed, and a prototype problem is solved to validate the idea of virtual gage analysis in Section 5.4. The generalized virtual gage analysis with slack variables is tested in Section 5.5, while Section 5.6 presents a summary drawn from this chapter.

5.1 INTRODUCTION

A virtual gage is a digital simulation that combines real data (measured from an artifact) with a computer representation of a condition or test that the data should satisfy. For example, the former might be a point-set extracted from the scanned data of a part while the latter might be the equation of a plane extracted for part model that represents an ideal material condition that must be satisfied by all points in the aforementioned point-set. The virtual gage assembles these entities into a common reference frame, creates the appropriate set of inequality conditions to be satisfied, and then checks them on each point in the point-set. Figure 5.1(a) shows a typical situation encountered in the test case of deciding the acceptability of a casting for final machining. To produce an acceptable finished part, a machining allowance of d_o is desired on each machined surface. The virtual gage software computes the rigid body transform (translation and rotation) of the point-set to produce each point in the set representing the casting surface satisfies the gage equation (given in Figure 5.1(b)). Obviously, a part will have several such specifications and the virtual gage software is expected to find a single rigid body transformation with which to simultaneously satisfy all of them. One may introduce conditional steps into the procedure, for example, if all the gages can be satisfied, then the software should "equalize" the allowances on all machined surfaces. As is apparent, each virtual gage introduces one or more sets of inequality constraints as does the conjoining of the reference frames of point-set and gage planes. Thus the problem becomes a constrained optimization problem, where an optimal solution satisfies the virtual gage constraints and minimizes or maximizers some objective (such as difference in machining allowance on all machined faces).

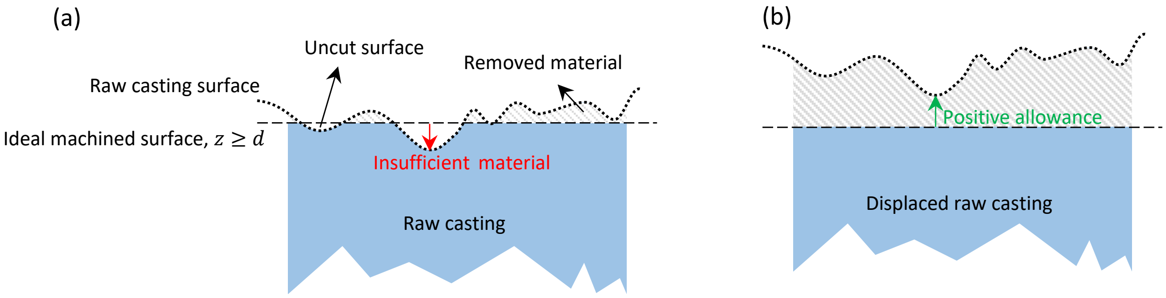

Figure 5.1: Figure 1. Virtual gage and point-set before (a) and after (b) adjustment

Thus, in summary, to setup a virtual gaging problem, a point-cloud representing the physical part and virtual gages idealized geometrical surfaces and representing the dimensional tolerance and allowance specifications are required. In the GD&T standard, dimensions and tolerances are defined based on a reference or datum coordinate frame. This may be different from the coordinate system in which machining is programmed. For simplicity, a homogeneous transformation is applied so that the datum coordinate system is made coincident with the machining programming coordinate system.

5.2 FORMULATION OF VIRTUAL GAGE

5.2.1 SINGLE VIRTUAL GAGE PROBLEM

Let $\vec{r}_i$ be the i^{th} point in the point-set, S that represents a planar surface on a part. After the rigid body transformation, T is applied, the shortest distance from $\vec{r}_i$ to some plane, A, with equation, $ax + by + cz + d = 0$ is given by,

$$e_i = [a \quad b \quad c \quad d]T \begin{bmatrix} \vec{r}_i \\ 1 \end{bmatrix}. \tag{5.1}$$

With small angle assumptions, a rigid body transformation can be written as,

$$T = \begin{bmatrix} 1 & -\alpha & \beta & \Delta x \\ \alpha & 1 & -\gamma & \Delta y \\ -\beta & \gamma & 1 & \Delta z \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{5.2}$$

where α, β and γ are the roll, pitch yaw angles in radian for angular motion and Δx, Δy and Δz describe linear motion.

Ω is defined as the set of all rigid body transformations where $-\delta_r \leq \alpha, \beta, \gamma \leq \delta_r$ and $-\delta_t \leq \Delta x, \Delta y, \Delta z \leq \delta_t$ (the angular rotations are limited by δ_r and the translations are limited by δ_t), so that the optimization problem can be bounded.

A 3-D planar virtual gage problem that attempts to make plane A, a support plane for S at a minimum distance of ε from it while minimizing the distance of the farthest point in S from it can be expressed as a constrained min-max linear programming problem [19], [62],

$$\min_{T \in \Omega} q \quad such\ that\ 0 \leq \varepsilon \leq e_i \leq q, \forall\, r_i \in S, \tag{5.3}$$

where q is the target function as well as an upper bound of all the distance of points r_i in S from A, and ε is the desired “clearance from” or “allowance for” the gage to the point-set as shown in Figure 5.2.

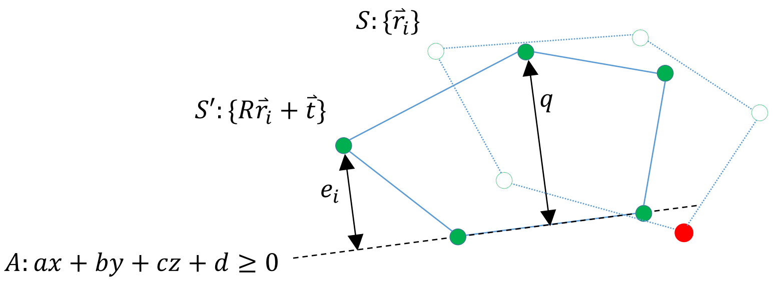

Figure 5.2: Optimized rigid body transformation is found by minimizing the maximal distance

5.2.2 MULTIPLE VIRTUAL GAGE PROBLEM

For a part with multiple virtual gages specified, a single homogeneous transformation matrix, T is used to simultaneously displace (rotate and translate) all the point-sets so that their correspoding gage planes become supporting. For example, three point-sets as shown in Figure 5.3 are cheched against three virtual gages. Two points (marked as solid circles) do not satisfy their corrsponding virtual gage. A single homogeneous transformation matrix, T, is used to simultaneously displace (rotate and translate) all the point-sets so that their corresponding gage planes equations are satisfied.

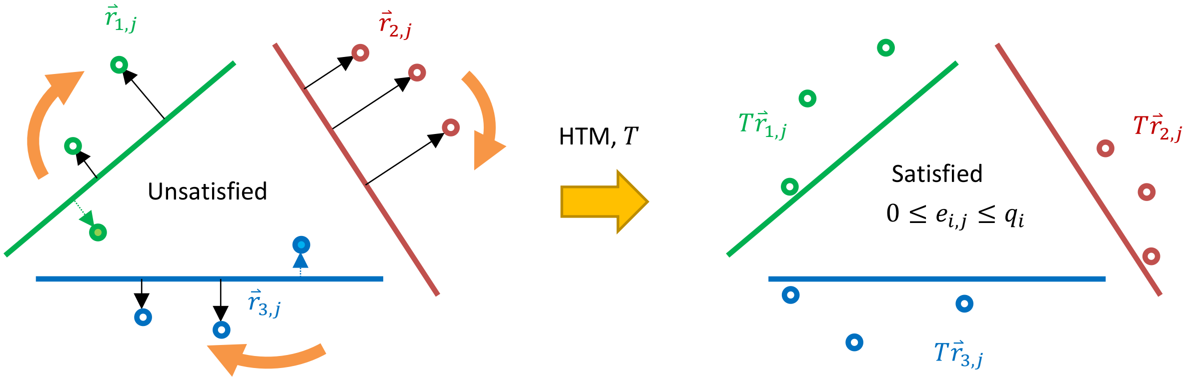

Figure 5.3: An HTM rigidly displaces the point-sets so that their corresponding gage planes are satisfied

In an n-gage problem, n pairs of point-sets and gage planes $\{S_i, p_i\}, i = 1, 2, \dots n$ are considered. The distance $e_{i,j}$ of the j^th^ point, $\vec{r}_{i,j}$ in i^th^ point-set, S_i to the i^th^ gage plane, p_i is given by:

$$e_{i,j} = \vec{p}_i^{\,T} T \begin{bmatrix} \vec{r}_{i,j} \\ 1 \end{bmatrix} \leq q_i, \tag{5.4}$$

where q_i represents an upper bound on $e_{i,j}$, and $\vec{p}_i^{\,T} = [a_i \quad b_i \quad c_i \quad d_i]$ is the coefficient vector of the i^th^ gage plane, p_i with plane equation $a_i x + b_i y + c_i z + d_i = 0$, where $\|[a_i \quad b_i \quad c_i]^T\| = 1$.

The algorithm seeks to minimize the weighted sum of n distance upper bounds,

$$\min_{T \in \Omega} \sum_{i=1}^{n} f_i q_i \ \ni 0 \leq e_{i,j} \leq q_i \ \forall \, r_{i,j} \in S_i, i = 1, \dots n, \tag{5.5}$$

where f_i, is a non-negative weighting coefficient for the distance measure, q_i of the i^th^ virtual gage has elements.

The constraints in Equation (5.5) can be written as linear inequalities,

$$\begin{bmatrix} b_i x_{i,j} - a_i y_{i,j} & a_i z_{i,j} - c_i x_{i,j} & c_i y_{i,j} - b_i z_{i,j} & a_i & b_i & c_i & -1 \\ a_i y_{i,j} - b_i x_{i,j} & c_i x_{i,j} - a_i z_{i,j} & b_i z_{i,j} - c_i y_{i,j} & -a_i & -b_i & -c_i & 0 \end{bmatrix} \vec{v} \leq \begin{bmatrix} -h_{i,j} \\ h_{i,j} \end{bmatrix} \tag{5.6}$$

where $\vec{v} = [\alpha \; \beta \; \gamma \; \Delta x \; \Delta y \; \Delta z \; q_i]^T$ is the vector of controllable variables, $h_{i,j} = a_i x_{i,j} + b_i y_{i,j} + c_i z_{i,j} + d_i$ is the discrepancy from $\vec{r}_{i,j}$ to the plane.

The linear program of Equation (5.5) is bounded because Ω, the domain of its decision space, is bounded. If the linear programming problem is found to be infeasible, there are two possible reasons:

1. Ω, the limitation on allowable transformations T is restricting the algorithm for obtaining a feasible solution:

The reason for the infeasibility can be verified by examining the Lagrange multipliers of the constraints at termination. If any of the constraints that correspond to the limits imposed by Ω have non-zero values, then one can attribute the infeasibility due to restricting the set of feasible rigid-body transformation. The situation can be resolved either by relaxing Ω or by using a sequential programming approach, which updates the problem linear program at the current value of T and restarting the optimization. As can be seen in Figure 5.4, the two gage problem cannot be solved since the point-set is not allowed to rotate about Z direction. There exists no translational matrix such that the two gages can be satisfied simultaneously. However, if rotation about Z direction is possible in the HTM, the problem is solvable.

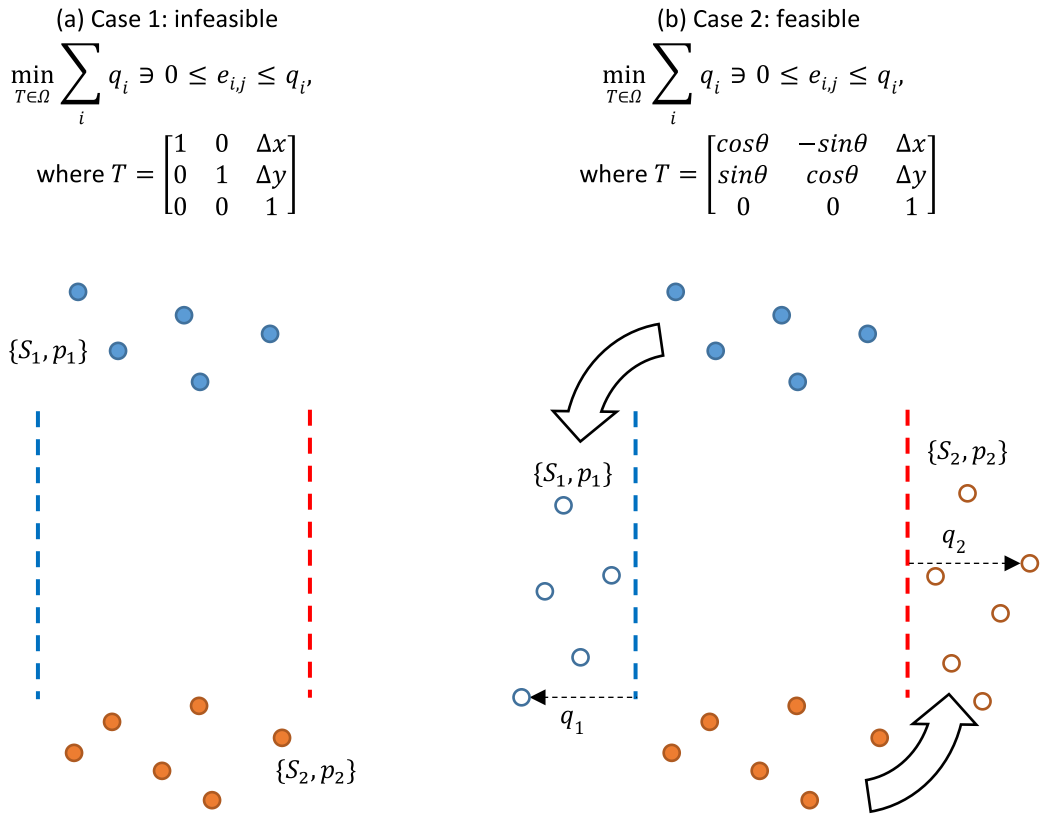

Figure 5.4: 2-D examples, problem is: (a) infeasible due to rotation is not allowed; (b) feasible

If none of the tight constraints are associated with Ω, then the infeasibility is due to dimensional defects (e.g., insufficient material). In the context of the casting metrology problem described earlier, one can choose to reduce clearances or machining allowance requirements, relax constraints post by gages deemed less important than other, or reject the casting.

2. The infeasibility is due to insufficient material on the part:

As can be seen in Figure 5.5, the 1st and 2nd pair of virtual gage and point-set cause the infeasibility of the problem formulated by Equation (5.5). Since the distance between gage 1 and 2 are shorter than the distance between the two closest points in S_1 and S_2, there does not exist any displacement that can satisfy these two gages at the same time. In this case, further analysis is required to detect the gages that cannot be satisfied.

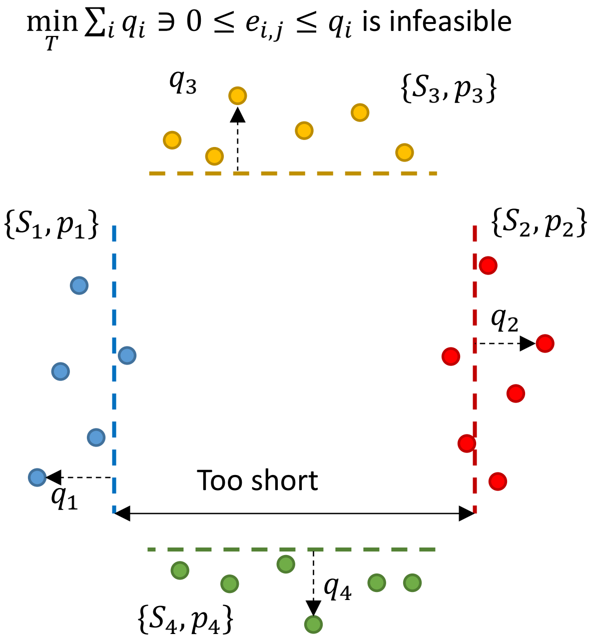

Figure 5.5: Problem is infeasible since insufficient material on features 1 and 2.

5.2.3 MULTIPLE GAGE PROBLEM WITH INSUFFICIENT MATERIAL

A general form of linear programming problem can be written as

$$\min_{x} f^{T} x \ni Ax \leq b, \tag{5.7}$$

where x is the vector of controllable variables, f is the coefficient vector, $Ax \leq b$ are the constraints of linear inequalities.

The existence of the solution to Equation (5.7) depends on only the constraints, $Ax \leq b$. The feasible region is the set of all solutions that satisfy $Ax \leq b$. There are three different cases of feasible region and the associated constraints:

(a) $Ax \leq b$ has infinitely many solutions. The constraints are loose.

(b) $Ax \leq b$ has unique solution. At least two constraints are tight (can be satisfied exactly).

(c) $Ax \leq b$ has no solution. At least two constraints are broken (infeasible).

A virtual gage is a set of linear constraints. Similarly, a virtual gage can be loose, tight or broken when it appears as constraints in a linear programming problem as shown in Figure 5.6.

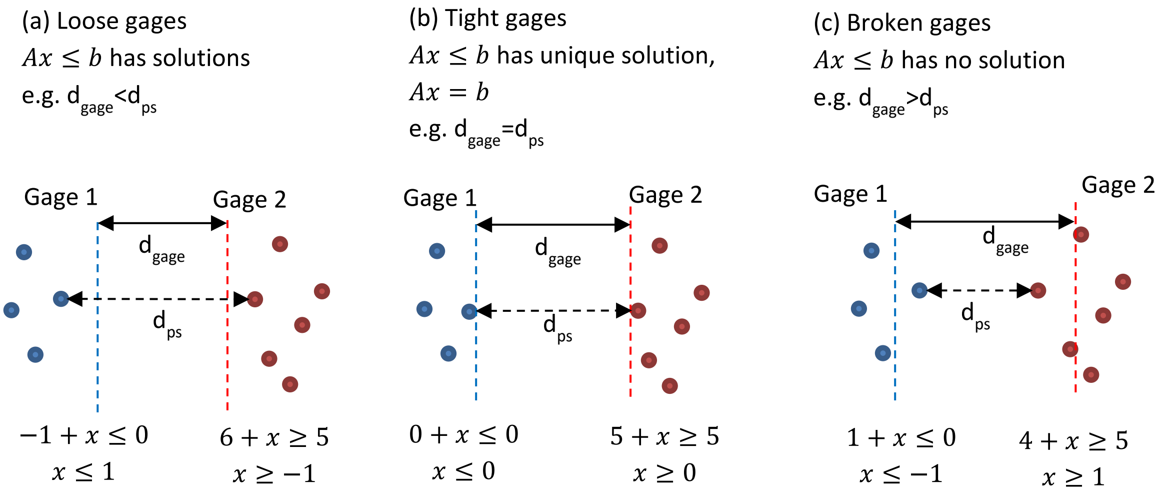

Figure 5.6: Schematics of three types of virtual gages: (a) loose; (b) tight; (c) broken

When there exists no feasible solution for Equation (5.5), mathematically speaking, the feasible region formed by the corresponding constraints does not exist. Without further analysis, the only conclusion can be made is that the raw casting has defectives, so the final casting will not satisfy all GD&T requirements. However, the defectives might be fixed by manually adding material on the defective surfaces and still gets acceptable finished part. If the features are too defective to be manually compensated, rejection could be considered. Therefore, more quantifiable information regarding the surfaces, where the material conditions cannot be satisfied are required to decide to reject or accept the part.

This examination of unsatisfiable feature's machining allowance is implemented by allowing the gages to move along the direction of their normal vectors by the values of their slack variables, which will be introduced to the linear programming formulation. A virtual gage with a positive slack variable cannot be satisfied by any HTM, and its slack variable represents the amount of negative machining allowance. In order to find the infeasible virtual gages, slack variables, denoted by s_i are introduced:

$$\min_{T\in\Omega}\sum_{i=1}^{n} g_i s_i \ \ni -s_i \leq e_{i,j} \leq q_i \ \forall \ r_{i,j} \in S_i, i = 1,2 \dots n, \tag{5.8}$$

where $e_{i,j}$ is the discrepancy of the jth point in the point-set paired with the ith virtual gage defined in Equation (5.4), and g_i is the ith non-negative weighting coefficient for the ith slack variable, s_i for the ith virtual gage.

The solution to Equation (5.8) suggests the least summation of the slack variables. The solution to the ith slack variables is found to be s_i^*, where $s_i^* > 0$. If $s_i^* > 0$, the ith virtual gage is infeasible. By releasing all infeasible gages by the values of the slack variables, Equation (5.5) becomes:

$$\min_{T\in\Omega}\sum_{i=1}^{n} f_i q_i \ \ni -s_i^* \leq e_{i,j} \leq q_i \ \forall \ r_{i,j} \in S_i, i = 1,2 \dots n, \tag{5.9}$$

where f_i represents non-negative weighting coefficient, $e_{i,j}$ is the discrepancy of the jth point in the point-set paired with the ith virtual gage defined in Equation (5.4) and s_i^* is the slack variable obtained from Equation (5.8).

Assume that the kth virtual gage is found to be infeasible with positive slack variable, s_k^*>0, as the 1st gage shown in Figure 5.7(a). There exists at least one gage associated with the kth gage that cannot be satisfied if the kth gage is not allowed to be released. As can be seen in Figure 5.7(b), the 2nd and 3rd gages must be moved, or the problem is unsolvable. The minimum group of the gages that cannot be satisfied simultaneously is called an infeasible gage group. The other members in the group can be found be solving a linear programming problem with equality constraint of $s_k^* = 0$:

$$\min_{T\in\Omega}\sum_{i=1}^{n} f_i q_i \ \ni -s_i^* \le e_{i,j} \le q_i \ \forall\, r_{i,j} \in S_i \ (i = 1{\sim}n), s_k^* = 0, \tag{5.10}$$

where f_i represents non-negative weighting coefficient, $e_{i,j}$ is the discrepancy of the jth point in the point-set paired with the ith virtual gage defined in Equation (5.4), s_i^* is the slack variable obtained from Equation (5.8) and s_k^* is forced to be zero.

The other gage members that cannot be satisfied with the kth gage are given by the gages with positive slack variables. There are always two or more gages in the infeasible group. Figure 5.7 shows an example of an infeasible gage group of three gages.

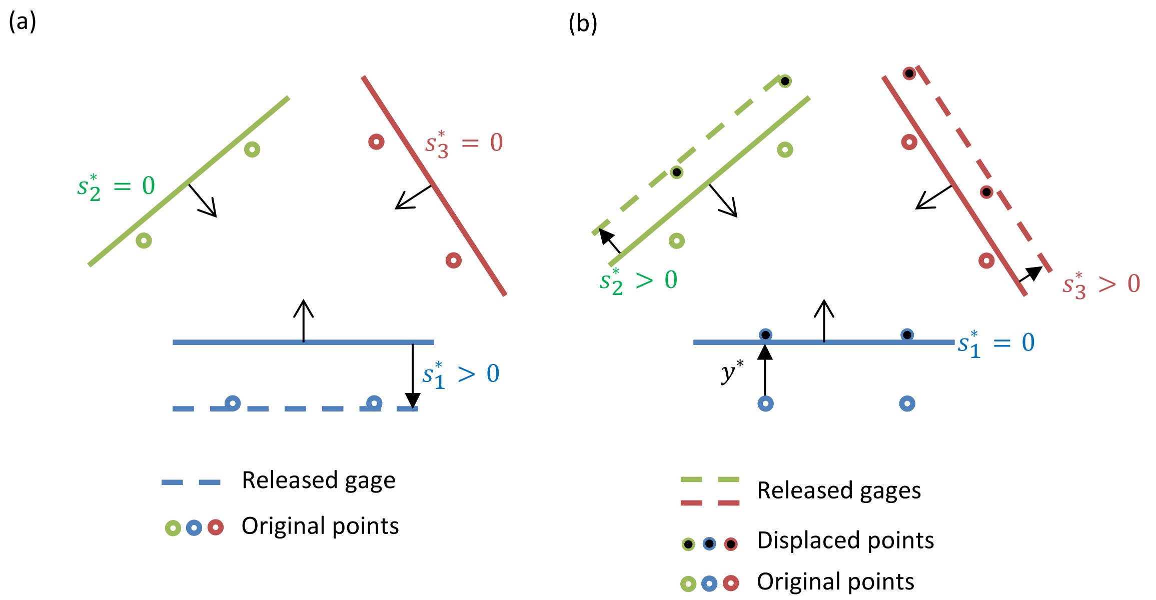

Figure 5.7: Schematics of infeasible group: (a) the 1st gage is infeasible; (b) the 2nd and 3rd gages must be released if 1st gage is not released

It must be noted that Equation (5.9) is always solvable since all infeasible virtual gages are released to be feasible. Although the released virtual gages can be satisfied, the allowances associated with these released gages will be all zero. As explained in Figure 5.6(b). The feasible region is a single point. Hence, the solution to Equation (5.9) must be that feasible point, which is no longer objective function dependent. In fact, the ways of releasing the constraints such that the feasible region exists are not unique, and they are all solutions of Equation (5.9), as shown in Figure 5.7. Releasing every constraint by same amount is a better strategy because it reduces the maximum amount of violation of the infeasible group and hence reduces the chance of rejection.

To conclude, the downsides of solving Equation (5.9) are:

(a) There are infinitely many ways to release the constraints.

(b) Some constraints may be over released.

(c) The optimal solution is independent on objective function.

Therefore, a generalized problem is given by minimizing the summation of the original objective function, all the slacks and the largest slack:

$$\min_{T\in\Omega}\sum_{i=1}^{n} f_i q_i + g_i s_i + u \ \ni -u \le -s_i \le e_{i,j} \le q_i \ \forall\, r_{i,j} \in S_i, i = 1,2 \dots n, \tag{5.11}$$

where $e_{i,j}$ is the discrepancy of the jth point in the point-set paired with the ith virtual gage defined in Equation (5.4), f_i and g_i are the non-negative weighting coefficients for the distance measure, q_i and the slack variable, s_i of the ith virtual gage and u is the largest slack variable.

Similarly, the constraints given by point $\vec{r}_{i,j}$ in are written as linear inequalities,

$$\begin{bmatrix} b_i x_{i,j} - a_i y_{i,j} & a_i z_{i,j} - c_i x_{i,j} & c_i y_{i,j} - b_i z_{i,j} & a_i & b_i & c_i & -1 & 0 \\ a_i y_{i,j} - b_i x_{i,j} & c_i x_{i,j} - a_i z_{i,j} & b_i z_{i,j} - c_i y_{i,j} & -a_i & -b_i & -c_i & 0 & -1 \end{bmatrix} \vec{v}' \le \begin{bmatrix} -h_{i,j} \\ h_{i,j} \end{bmatrix} \tag{5.12}$$

where $\vec{v}' = [\alpha\ \beta\ \gamma\ \Delta x\ \Delta y\ \Delta z\ q_i\ s_i]^T$ is the vector of controllable variables, $h_{i,j} = a_i x_{i,j} + b_i y_{i,j} + c_i z_{i,j} + d_i$ is the discrepancy from $\vec{r}_{i,j}$ to the plane.

Solution to Equation (5.11) not only detects the features that cannot be satisfied but also takes into consideration on optimizing the original objective function. More importantly, Equation (5.11) always has solution, even when some material conditions of the part cannot be satisfied. would also equalize the amounts of release on infeasible gages because the largest slack in the objective, u is introduced in the objective function. If some slack variables are found to be positive, their corresponding gage plane will be violated by the same amount and the material condition cannot be satisfied. Decision by human is required to determine the acceptability of a defective part based on the functionalities of the defective features and the values of the slack variables. If all slack variables are found to be zeros, all material conditions are satisfiable. The conformity of the part is guaranteed.

5.3 EXPERIMENTAL WORK

A test part is designed to verify the feasibility of virtual gage analysis and shown in Figure 5.8(a) is scanned to verify the feasibility of virtual gage analysis. A Keyence LJ-V7200 laser scanner is mounted on the spindle of a machine tool, and Figure 5.8(b) schematically shows the setup. The scan paths are generated by VERICUT®, a machine tool simulation tool and verified on the actual machine to ensure that no collision occurs. During the experiment, the machine tool is commanded to move the laser scanner along the scanning paths at the feed rate 100 mm/s with scan frequency of 1000 Hz.

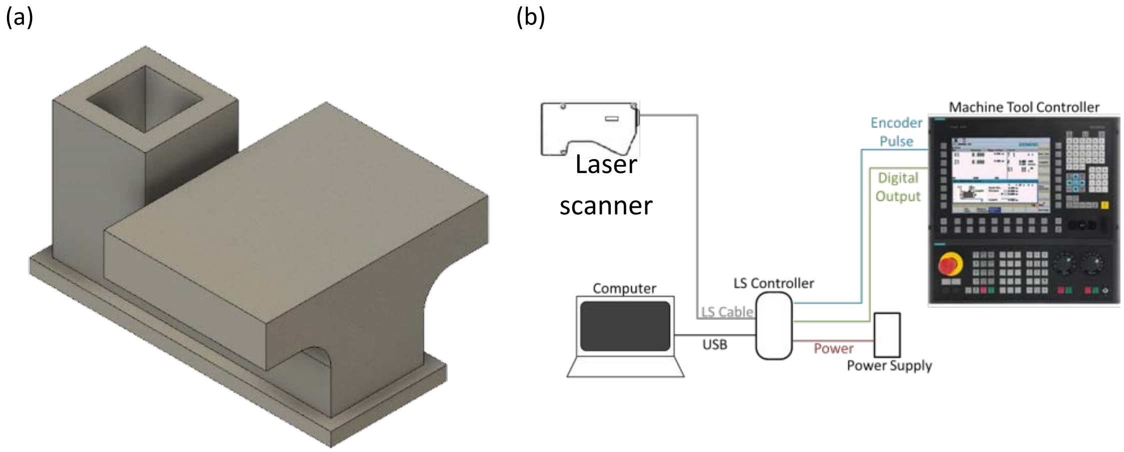

Figure 5.8: (a) Nominal model of the raw casting; (b) Schematic of using machine tool to move laser scanner and collect data

To test the validity of the virtual gage analysis for the casting with fixturing error, spacers were placed between locator X_1 and casting as well as between the Y-locators and casting to simulate the translation error as shown in Figure 5.9(a). Similarly, a small angular error about the Z-axis was introduced by the spacers between the locator Z_1, Z_2 and the casting. The casting with fixturing error was scanned and 130976 points were collected as shown in Figure 5.9(b).

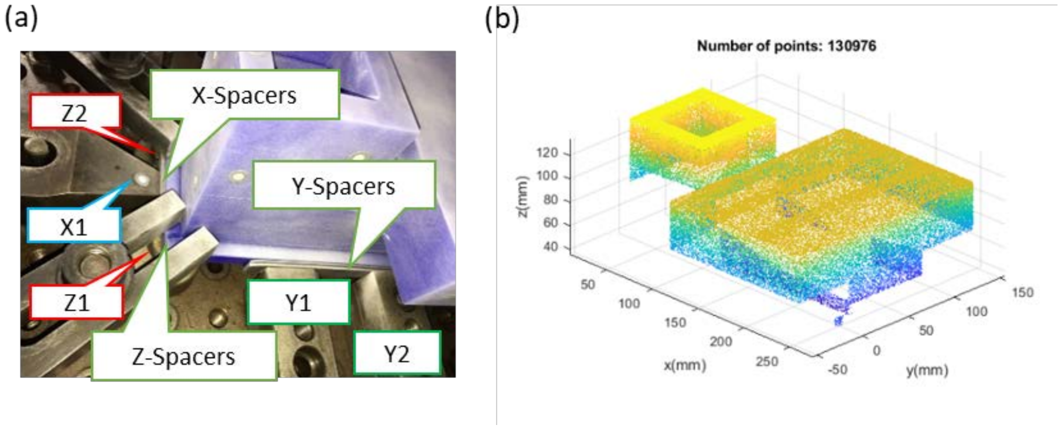

Figure 5.9: (a) Locators and spacers setup; (b) Collected point cloud of the casting

5.4 VIRTUAL GAGE ANALYSIS

A 2-step virtual gage analysis procedure is developed as shown in Figure 5.10. First, the point cloud data is aligned with the raw casting's CAD model and divided into sub-point-sets to represent functional features. The redundant points in the sub-point-sets are removed using a convex hull filter. Second, virtual gages are parameterized as $a_i x + b_i y + c_i z + d_i \geq 0$ based on CAD model and GD&T requirements of the final casting. The parameterized gages and extracted sub-point-sets are used to formulate linear programming problem with constraints. The optimal solution obtained can be used to correct the machining coordinate system and predict the machining allowance of each feature. Each step is explained in detail as follows.

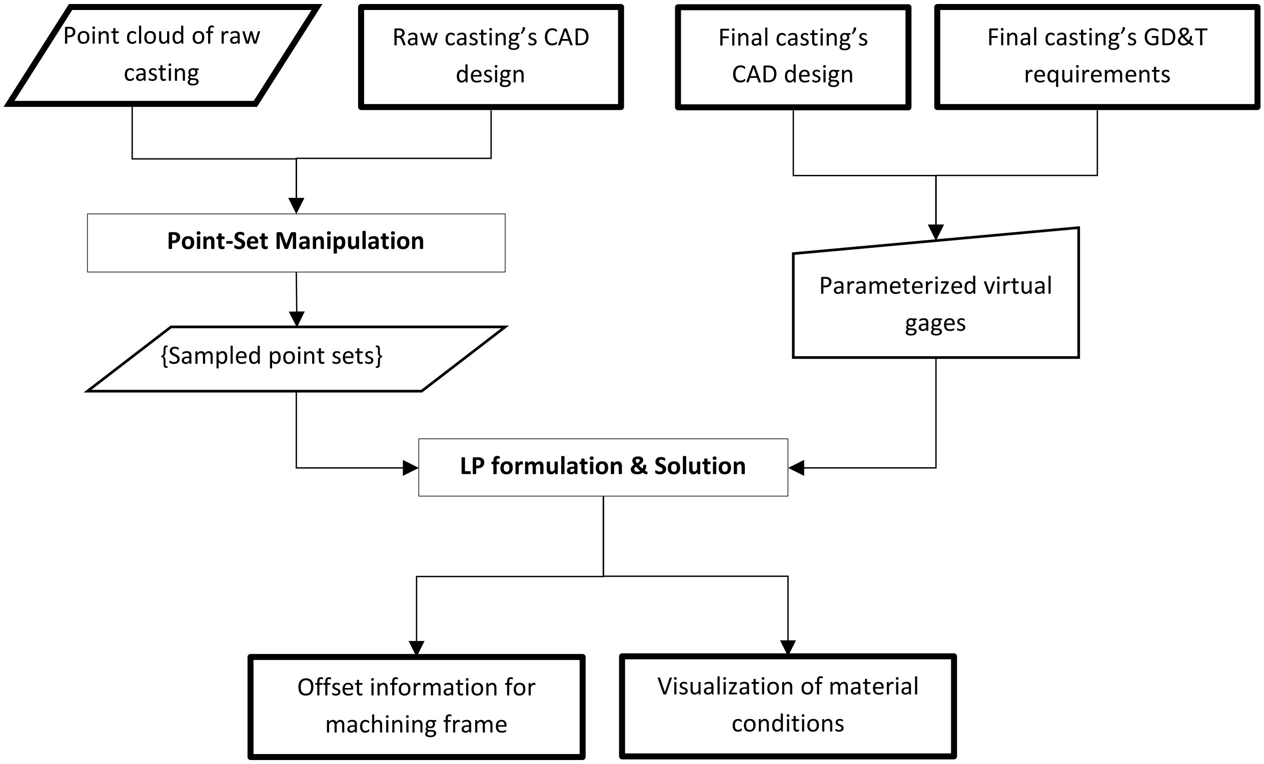

Figure 5.10: 2-step procedure of virtual gage analysis

5.4.1 POINT-SET MANIPULATION

Data alignment

The virtual gage integrates the gage planes with sample data (point cloud) from the physical part. The gage planes are extracted from CAD models, and compatible reference frames between the point cloud data and the CAD model are established. In this example, a 3-2-1 location scheme to identify three orthogonal planes that serve as primary, secondary and tertiary data planes are used. To locate the reference frame of the part, the features of the locators are also scanned before the part is fixed on the locators. A locator frame is fitted using the point-set representing the locators.

Data that corresponds to locator planes is extracted from the point cloud. The primary datum or locating plane, by convention, is selected as the XY plane of the reference frame that defines its z-direction. Likewise, the XZ-plane and YZ-plane are the secondary and tertiary datum reference, respectively. Each of these locator planes is obtained by identifying best-fit support planes for their corresponding point-sets. The planes are identified sequentially, with the primary (XY) plane being identified first, the secondary (XZ) plane next with the additional constraint that it is perpendicular to primary location plane, and finally the tertiary (YZ) plane is identified with constraints added to ensure that it is perpendicular to the other locator planes.

The primary datum plane, $a_z x_i + b_z y_i + z_i + d_z = 0$ is obtained by using the point-sets associated with the location surfaces of the primary datum and finding the best supporting plane with the following linear optimization problem:

$$\underset{\vec{p}_z, d_z}{min} \max(\vec{p}_z{}^T \cdot \vec{r}_i + d_z)\ s.t. \vec{p}_z \cdot \vec{r}_i + d_z \geq 0;\ \|\vec{p}_z\| = 1\ \forall i, \tag{5.13}$$

where $\vec{p}_z{}^T = [a_z \quad b_z \quad c_z]$ is the normal to the identified plane and d_z locates it in space and $\vec{r}_i$ represents a point in the point-set(s) associated with the primary location plane. The above optimization problem can be linearized and solved as a sequential linear program by preprocessing the data (For example, setting up the search for the optimal support plane is as a small perturbation on the best-fit least square plane for the given data). It should be noted that this formulation can be reduced to the linear programming formulation given in Equation (5.5). Additionally, since a support plane is searched, the point-set can be reduced (thus reducing the number of constraints), by only retaining the points on the convex hull of the point-set.

After the primary datum plane is fitted, the secondary datum plane can be obtained by same formulation given in Equation (5.13) but with the addition of a constraint to enforce the perpendicularity requirement between the identified primary datum and the secondary datum. Thus, the following constrained optimization problem is considered:

$$\min_{\vec{p}_y, d_y} max(\vec{p}_y{}^T \cdot \vec{r}_j + d_y)\ \ni\ \vec{p}_y \cdot \vec{r}_j + d_y \geq 0;\ \|\vec{p}_y\| = 1;\ \vec{p}_y{}^T \cdot \vec{p}_z = 0\ \forall j\ , \tag{5.14}$$

where $\vec{p}_y{}^T = [a_y \quad b_y \quad c_y]$ is the normal to the identified secondary normal, and $\vec{r}_j$ represents a point in the point-cloud from the surfaces associated with the secondary location plane.

The normal vector of the tertiary datum plane is fixed as it must be the cross product of $\vec{p}_y$ and $\vec{p}_z$, i.e, $\vec{p}_x = \vec{p}_y \times \vec{p}_z$. The optimization of the tertiary plan can then be written as:

$$\min_{d_x} \max_{k} d_x \ni \vec{p}_x \cdot \vec{r}_k + d_x > 0 \; \forall k, \tag{5.15}$$

where $\vec{p}_x = \vec{p}_y \times \vec{p}_z$, $\vec{p}_x{}^T = [a_x \quad b_x \quad c_x]$ is the normal vector of the virtual X-plane and $\vec{r}_k$ represents a point in the point-cloud of the X-locators.

After three datum planes are fitted, the locator frame can be constructed. The origin of the frame is given by solving three plane equations,

$$\begin{bmatrix} o_x \\ o_y \\ o_z \end{bmatrix} = -\begin{bmatrix} a_x & b_x & c_x \\ a_y & b_y & c_y \\ a_z & b_z & c_z \end{bmatrix}^{-1} \begin{bmatrix} d_x \\ d_y \\ d_z \end{bmatrix}. \tag{5.16}$$

A coordinate system can be represented by a 4 by 4 homogeneous transformation matrix (HTM). The locator frame identified above, in the coordinate systems of the point-sets (scanner's coordinate system) is given by,

$$C_L^S = \begin{bmatrix} a_x & a_y & a_z & o_x \\ b_x & b_y & b_z & o_y \\ c_x & c_y & c_z & o_z \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{5.17}$$

As an example of the solid model is constructed in a modeling frame, C_L^C the with the same primary, secondary and tertiary locator surfaces, then, located in the point-cloud data reference system, it should be a 4 by 4 identity matrix. Brought into the point-cloud reference frame (scanner frame), it would locate the part at its origin with its locator surfaces, aligned with the principal (XY, YZ, and ZX) planes. When the part's CAD model is brought into the point-cloud coordinate system, it is aligned with the locator frame, but situated at the origin. The coordinate transformation

H moves the point-cloud so that its locator frame is aligned with the locator frame of the part model. The point-cloud locator frame, C_L^S, identified by the aforementioned procedure coincident with the locator frame attached to the model is given by:

$$H \times C_L^S = C_L^C = I_4. \tag{5.18}$$

Thus,

$$H = (C_L^S)^{-1}. \tag{5.19}$$

Therefore, HTM, H brings the point-cloud into alignment with the CAD model frame with the same datum planes.

Feature extraction

After the point-cloud is aligned with the part CAD model, the point-cloud can be segmented in to point-sets, such that each set is associated with a virtual gage. Each of these point-sets can then be processes to remove redundant points from the set.

The extraction of points from the point cloud to form a point-set for a virtual gage is accomplished by creating sampling volumes and classifying (deciding whether a point is in or out) the points against these volumes. These sampling volumes are associated with important features and constructed with the part CAD model (because the CAD environment has the appropriate tools to create and locate them relative to a face in part CAD model that will become a gage plane in the virtual gage). Besides identifying the points to be included in the point-set for a virtual gage, the sampling volumes are used for the removal of scanning artifacts (especially those produced near the edges of a surface during scanning). Figure 5.11(a) schematically depicts the use of sampling volumes to extract a set of points from the point cloud, and Figure 5.11(b) shows the sampled point-set (green) from the point cloud(as shown in Figure 5.9(b)) after applying the sampling volume on a face. In the current stage, only rectangular boxes can be used as sampling volumes.

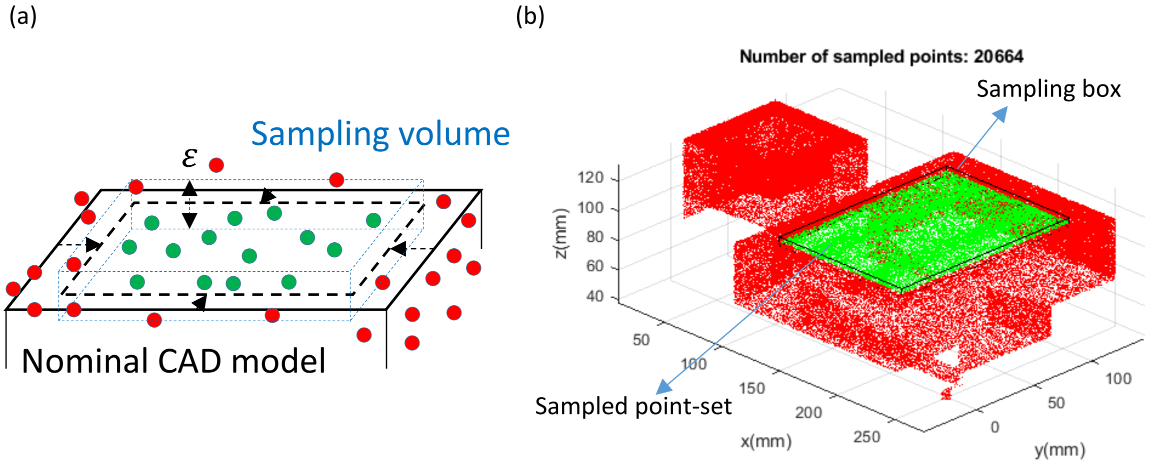

Figure 5.11: (a) Schematic segmentation of point-cloud data using sampling volumes; (b) An example of extracting a point-set for a virtual gage from the part's point-cloud

Data filtering

The constrained optimization algorithms that implement the virtual gages are computationally intensive. Dense point-sets generate many constraints for a virtual gage, many of which are redundant. To reduce the computational time required to check a virtual gage, reducing the number of constraints by thinning down the associated point-sets by identifying and eliminating redundant points is required.

Since virtual gages essentially identify optimal support or classifying planes for point-sets (i.e., planes that define half-spaces that either contain all the points or none), convex closures of the point-set play an important role in characterizing them relative to the gage planes. Therefore, only those points involved in the definition of a convex closure or hull (i.e., its vertices) need be considered. Other points, interior to the closure can be eliminated without fear of changing any metrics relative to the gage planes.

The convex hull of a finite point-set, S is defined by the convex combination,

$$CH(S) = \left\{\sum_{i=1}^{|S|} \alpha_i x_i \,\middle|\, (\forall i: \alpha_i \geq 0) \wedge \sum_{i=1}^{|S|} \alpha_i = 1\right\}, \qquad (5.20)$$

where x_i is the i^th^ point in S.

A finite point-set has unique convex hull, whose vertices, S' are the minimal possible subset of S that share the same convex hull. Thus,

$$CH(S) = CH(S'). \tag{5.21}$$

All extremal (minimum or maximum) distances between the S and a support or classifying plane are defined by points in S'. Hence, only the vertices of the convex hull need be considered in the virtual gage algorithms. In general, $|S'| \ll |S|$ can be expected, and the computing of $CH(S)$ is computationally much less intensive than solving the constrained optimization problem of the virtual gages. Thus, by using the vertices of convex hulls of the point-sets instead of the entire point-set for the implementation of the virtual gage algorithm is able be to greatly reduce the computational effort. Figure 5.12 shows the use of a convex hull filter in replacing the point-set extracted by a sampling volume in Figure 5.11.

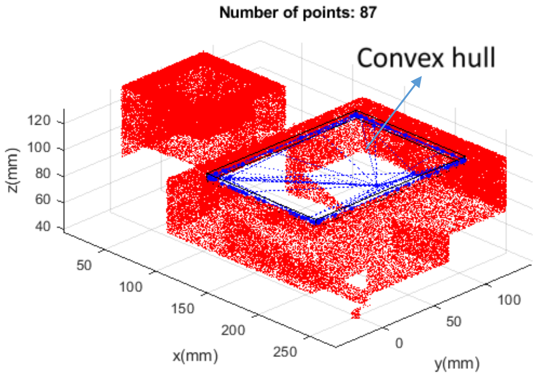

Figure 5.12: A point-set shown in Figure 5.11(b) of 20664 points, filtered to 87 points

5.4.2 LINEAR PROGRAMMING FORMULATION

A 3-D virtual gage analysis requires to setup a linear programming problem that minimizes the sum of maximal distances from gage planes to sub-point-sets. However, before formulating a linear programming problem, for every sub-point-set collected in Section 5.4.1, at least one corresponding virtual gage plane needs to be defined. Since the virtual gage usually cannot be seen in the CAD model, it should be referred to a reference feature (a reference plane or a datum surface), which can be found and defined in CAD software.

Having known the sub-point-sets and virtual gages, constraints for linear programming are defined. Every point provides two constraints, which must be converted to linear inequalities for the linear programming solver as shown in Equations (5.6) and (5.12). After all constraints are defined, a linear programming problem is formulated.

5.4.3 RESULTS AND DISCUSSION

As shown in Figure 5.14(a), three separate locators (Z1, Z2, Z3,) provided data for fitting the primary datum as the XY plane of the reference frame that defines the z-direction, followed by fitting the secondary and tertiary datum, XZ and YZ planes shown in Figure 5.13(b) and (c). All six scanned locators are shown in Figure 5.13(d).

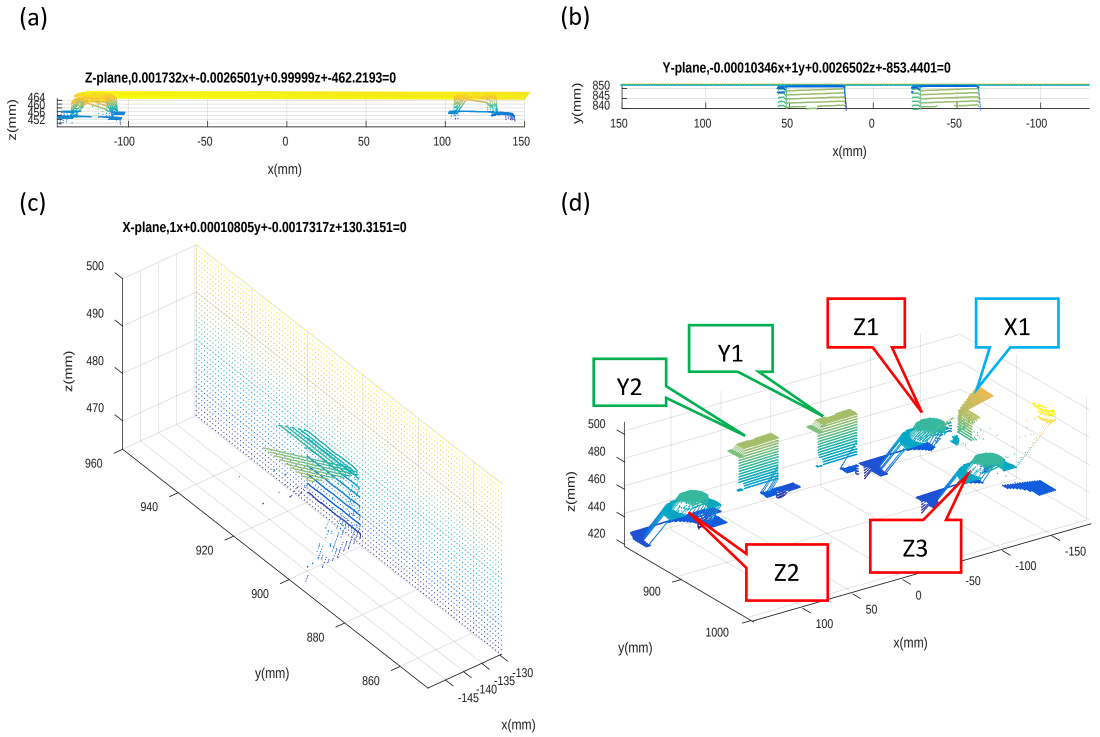

Figure 5.13: Fitted datum surfaces: (a) Z-plane; (b) Y-plane; (c) X-plane; (d) the scan data of the 3-2-1 locators

Figure 5.14(a) virtually shows the placement of the CAD model on the 3-2-1 locators. Equations (5.13)-(5.17) are solved to fit the point cloud locator frame, C_L^S. Figure 5.14(b) shows the point cloud and two their reference frames, C_L^S and C_L^C. The homogeneous transformation, H between two frames is given by solving Equations (5.18) and (5.19). By applying H to displace the point cloud, the point cloud is aligned with the CAD model as shown in Figure 5.14(c).

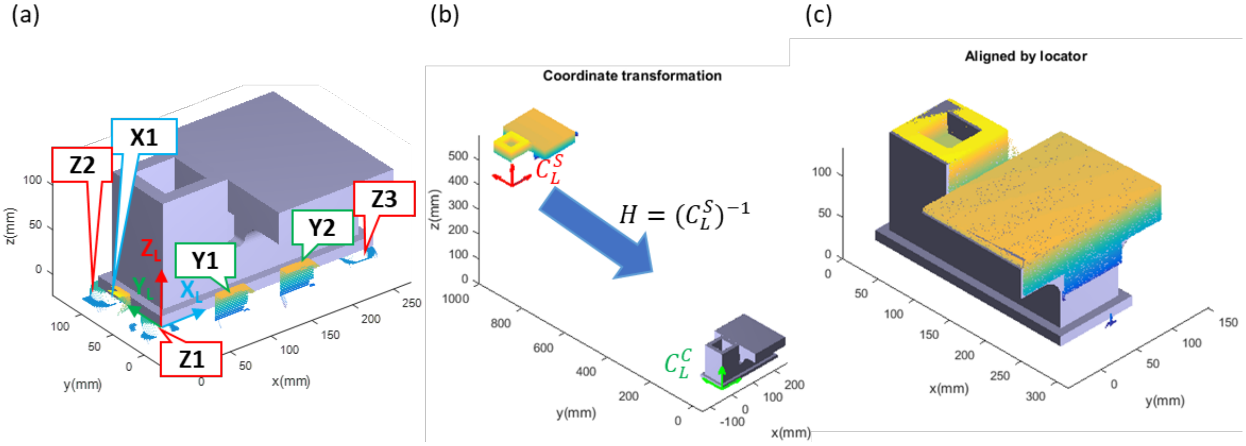

Figure 5.14: a)Virtually located CAD model in point cloud data scanned from the location surfaces of a fixture; (b) The raw point cloud and the CAD model; (c) The point cloud given in CAD model's frame

Initially, thirteen point-sets are extracted using sampling volumes to represent 13 associated planes of the casting as shown in Figure 5.15(a). However, seven of them, including the top and side of the flange, two inner walls and two outer walls and the top of the tower as shown in Figure 5.15(b) are used in the virtual gage analysis since they represent the critical features with tolerance specifications according to the print. Note that six of the seven faces including top of flange, four inner wall faces and top of the tower restrict only the linear translation along X and Z-axes as linear constraints. If all normal vectors of virtual gages are perpendicular to y-axis, then the motion along y-axis is unconstrained, and the solution for translation along y-axis, Δy is not unique. Hence, the side face of flange is also considered to provide restriction on translation along Y-direction for the uniqueness of solution.

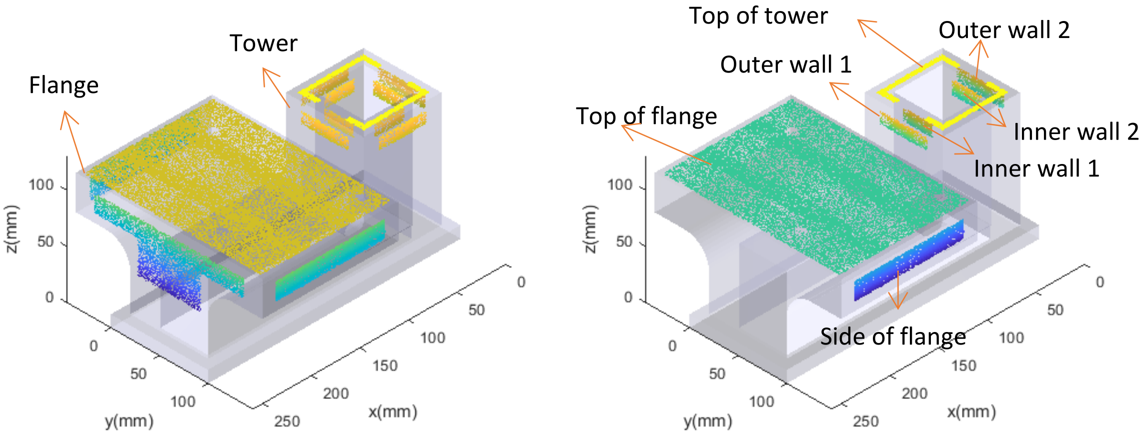

Figure 5.15: (a) Thirteen extracted point-sets; (b) seven point-sets used in virtual gage analysis

The GD&T requirements specified in the print include the minimum thickness of two walls and with a plus and minus 1.1 mm tolerance region for the machined flange surface. Hence, the set of points representing the top of the flange surface, as can be seen in Figure 5.15(b), is checked against the virtual gages, $z \geq 109.3$ and $z \leq 111.5$ for the specified tolerance region to determine if the flange surface can be machined properly. Other GD&T specifications are all minimum material conditions, which are checked only by single gage. The sets of parameters of all virtual gages obtained from CAD model are given in Table 5.1. All extracted point-sets are filtered by convex hull to remove redundant points. The size of seven point-sets representing seven features with and without filtering is given in Table 5.2.

Table 5.1: Virtual gage parameter sets, $a_i x + b_i y + c_i z + d_i \geq 0$

Feature	a_i	b_i	c_i	d_i(mm)
Top of flange	0	0	1	-109.30
Top of flange	0	0	-1	111.50
Side of flange	0	1	0	-145.80
Outer wall 1	1	0	0	-87.43
Outer wall 2	-1	0	0	12.42
Inner wall 1	-1	0	0	75.62
Inner wall 2	1	0	0	-23.48
Top of tower	0	0	1	-124.20

Table 5.2: Point-set size with and without filtering

Feature	Size before filtering	Size after filtering
Top of flange	20664	87
Side of flange	5361	79
Outer wall 1	841	57
Outer wall 2	1024	59
Inner wall 1	661	47
Inner wall 2	767	53
Top of tower	3482	97

Since the machine used in the test has three linear axes and one rotary axis, which only allows the table to rotate about Y-axis, the HTM is limited to four degrees-of-freedom,

$$T = \begin{bmatrix} 1 & 0 & \beta & \Delta x \\ 0 & 1 & 0 & \Delta y \\ -\beta & 0 & 1 & \Delta z \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{5.22}$$

where translational offsets, Δx, Δy and Δz are in mm and rotary offset of Y-axis, β is in radian and restricted by the small angular assumption, $-0.05 \leq \beta \leq 0.05$.

Since all virtual gages are considered equally important, the weighting coefficient f_i is chosen to be 1. A linear programming problem given by Equation (5.5) can be written as,

$$\min_{T \in \Omega} \sum_{i=1}^{8} q_i \ \ni 0 \leq [a_i \ b_i \ c_i \ d_i] T \begin{bmatrix} \vec{r}_{i,j} \\ 1 \end{bmatrix} \leq q_i, i = 1 \sim 8 \tag{5.23}$$

where $e_{i,j}$ and q_i are defined in Equation (5.4).

The linear programming problem is formulated with the linear constraints as given in Equation (5.6), and the linear programming solver returns a unique, optimal HTM for the 4-axis machine as,

$$T^* = \begin{bmatrix} 1 & -\alpha^* & \beta^* & \Delta x^* \\ \alpha^* & 1 & -\gamma^* & \Delta y^* \\ -\beta^* & \gamma^* & 1 & \Delta z^* \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -0.0267 & -3.12 \\ 0 & 1 & 0 & -3.09 \\ 0.0267 & 0 & 1 & -7.19 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{5.24}$$

where α^*, β^* and γ^* are angular offsets in radian, α^* and γ^* are restricted to be zeros since the table can only be rotated about Y-axis and Δx^*, Δy^* and Δz^* are linear offsets in mm.

As can be seen in Equation (5.24), T^* suggests a rotation of the casting by -0.0267 radian (-1.298 degrees) along Y-axis, which is close to the angular offset, generated by the spacers placed between spacer Z1, Z2 and the casting. The point-set, which represents the top of the flange, is tilted, but the effect of the rigid body transformation rotates the point-set back to the horizontal position, as shown in Figure 5.16 and Figure 5.17. The whole point-set is translated by -3.12 mm along X-direction and -3.09 mm along Y-direction, respectively. These translated distances are close to 3.00 mm, the thickness of the spacers placed between the casting part and the Y-spacers.

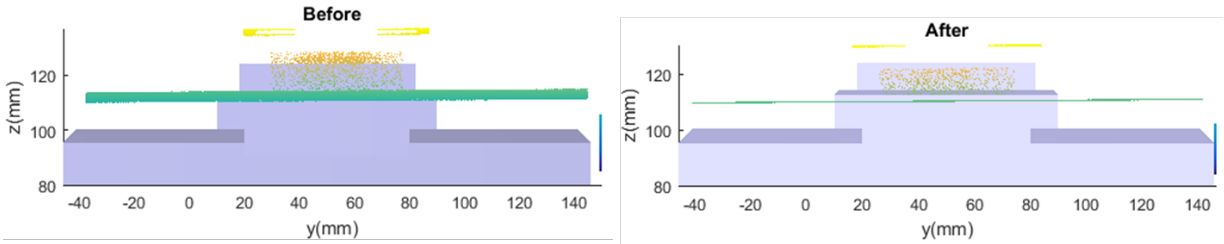

Figure 5.16: Front view of the seven point-sets (a) before T^* is applied; (b) after T^* is applied

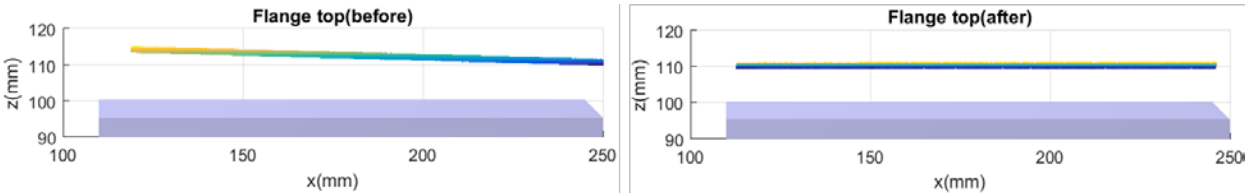

Figure 5.17: Side view of the point-set that represents flange top (a) before T^* is applied; (b) after T^* is applied

Similarly, the point-set for the top face of the neck must lie above the cutting line of the top face, which is already satisfied. Figure 5.18 shows how T^* compensates for the angular error about Y-axis and linear error along Z-axis to make the point-set approximately sit on the horizontal plane, z=130 mm.

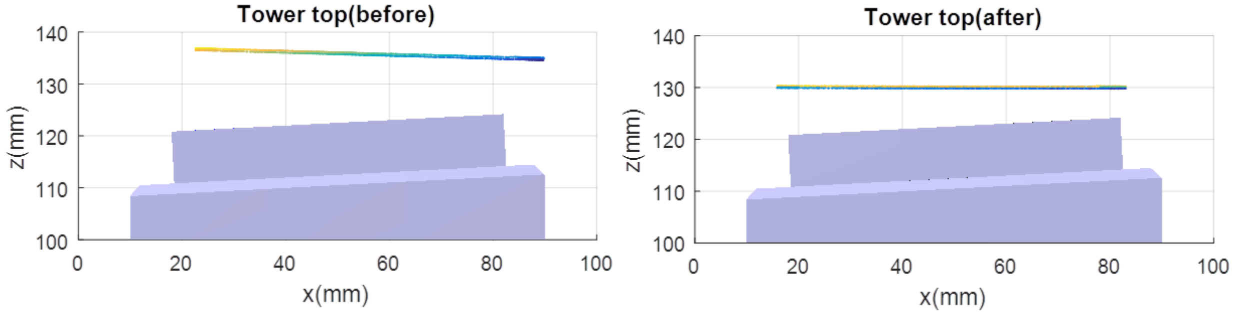

Figure 5.18: Side view of the point-set that represents tower top (a) before T^* is applied; (b) after T^* is applied

Also, four virtual gages, shown as four vertical lines in Figure 5.19 are deployed to check the two walls' thickness. By applying T^* to compensate the rotary errors and translational error along X-axis, the point-sets moved to the new positions without touching the virtual gages. It shows that the requirements on thickness can be satisfied.

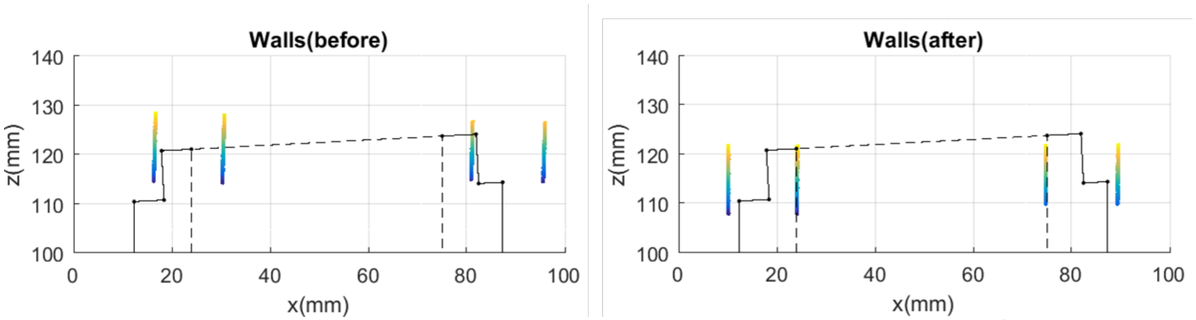

Figure 5.19: Side view of the point-sets that represents four walls (a) before T^* is applied; (b) after T^* is applied

To check the validity of the virtual gage analysis, the machining paths in the G-code was modified according to the transformation of the machining coordinate frame using T^*. The error map of the finished part was built by comparing the scan data of the finished part with the nominal CAD model. The magnitude of error vector is visualized using color code and shown in Figure 5.20. All machined surfaces on flange and tower are shown as green (error magnitude is between -0.4 and 0.4 mm). Since all locations of machined faces are within tolerance of 1 mm, the machined part is conforming.

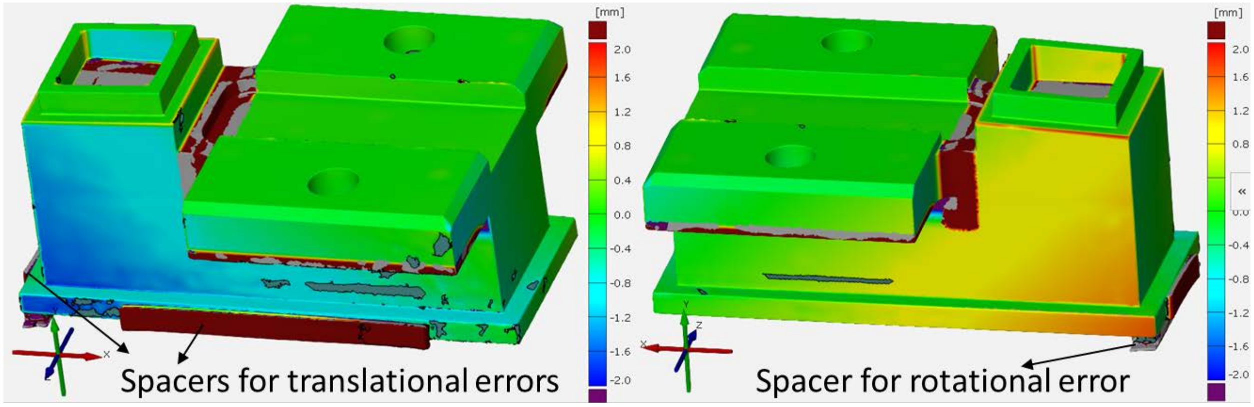

Figure 5.20: Error map of the finished part (compared with nominal CAD model)

5.5 EXAMPLE: CASTING WITH UNSATISFIABLE MATERIAL CONDITIONS

The purpose of this example is to check if the virtual gage analysis developed in Section 5.2.3 could identify the defective casting where the material conditions specified are not met for some features. An industrial part of approximate size of 1.5 m x 1.5 m x 1.5 m is scanned. The part has many features including steps, slots, pockets and holes, etc. However, the example part has two specific features, represented as type-A (pocket with slot) and type-B (step) that are shown in Figure 5.21 and Figure 5.22, respectively.

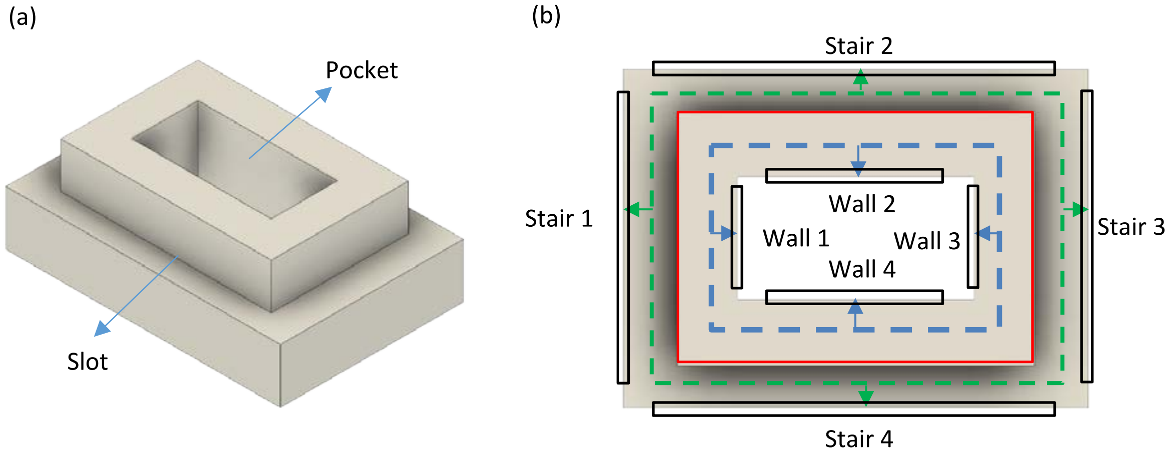

Figure 5.21: Type-A feature: (a) CAD model; (b) top view, eight sampling volumes and eight virtual gages

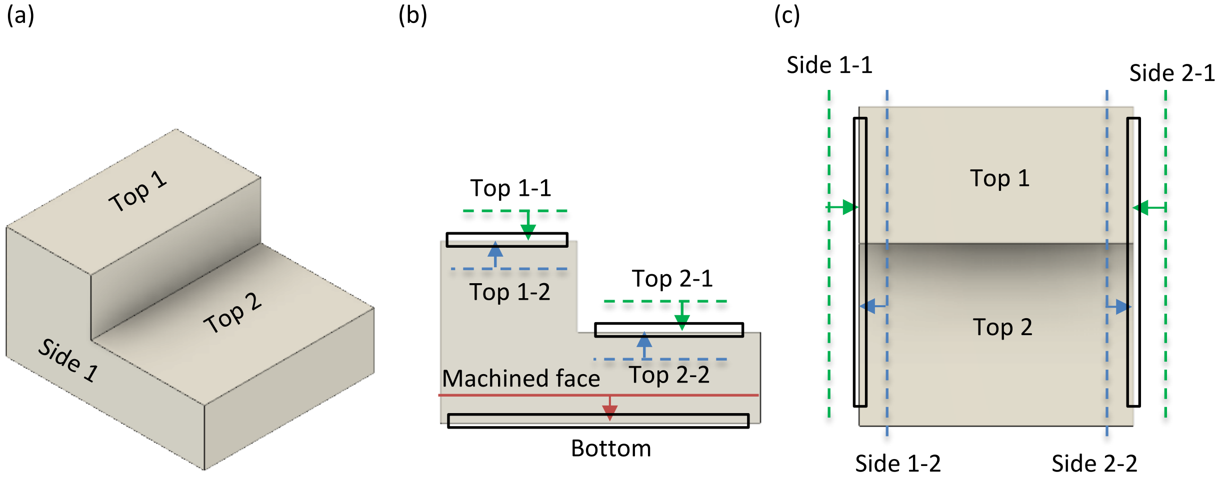

Figure 5.22: Type-B feature: (a) CAD model; (b) side view, three sampling volumes and five gages; (c) top view, two sampling volumes and four gages

As can be seen in Figure 5.21 (a), the type-A feature is a pocket with slot. Figure 5.21 (b) shows the top view of type-A feature and eight sampling volumes and virtual gages, represented by the black boxes and dashed lines, respectively. Eight point-sets are extracted to represent eight faces (Wall 1~4 and Stair 1~4). The blue gages are used to check four wall thicknesses, and the green gages examine four depths of steps. All eight gages are built by offsetting the reference surfaces, marked as red lines, and the offsets represent the minimum acceptable thicknesses and depths. Note the arrows in Figure 5.21(b) and Figure 5.22(b) and (c) point the side of feasible half-space. If the entire point-set is within the feasible region, the feature has positive allowance and has enough material for machining.

Type-B feature is a step shape feature, which is schematically depicted in Figure 5.22(a). Three sampling volumes (black boxes) as shown in Figure 5.22(b) are applied to extract point-sets representing two top faces and the bottom face. The ideal machined face (the red line in Figure 5.22(b)) is defined as a virtual gage to check if the bottom face has enough material to be machined properly. To examine if the thicknesses are within a plus/minus tolerance, two pairs of virtual gages (Top 1-1, 1-2 and Top 2-1, 2-2) are defined and shown using the side view of type-B feature (see the dashed lines in Figure 5.22(b)). These four gages are all defined by offsetting the nominal top faces, and the distances between blue/green dashed lines to red line are the minimum/maximum acceptable thicknesses. Similarly, two sampled point-sets extracted by the two sampling volumes as shown in Figure 5.22(c) are checked against four virtual gages, which are defined to control the width of type-B feature. Distances between blue/green dashed lines represent minimum/maximum acceptable widths.

The point cloud data is collected with a blue light scanner by scanning four main features. 636368 points are collected. A similar 2-step procedure as mentioned in Section 5.4 and Figure 5.10 is used to execute the virtual gage analysis with slack variables. As shown in Figure 5.21 (b), a type-A feature is represented by eight point-sets, which are checked by eight virtual gages. Figure

5.22(b) and (c) show type-B feature has five sampled point-sets, which are checked by nine gages. The part has three type-A features and one type-B feature. Therefore, 29 point-sets are extracted, and 33 virtual gages are defined in total. The linear programming problem with slack variables is formulated as,

$$\min_{T\in\Omega}\sum_{i=1}^{33} q_i + s_i + u \ \ni -u \le -s_i \le [a_i\ b_i\ c_i\ d_i]T\begin{bmatrix}\vec{r}_{i,j}\\ 1\end{bmatrix} \le q_i\ , i = 1\text{~}33, \tag{5.25}$$

The linear constraints are given in Equation (5.12). T the optimal HTM using 4-axis machine is obtained by solving Equation (5.25),

$$T^* = \begin{bmatrix} 1 & -\alpha^* & \beta^* & \Delta x^* \\ \alpha^* & 1 & -\gamma^* & \Delta y^* \\ -\beta^* & \gamma^* & 1 & \Delta z^* \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -0.0037 & -2.25 \\ 0 & 1 & 0 & -8.11 \\ 0.0037 & 0 & 1 & -5.87 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{5.26}$$

where α^*, β^* and γ^* are angular offsets in radian, α^* and γ^* are restricted to be zeros since the table can only be rotated about Y-axis and Δx^*, Δy^* and Δz^* are linear offsets in mm.

The machining allowances with and without compensation on three Type-A features are shown in Table 5.3. If the machining coordinate system is not adjusted, the finished casting would have four unmet material conditions (not enough material for machining) on three type-A features with negative allowances in depth, which are marked as bold in Table 5.3. With the compensation of T^*, all 24 material conditions of three Type-A features are satisfied.

Table 5.3: Machining allowances of three Type-A features (mm) before and after compensation

No.	Gage	Depth (before)	Depth (after)	Gage	Thickness (before)	Thickness (after)
1	Stair 1	**-8.2**	0	Wall 1	15.0	7.0
	Stair 2	7.1	0	Wall 2	6.2	13.0
	Stair 3	14.0	5.6	Wall 3	4.3	13.0
	Stair 4	1.8	8.9	Wall 4	8.2	1.1
2	Stair 1	**-6.8**	1.0	Wall 1	18.0	10.0
	Stair 2	4.4	3.9	Wall 2	3.4	3.8
	Stair 3	**-0.4**	0	Wall 3	13.0	19.0
	Stair 4	6.6	0	Wall 4	6.2	5.8
3	Stair 1	5.8	12.0	Wall 1	22.0	16.0
	Stair 2	4.1	0.2	Wall 2	6.7	10.0
	Stair 3	8.6	0.6	Wall 3	12.0	20.0
	Stair 4	**-1.7**	2	Wall 4	9.7	5.9

For Type-B feature, as can be seen in Table 5.4 and Figure 5.23(a), if the compensation HTM, is not applied, the Type-B feature has only one unmet material condition on side 1-1 but is placed off-center. With the compensation of T^*, the Type-B feature is aligned by center. Two slack variables in the analysis are found to be positive, and the two corresponding material conditions of maximum acceptable width cannot be satisfied as shown in Figure 5.23(b), which schematically shows the unsatisfiable material conditions on sides 1-1 and 2-1. This also shows the width of the type-B feature is larger than the maximum acceptable width by 5 mm (sum of two negative allowances in the third column of Table 5.4), which is further verified by direct measurement.

Table 5.4: Material allowances of Type-B feature (mm)

Gage	Thickness/width(before)	Thickness/width(after)
Top 1-1	4.4	2.4
Top 1-2	3.7	6.6
Top 2-1	4.5	2.3
Top 2-2	3.9	6.6
Bottom	2.5	0.4
Side 1-1	**-7.9**	**-2.4**
Side 1-2	16.9	11.3
Side 2-1	2.6	**-2.7**
Side 2-2	6.1	11.8

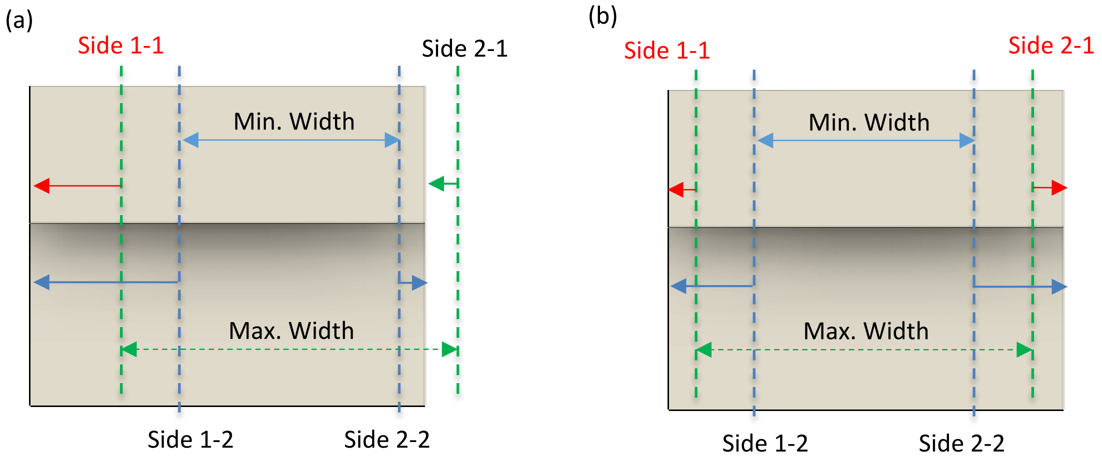

Figure 5.23: Type-B feature with unsatisfiable material conditions on maximum width: (a) before compensation; (b) after compensation

5.6 SUMMARY

In this chapter, the virtual gage analysis is proposed to determine the acceptability of a raw casting for the machining process. The concept of virtual gage is proposed by a parameterized plane and its half space, which represents a feasible region for the point-cloud data. The analysis seeks to displace the point-cloud using a single HTM sot that every defined virtual gage can be satisfied simultaneously. If such an HTM exists, the conformity of the part can be guaranteed. The analysis is also extended by introducing slack variables to deal with the part without enough material. Even the part cannot be properly machined, the HTM given by the analysis can still improve the conformity.

CHAPTER 6. TOLERANCE VERIFICATION OF CYLINDRICAL SURFACES

In Chapter 5, the virtual gage analysis is introduced to virtually check the material conditions of multiple planner surfaces. The concept of point-set based metrology is extended to cylindrical surfaces in this chapter. However, the difficulty of examining cylindrical surfaces is higher because a cylindrical surface has more degrees-of-freedom (five for a cylindrical surface and three for a planar surface). More importantly, the feasible space of plane fitting is always a convex space, which can be used to reduce the size of constraints and thus reduce the complexity of the problem. However, not every type of cylinder fitting problem has convex feasible space. For example, the fitting of maximum possible inscribed cylinder has non-convex feasible space, which makes the complexity of problem grows exponentially with the size by using traditional optimization solver. Therefore, different strategies are required.

Section 6.1 introduces the problem. In Section 6.2, the 2-D circular fitting problems are discussed, and the corresponding computational geometry-based approaches are developed. The projection model of 3-D point-set is built in Section 6.3, followed by example problems using available data sets in the literature demonstrated in Section 6.4. Section 6.5 summarizes this chapter.

6.1 INTRODUCTION

Modern metrology techniques make the measurements of surface profile efficiently with introduction of new measuring equipment such as laser scanner, which allows metrologist to get accurate and dense measurement data set. As the accuracy of measuring machine improves, requirements on manufacturing tolerance become more rigorous. However, the paradigm of workpiece metrology in the industry remains unchanged for decades. Without a good algorithm, metrologists are not able to process a larger data sets specifically obtained for the cylindrical surface, which may lead to overestimation of the tolerance, rejecting the acceptable parts and increasing the cost. Thus, a quick and accurate algorithm that judge the conformity of cylindrical surface is critically needed.

According to ASME Y14.5 [12], common tolerance specifications of cylindrical surfaces include minimum/maximum possible cylinder radii and the cylindricity error. These specifications are difficult to be measured or evaluated directly using CMMs since they are controlled by a three-dimensional data set. However, minimum and maximum possible cylinder radii can be estimated by the radii of maximum inscribed cylinder (MIC) and minimum circumscribed cylinder (MCC), and the cylindricity can be modeled using minimum zone cylinder (MZC) using optimization algorithms. However, the optimization problems cannot be solved easily due to the nonlinearities caused by the rotation and their non-differentiable target functions caused by discrete point-set.

In this chapter, a simplified approach for verifying cylindrical surface's tolerance specifications is proposed. Unlike the reported works, which directly use intelligent searching algorithms to find all five parameters of the best-fit cylinder (including two parameters representing the orientation of cylinder axis and three parameters for linear offset of cylinder axis), the proposed methodology only searches for two parameters that control the orientation of the cylinder. This is done by casting projection of the 3-D point-set along different directions to get different 2-D projected point-sets and their corresponding 2-D tolerance specifications, which are computed by computational geometry-based approaches. After the 2-D model is built by 2-D circular fitting problems, particle swarm optimization (PSO) is applied to find the cylinder axis's orientations (specified by azimuthal and polar angles) that optimize the corresponding 2-D tolerance specifications. By reducing the number of optimization parameters, the efficiency and accuracy of the tolerance verification procedure can thus be improved.

6.2 2-D CIRCULAR FITTING PROBLEMS

These tolerance specifications of cylindrical surface are usually difficult to be measured or evaluated directly using CMMs since they are controlled by a three-dimensional data set. For simplicity and efficiency, only a portion of the cylinder is measured, and a two-dimensional data set is collected around a circle and used to represent the entire cylinder. This method greatly simplifies the problems by reducing the dimension. The tolerance specifications in 3-D can be approximated using 2-D data set and its 2-D specification, i.e. minimum/maximum possible radii and the roundness error of the 2-D point-set. The verifications of these three 2-D tolerance specifications can be done in numerical or computational geometry-based approaches, which are explained in the following sections.

6.2.1 MAXIMUM RADIUS AND MINIMUM CIRCUMSCRIBED CIRCLE

MC is defined by the smallest possible circle that can be fitted around the roundness profile. Radius of MC represents the maximum possible radius of a circular profile. A min-max optimization can be used to define MC of a point-set,

$$\min_{x,y} r_{MC} \ni \ r_{MC}^2 \geq (x - x_i)^2 + (y - y_i)^2 \ \forall i = 1 \dots n, \tag{6.1}$$

where r_{MC} and $[x, y]$ are the radius and center of MC and the coordinate of i$^{\text{th}}$ point in the point-set is given by $[x_i \ y_i]$.

Equation (6.1) is a quadratically constrained quadratic program problem (QCQP), which generally is an NP hard problem [101]. However, as a special case of QCQP, Equation (6.1) can be simplified and solved using interior point method since the feasible region is a circle, which is a convex region. The two-dimensional optimization can be also solved by optimization solvers numerically.

The geometrical-based algorithm developed by Welzl [86] is able to solve MC fitting problem the problem. The algorithm starts with a MC candidate defined by two or three points in the point-set. If a point in the point-set is inside of the circle, the point is ignored in the following iteration as shown in Figure 6.1(a). Figure 6.1(b) shows the case that the point, which is outside the current circle is used to form the new circle. When a new point is considered, the circle becomes larger and circumscribes more points in the point-set. The iteration stops when all points in the point-set is either ignored or used to form the circle. The algorithm finds the control points that determine MC and provides the solution for Equation (6.1), and its complexity is proportional to the size of the point-set.

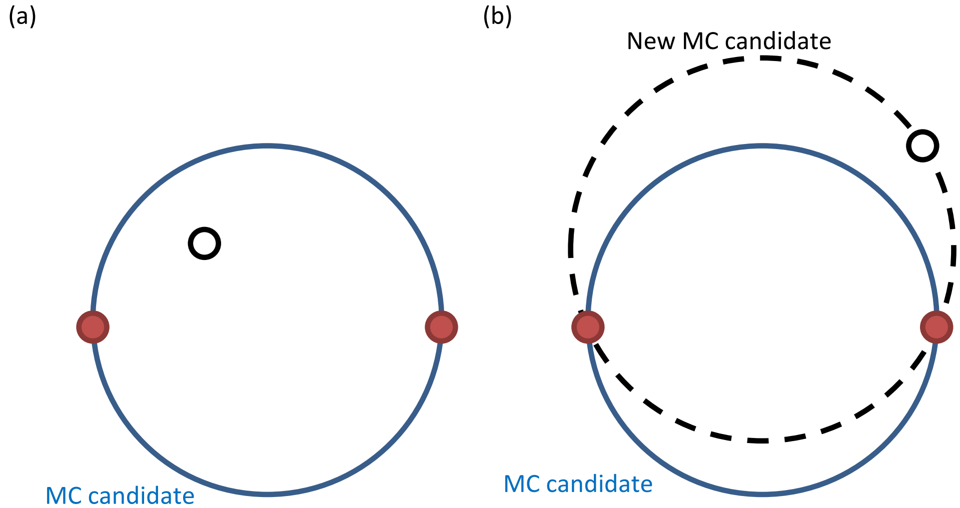

Figure 6.1: Welzl's algorithm on finding MC: (a) Case 1: remove a surrounded point; (b) Case 2: fit a larger circle using the point out of current circle

Sorting by the number of control points, there are four outcomes of finding MC for a given finite point-set, P. The center and radius of MC, and the i$^{\text{th}}$ point in P are denoted by c_{MC}, r_{MC} and $\vec{p}_i$, respectively.

1. Number of control points of MC is 1.

It is a special case. Either every point in P coincides or there is only one point in P. In this case, $c_{MC} = \vec{p}_1, r_{MC} = 0$.

2. Number of control points of MC is 2.

Two control points, the ath and bth points in P define the diameter of MC. If there are more than three points in P, the triangle determined by the control points and any other point in P is an obtuse triangle as shown in Figure 6.2(a). In this case, $c_{MC} = (\vec{p}_a + \vec{p}_a)/2, r_{MC} = \|\vec{p}_a - \vec{p}_b\|_2$.

3. Number of control points of MC is 3.

The three control points (ath, bth and cth points in P) form an acute triangle and its circumscribed circle is MC of P. Figure 6.2(b) shows an example that three points determine the point-set's MC.

4. Number of control points of MC is 4 or larger than 4.

As can be seen in Figure 6.2(c), it is a special case like case 3 that all four (or more) control points are concyclic.

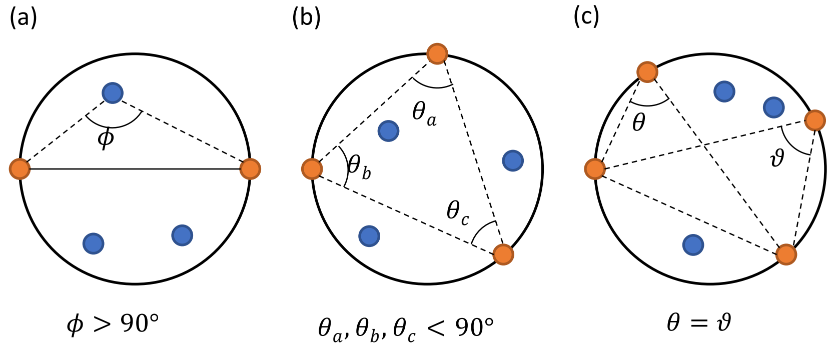

Figure 6.2: Circumscribed circle determined by: (a) 2 points; (b) 3 points; (c) 4 or more points

6.2.2 MINIMUM RADIUS AND MAXIMUM INSCRIBED CIRCLE

According to ISO 12181-1:2011 standard [102], MI is defined by the largest possible circle that can be fitted within the roundness profile. For a discrete point-set data, the roundness profile is not a continuous boundary and cannot be used to restrict the location of MI. Therefore, the roundness profile is approximated by MC, which surrounds the point-set. This implies that MI must be surrounded by MC. For example, if all points in the point-set are concyclic and form a perfect circle, MI and MC of the point-set are identical. MI of a finite point-set can thus be formulated using the following max-min optimization [20], [77]:

$$\max_{x,y} r_{MI} \ni r_{MI} + |\vec{c}_{MI} - \vec{c}_{MC}| \leq r_{MC}, r_{MI}^2 \leq (x - x_i)^2 + (y - y_i)^2 \; \forall i = 1 \ldots n, \tag{6.2}$$

where r_{MC} and $\vec{c}_{MC} = [x_{MC} \; y_{MC}]$ are the radius and center of MC, r_{MI} and $\vec{c}_{MI} = [x \; y]$ are the radius and center of MI and the coordinate of ith point in the point-set is given by $[x_i, y_i]$.

Equation (6.2) also forms a quadratically constrained quadratic program problem like Equation (6.1). However, the feasible region formed by the constraints is a ring-shape area, which is non-convex, and the problem is hence NP hard [101]. In the other word, without relaxation of constraints, the complexity and computation time of the problem grow exponentially with the size of problem. Therefore, using traditional optimization solver to solve Equation (6.2) might be inefficient for point-set with larger size. Thus, intelligent searching algorithms such as using particle swarm optimization (PSO) can be used to solve Equation (6.2) [103]. A generalized problem form that intelligent searching algorithms deal with is formulated as,

$$\max_{v} f(v) \; \ni l_i \leq v_i \leq u_i \; \forall i, \tag{6.3}$$

where $v \in \mathbb{R}^n$ is the vector with n variables, $f(v)$ is the objective function, l_i and u_i are lower and upper bounds for the ith variable v_i.

Equation (6.2) is rewritten without the constraint on MI's location and size so that intelligent searching algorithms can be applied,

$$\max_{x,y} \min_{i} [(x - x_i)^2 + (y - y_i)^2] \; \forall i \; \ni l_x \leq x \leq u_x, l_y \leq y \leq u_y, \tag{6.4}$$

where l_x and u_x are the lower and upper bound for x and l_y and u_y are the lower and upper bound for y.

However, the constraint on MI's location and size given in Equation (6.2) cannot be fully considered in Equation (6.4). The optimal solution might be inaccurate. Alternatively, the nearest Voronoi diagram, $NV(P)$ for point-set P can also be used to find all inscribed circles and MI. $NV(P)$ is the intersection of all the nearest Voronoi convex sets. The nearest Voronoi convex set associated with the i^{th} point in P is the set of all the points that are closer to the i^{th} point in P than to any other point. The nearest Voronoi diagram is composed of Voronoi edges and has two important properties [87], which are shown schematically in Figure 6.3:

1. The vertices of $NV(P)$, denoted by $V_N(P)$ are all points equidistant from at least three points in P and farther to all the others points in P. Every point in $V_N(P)$ can be the center of a circle that passes through at least three points in P without circumscribing any other points in P.
2. A set of points equidistant from two points in P and farther to all the others in P depict a nearest Voronoi edge, and the set of all nearest Voronoi edges are denoted by $E_N(P)$. Every edge is a part of perpendicular bisector of two points in P. Every point on $E_N(P)$ can be the center of a circle that passes through exact two points in P without circumscribing any other points in P.

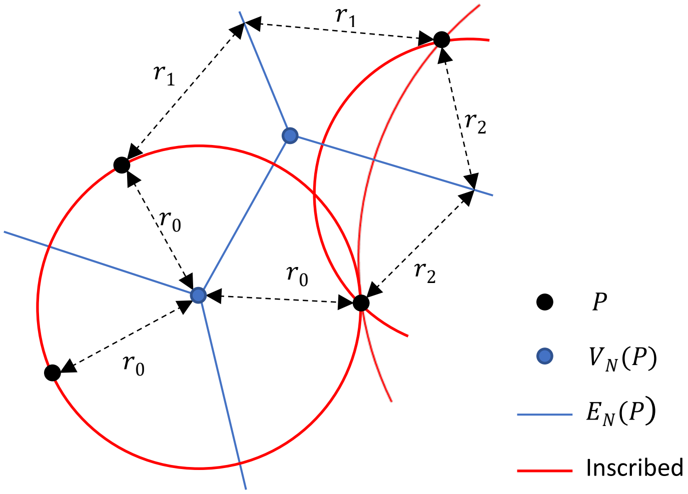

Figure 6.3: Schematic of nearest Voronoi diagram and its properties

By taking advantage of the first property of nearest Voronoi diagram, one can examine every potential inscribed circle that passes through at least three points in P but does not circumscribe any other points. The following steps are used to find all inscribed circles defined by at least three points:

1. Find minimum circumscribed circle of P, as shown in the black circle in Figure 6.4(a).
2. Construct nearest Voronoi diagram, $NV(P)$ with edges, $E_N(P)$ shown as blue lines in Figure 6.4(a).
3. For each vertex of $NV(P)$ (each point in $V_N(P)$), shown as blue points in Figure 6.4(a),
    (a) Find the three points, which are equally distant to the vertex.
    (b) Define a corresponding inscribed circle using three points found.
    (c) As shown in Figure 6.4(a), check if the potential inscribed circle is inside of MC. If yes, add it into the candidate list.
4. Find the largest potential inscribed circle left from step 3.

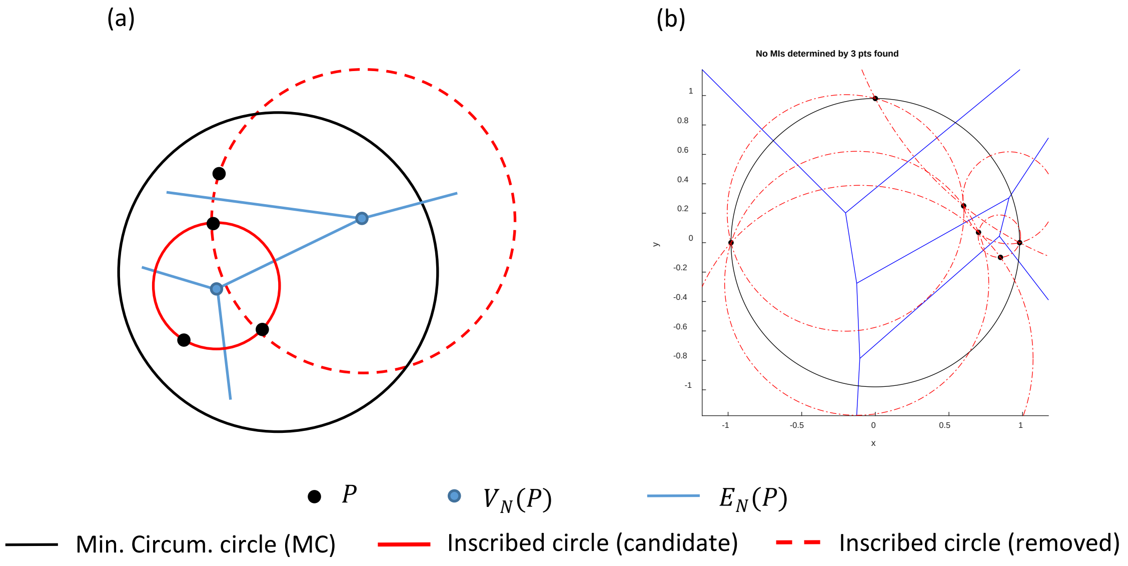

Figure 6.4: (a) Schematic of finding inscribed circles defined by 3 points; (b) every inscribed circle defined by 3 points is out of the minimum circumscribed circle

In general, using algorithms such as PSO to solve Equation (6.4) or constructing Voronoi diagram using at least three points in the point-set can find an inscribed circle for finding MI. However, it is possible that the found inscribed circle from these approaches is out of the roundness profile depicted by MC as shown in Figure 6.4(b). In general, this happens for point-set that hardly depicts a circular shape (e.g. the black points shown in Figure 6.4(b)). Hence, two additional situations (case I and II) are proposed and their solutions based on Voronoi diagram are developed.

Case I: MI is defined using 2 points in P and 1 tangent point on MC

Every point on the edge of Voronoi diagram can be the center of an inscribed circle. The inscribed circle passes through exactly two points in the point-set without circumscribing any other points. Since the inscribed circle should be as large as possible and within MC, the inscribed circle of the point-set must be also the internally tangent circles (ITC) of MC. The following steps explain how to find all circles that are not only ITCs of MC but also inscribed circles of the point-set:

1. Remove the points on MC and construct the new Voronoi diagram for the smaller point-set, P'.
2. Find all edges in $E_N(P')$.
3. For every edge found in step 2, it has two ends, V_1 and V_2 as shown in Figure 6.5(a).
    (a) Find A and B, the two points in P', which are closest to the edge.
    (b) Given A and B, solve two tangent points, Q_1 and Q_2 on MC so that two circles determined by ABQ_1 and ABQ_2 are two ITCs of MC.
    (c) Determine C_1 and C_2, the centers of two ITCs. Note that V_1, V_2, C_1 and C_2 are collinear.
    (d) Check if $\angle AC_iB > \angle AV_iB$ (i=1,2). If yes, add the ith ITC into the candidate list. Otherwise, remove the ith ITC because it circumscribes other points in P'. For example,

ITC 1 in Figure 6.5 (a) is a candidate because $\angle AC_1B > \angle AV_1B$. ITC 2 in Figure 6.5(a), on the other hand, will not be considered in the next step.

4. Find the largest circle in the candidate list found in step 3.

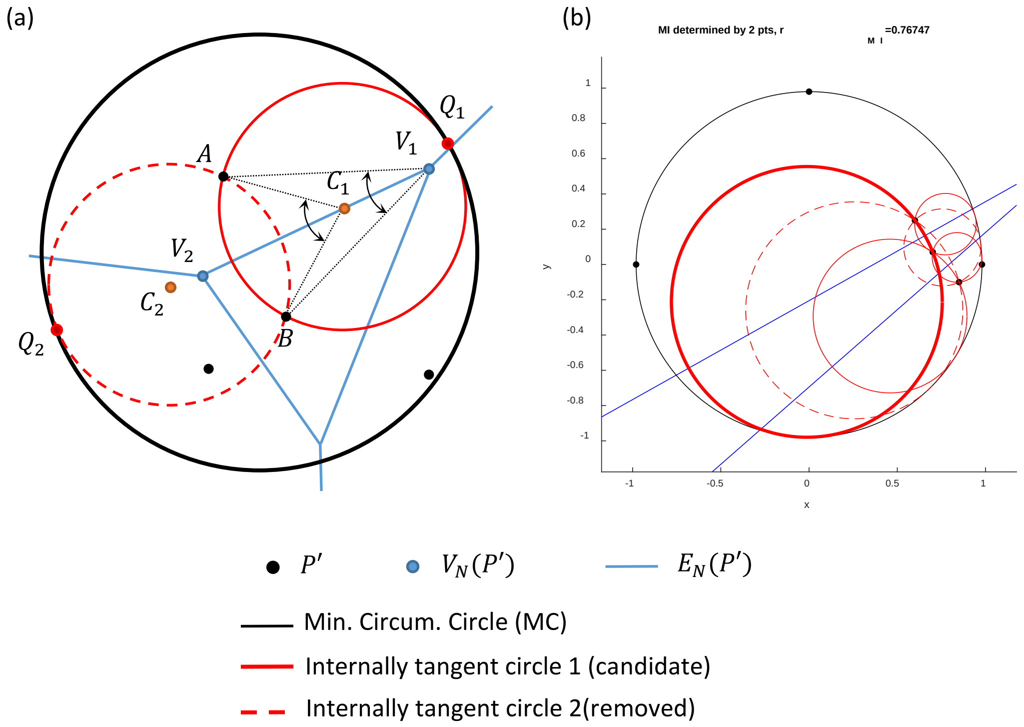

Figure 6.5: (a) Schematic of finding ITC using two points; (b) largest ITC in Case II

Case II: MI is defined by 1 point from point-set and 1 tangent point on MC.

A circle can also be determined using two points if these two points also define the diameter. To maximize the diameter defined in this way, the ITC can be defined using a point inside of MC and the farthest possible point. With the assistance of nearest Voronoi diagram, all ITC candidates can be found using the following steps:

1. Remove the points on MC and construct the new Voronoi diagram for the smaller point-set, P'.

2. For every point in P', do the following steps (details are shown in Figure 6.6(a)):

(a) Find the diameter of MC that passes through the point, A. The two ends of the diameter are N and F. Two ITCs can be found using diameters defined by $\overline{AN}$ and $\overline{AF}$.

(b) Check if $\overline{AN}$ and $\overline{AF}$ intersects with any of the Voronoi edges, $E_N(P')$. If yes, the defined ITC circumscribes other points in P' and hence is not a candidate. As can be seen in Figure 6.6(a), $\overline{AF}$ intersects one edge in $E_N(P')$ at T. The circle defined by $\overline{AF}$ will not be considered in step 3.

(c) If $\overline{AN}$ does not intersect with any Voronoi edges, add the ITC with diameter of $\overline{AN}$ to the list of candidates as shown in Figure 6.6(a). Similarly, check if the ITC defined by $\overline{AF}$ should be added.

3. Find the largest circle in the candidate list found in step 2.

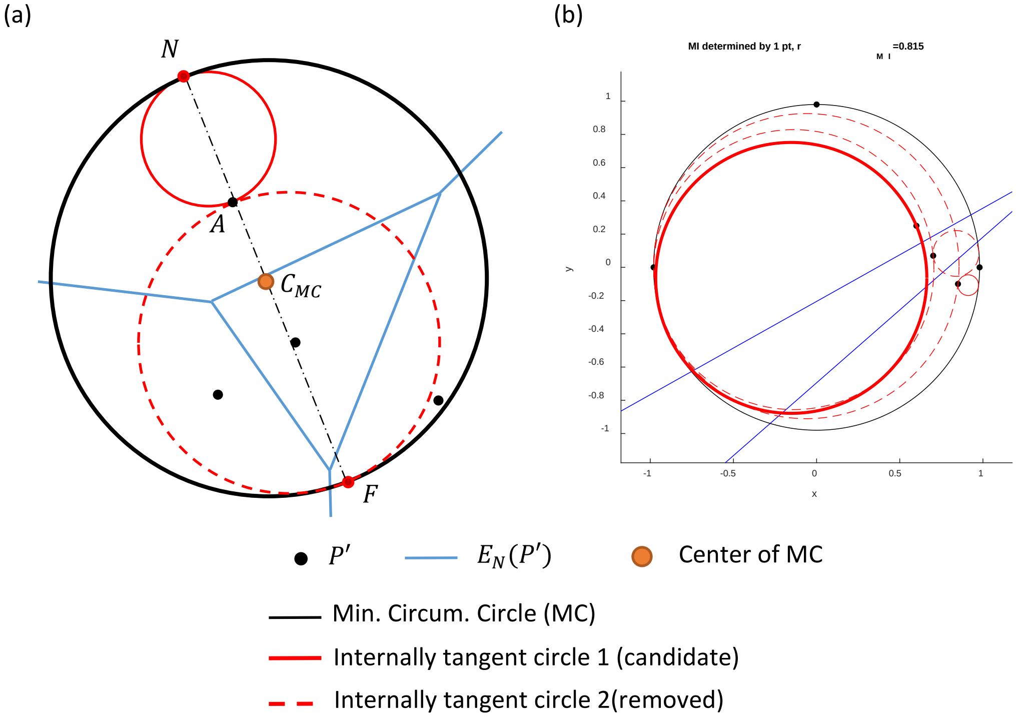

Figure 6.6: (a) Schematic of finding ITC using one point; (b) largest ITC in Case III

To find MI of a point-set, all cases including MI determined by at least three points should be considered. The largest inscribed circle given by one of the three cases is MI. For example, MI of the six points shown in Figure 6.4(b) cannot be defined using at least three points. Figure 6.5(b) shows the inscribed circle defined by Case I with radius of 0.768. The inscribed circle defined by Case II as shown in Figure 6.6(b) has radius of 0.815. Therefore, MI of the point-set has radius of 0.815.

6.2.3 ROUNDNESS AND MINIMUM ZONE CIRCLE

Roundness, also known as circularity is described by the tolerance zone bounded by two concentric circles [12], [104]. The minimum zone circle (MZ) fitting problem of a point-set, P with n points can be formulated with a min-max problem [20], which is similar to MI fitting problem as stated in Equation (6.4):

$$\min_{x,y} f_r(x,y) \tag{6.5}$$

where $f_r(x,y) = \max_i \sqrt{(x-x_i)^2 + (y-y_i)^2} - \min_j \sqrt{\left(x-x_j\right)^2 + \left(y-y_j\right)^2}$ $(i,j = 1\sim n)$ is the objective function of the roundness deviation controlled by x and y, the center position of the concentric circles and $[x_i \; y_i]$ is the i$^{\text{th}}$ point in P.

The roundness is given by $f(x^*, y^*)$ where x^* and y^* are the solution of Equation (6.5) as well as the positions of MZ. Equation (6.5) can also be numerically solved using intelligent searching algorithms such as PSO [105] and genetic algorithm (GA) [84]. However, a computational geometry method based on nearest and farthest Voronois diagrams [87], [88], [90] provides the global optimum of Equation (6.5) geometrically. Since a pair of concentric circles has four degrees of freedom, four or more control points within the point-set that uniquely determine the concentric circles are searched. According to Roy and Zhang's works [87], [88], one must individually find three pairs of concentric circles, whose outer and inner circles are defined by three and one points (case 3+1), two and two points (case 2+2), and one and three points (case

1+3). The pair with smallest amount of radial separation among three cases to be MZ. Roundness error can thus be obtained by radial difference of the concentric circles.

The roundness is given by $f_r(x^*, y^*)$, where x^* and y^* are the solution of Equation (6.5) as well as the center position of MZ. Equation (6.5) can also be numerically solved using intelligent searching algorithms such as PSO and genetic algorithm (GA). However, a computational geometry method based on nearest and farthest Voronois diagrams provides the global optimum of Equation (6.5) geometrically [87], [88], [90]. Since a pair of concentric circles has four degrees of freedom (two for center position and two for radii of two circles), four or more control points within the point-set that uniquely determine the concentric circles are searched. According to Roy and Zhang's works [87], [88], one must individually find three pairs of concentric circles, whose outer and inner circles are defined by three and one points (case 3+1), two and two points (case 2+2), and one and three points (case 1+3), respectively. The pair with smallest amount of radial separation among three cases to be MZ. Roundness error can thus be obtained by radial difference of the concentric circles. The flowchart of computational geometry-based algorithms for MC, MI and MZ fitting is shown in Figure 6.7.

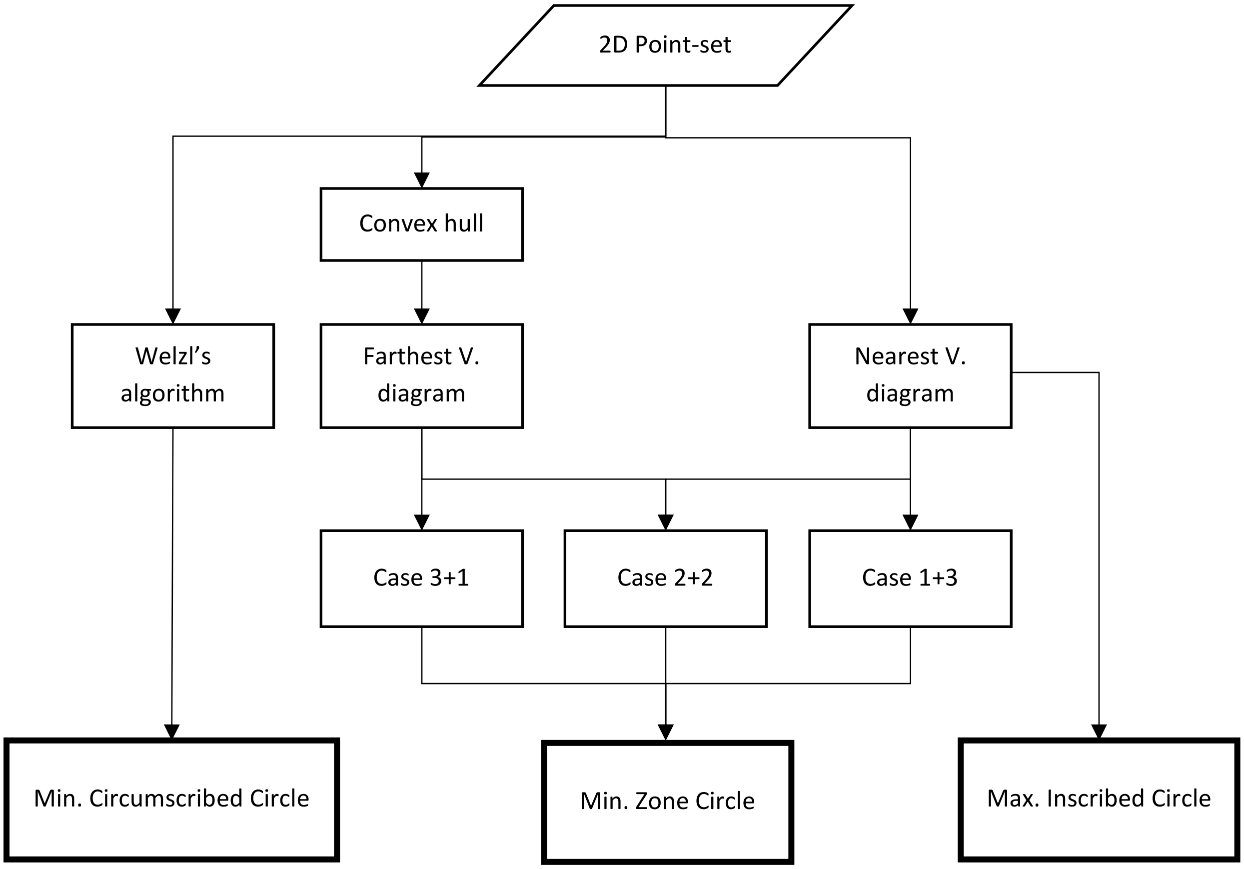

Figure 6.7: Computational geometry-based algorithms for fitting MC, MI and MZ [87]

In case 1+3, center of the concentric circles is determined by an inscribing circle of P. A potential inscribing circle must passes through at least 3 points in the point-set and does not circumscribe any other points in the point-set [88]. Hence, a potential inscribing circle must center on a vertex of the nearest Voronoi diagram. Equation (6.5) can be modified by searching the optimal concentric circles in case 1+3,

$$\min_{k_N} f_r(x_{k_N}, y_{k_N}), \tag{6.6}$$

where $f_r(x_{k_N}, y_{k_N}) = \max_i \sqrt{(x_{k_N} - x_i)^2 + (y_{k_N} - y_i)^2} - \min_j \sqrt{(x_{k_N} - x_j)^2 + (y_{k_N} - y_j)^2}$,

$\forall i, j = 1 \sim n$ and $[x_{k_N}\ y_{k_N}]$ is one of the vertices in $V_N(P)$.

Case 3+1, on the other hand, defines the center of the concentric circles by a circumscribing circle of P. As discussed in Section 6.2.1, MC of P might be defined by only two points in P. The potential circumscribing circle in Case 3+1, on the other hand, must be defined by 3 or more points [88]. In order to find the potential circumscribing circles, the farthest Voronoi diagram of P, $FV(P)$ is constructed. The center of a potential circumscribing circle is one of the vertices $FV(P)$ denoted by $V_F(P)$. Similarly, the formulation of case 3+1 is given by,

$$\min_{k_F} f_r(x_{k_F}, y_{k_F}), \tag{6.7}$$

where $f_{\mathrm{r}}(x_{k_F}, y_{k_F}) = \max_i \sqrt{(x_{k_F} - x_i)^2 + (y_{k_F} - y_i)^2} - \min_j \sqrt{(x_{k_F} - x_j)^2 + (y_{k_F} - y_j)^2}$,

$\forall i, j = 1 \sim n$ and $[x_{k_F}\ y_{k_F}]$ is one of the vertices in $V_F(P)$.

It must be noted that the outer and inner circle found in case 3+1 and 1+3 might be different circles from MC and MI of P.

In case 2+2, both concentric circles pass through two or more points. The inner circle cannot circumscribe any points, while the outer one should circumscribe the entire point-set. The center of such a pair of concentric circles is determined by the intersections of nearest and farthest Voronoi edges, $E_N(P)$ and $E_F(P)$ [87], [88]. The problem can be written as,

$$\min_{k_I} f_r(x_{k_I}, y_{k_I}), \tag{6.8}$$

where $f_2(x_{k_I}, y_{k_I}) = \max_i \sqrt{(x_{k_I} - x_i)^2 + (y_{k_I} - y_i)^2} - \min_j \sqrt{(x_{k_I} - x_j)^2 + (y_{k_I} - y_j)^2}$,

$\forall i, j = 1 \sim n$ and $[x_{k_I}\ y_{k_I}]$ is one of the intersections of $E_N(P)$ and $E_F(P)$.

Example problem

A randomly generated point-set is use to test the PSO-based and Voronoi diagram-based methods. PSO algorithm is used to solve Equation (6.5) with particle size of 20 and maximum number of iterations of 2000. For the Voronoi-based method, three cases of radial deviation are individually computed using nearest and farthest Voronoi diagrams. Three pairs of concentric circles are obtained and shown in Fig. 7. The roundness is found to be 1.055088 determined by case 2+2. As seen in Table 6.1, two methods have consistent results on both the center of MZ and the roundness.

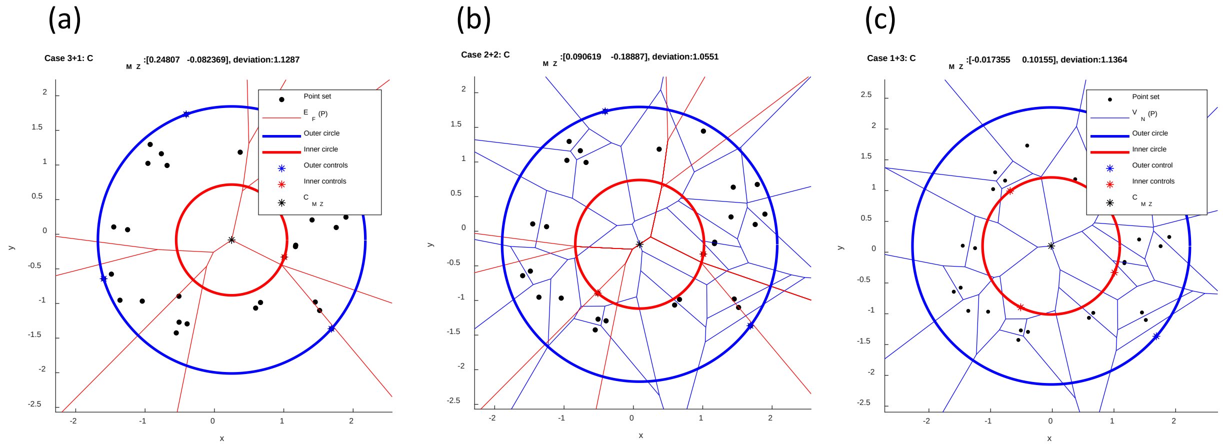

Figure 6.8: Three cases of finding minimum zone circle with smallest radial separation: (a) case 3+1; (b) case 2+2; (c) case 1+3

Table 6.1: Roundness and MZ fitting of a point-set using two proposed methods

Method	X_c	Y_c	Roundness
Voronoi (case 2+2)	0.0906190	-0.1888661	1.0550880
PSO	0.0906191	-0.1888664	1.0550879

6.3 3-D CYLINDRICAL DATA FITTING PROBLEMS

A cylinder has five degrees-of-freedom including two parameters for orientation of cylinder axis vector, two parameters for the offset of the axis vector and one to specify the radius. The schematic of using Plucker coordinate [20] to fit a cylinder is shown in Figure 6.9. The optimization formula of fitting MCC and MZC are given by [20],

$$\min_{\vec{a},\vec{L}} r_{MCC}(\vec{p}_i,\vec{a},\vec{L}) = \min_{\vec{a},\vec{L}} \max_{i} \left|\left(\vec{p}_i - \vec{L}\right) \times \vec{a}\right|, i = 1 \sim n, \tag{6.9}$$

$$\min_{\vec{a},\vec{L}} f_{MZC}(\vec{p}_i,\vec{a},\vec{L}) = \min_{\vec{a},\vec{L}} \left[\max_{i}\left|\left(\vec{p}_i - \vec{L}\right) \times \vec{a}\right| - \min_{j}\left|\left(\vec{p}_j - \vec{L}\right) \times \vec{a}\right|\right], i = 1 \sim n, \tag{6.10}$$

where $\vec{a}$ is axis direction, $\vec{L}$ is locating position vector, $\vec{a}^T\vec{L} = 0$, $|\vec{a}| = 1$ is the orientation vector of cylinder axis, $\vec{L}$ is a point on cylinder axis, $\vec{p}_i$ is the ith point in P.

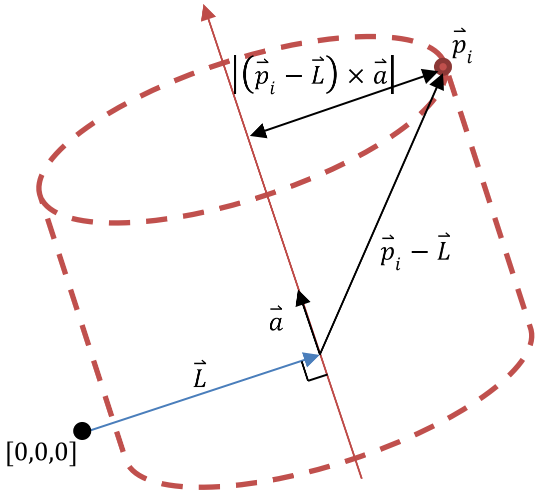

Figure 6.9: Define axis vector and the distance from points to axis using Plucker coordinate

MIC fitting problem of point-set is an unbounded maximizing problem like MI fitting problem. Therefore, two additional constraints to limit the axis location and orientation are required for the existence of solution.

$$\max_{\vec{a},\vec{L}} r_{MIC}(\vec{p}_i, \vec{a}, \vec{L}) = \max_{\vec{a},\vec{L}} \min_i \left|(\vec{p}_i - \vec{L}) \times \vec{a}\right|, i = 1\sim n \ \ni \vec{L} \in \Gamma_L, \tag{6.11}$$

where Γ_L limits the location of the axis within the cylindrical profile.

As can be seen in Equations (6.9)~(6.11), the three problems seek to optimize variables in $\vec{a}$ and $\vec{L}$ that specify position and orientation of the axes [20], [95], [99]. The number of variables in $\vec{a}$ can be reduced from 3 to 2 by (a) adjusting the axis coordinate by aligning it with Z-axis and (b) assuming the angular displacements of axis are sufficiently small. The simplified model of $\vec{a}$ is given by $\vec{a} = [\Delta\beta \quad -\Delta\gamma \quad 1\,]^T$, where $\Delta\beta$ and $\Delta\gamma$ are small angular displacements about Y and X axes. Similarly, $\vec{L}$ has 3 variables but and can be simplified by coordinate adjustment. Since the axis, $\vec{a}$ is assumed to be aligned with Z-axis, $\vec{L} = [L_x \quad L_y \quad 0\,]^T$ is given to reduce the number of variables. With small variable assumptions, Equations (6.9)~(6.11) are solvable sequentially (e.g., by updating $\Delta\beta$, $\Delta\gamma$, L_x and L_y step by step until the objective cannot be improved anymore) [20], [77]. Thus, intelligent searching algorithms such as PSO is applied to find global optimum for five parameters (including two parameters in axis direction $\vec{a} = [q_1\ q_2\ 1]^T$, and three parameters in locating positon vector ,$\vec{L} = [x_0\ y_0\ z_0]^T$ as shown in Figure 6.9) [99]. Cheraghi et al. [94] constructed a perturbation model for 3-D point-set's rotation and projection. Although the global optimum is not promised due to the nonlinearities of rotations, the projection model can be used to simplify the problem. The radii of MIC and MCC and cylindricity error of a 3-D point-set are formulated by maximum radius of MI, minimum radius of MC and minimum roundness error of the rotated and projected 3-D point-set, which can be solved using numerical and geometrical methods as discussed in Section 6.2. If the projection model is used, parameters in locating positon vector,$\vec{L}$ are redundant, and only two variables (two angles to specify an arbitrary rotation) are required to determine MIC, MCC and MZC as shown in Figure 6.10. After the projection model is built, PSO is used to find the optimal rotation (specified by two angles) that optimizes maximum radius of MI, minimum radius of MC and minimum roundness error of the rotated and projected 3-D point-set.

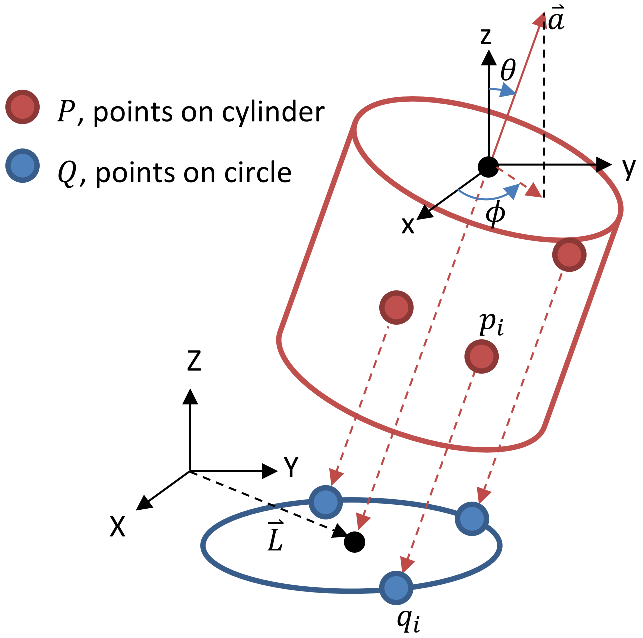

Figure 6.10 Project a 3-D point-set, P along $\vec{a}$ onto x-y plane to be Q

An arbitrary rotation applied to a point can be described by two consecutive rotations about Y and Z axes using polar and azimuthal angles. As shown in Figure 6.10, the projection of the rotated point onto x-y plane can be expressed as,

$$\vec{q}_i = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix} \vec{p}_i, \tag{6.12}$$

where $\vec{p}_i \in \mathbb{R}^3$ is the ith point in P, $\vec{q}_i \in \mathbb{R}^2$ is the ith point Q, the projected point-set of P, $\theta \in [0\ \pi/2]$ is the polar angle and $\phi \in [-\pi\ \pi]$ is the azimuthal angle.

As can be seen in Figure 6.10, the projection of rotated P is Q, a 2-D point-set whose MC, MI and MZ can be solved using the methods reviewed in Section 6.2 for given θ and ϕ. Equations (6.9)~(6.11) can be rewritten based on Equations (6.1), (6.4) and (6.5):

$$\min_{\theta,\phi} r_{MCC} = \min_{\theta,\phi} r_{MC} = \min_{\theta,\phi} \left[\min_{\vec{c}_{MC}} \max_{i} |\vec{c}_{MC} - \vec{q}_i| \right], \tag{6.13}$$

$$\min_{\theta,\phi} f_{MZC} = \min_{\theta,\phi} f_{MZ} = \min_{\theta,\phi} \left[\min_{\vec{c}_{MZ}} \left(\max_{i} |\vec{c}_{MZ} - \vec{q}_i| - \min_{j} |\vec{c}_{MZ} - \vec{q}_j| \right) \right], \tag{6.14}$$

$$\max_{\theta,\phi} r_{MIC} = \max_{\theta,\phi} r_{MI} = \max_{\theta,\phi}\left[\max_{\vec{c}_{MI}} \min_{i} |\vec{c}_{MI} - \vec{q}_i|\right] \ni \\ r_{MI} + |\vec{c}_{MI} - \vec{c}_{MC}| \leq r_{MC} \ \forall \theta, \phi \tag{6.15}$$

where $\vec{q}_i$ is the i^{th} point in Q, r_{MCC} and r_{MIC} are the radii of minimum circumscribed and maximum inscribed cylinder of P, f_{MZC} is the cylindricity of P, r_{MC} and r_{MI} are radii of MC and MI of Q and $\vec{c}_{MC}, \vec{c}_{MZ}, \vec{c}_{MI}$ are centers of MC, MZ, MI.

In this chapter, PSO is applied to solve θ and ϕ in Equation (6.13)~(6.15). In the main loop, 2-D tolerance specifications are computed using the computational geometry-based methods as shown in Figure 6.7. The flowchart of operation is given in Figure 6.11. It must be noted that one can also use PSO-based fitting in the main loop, since PSO is also capable of fitting circles of the projected point-set as demonstrated in Section 6.2.

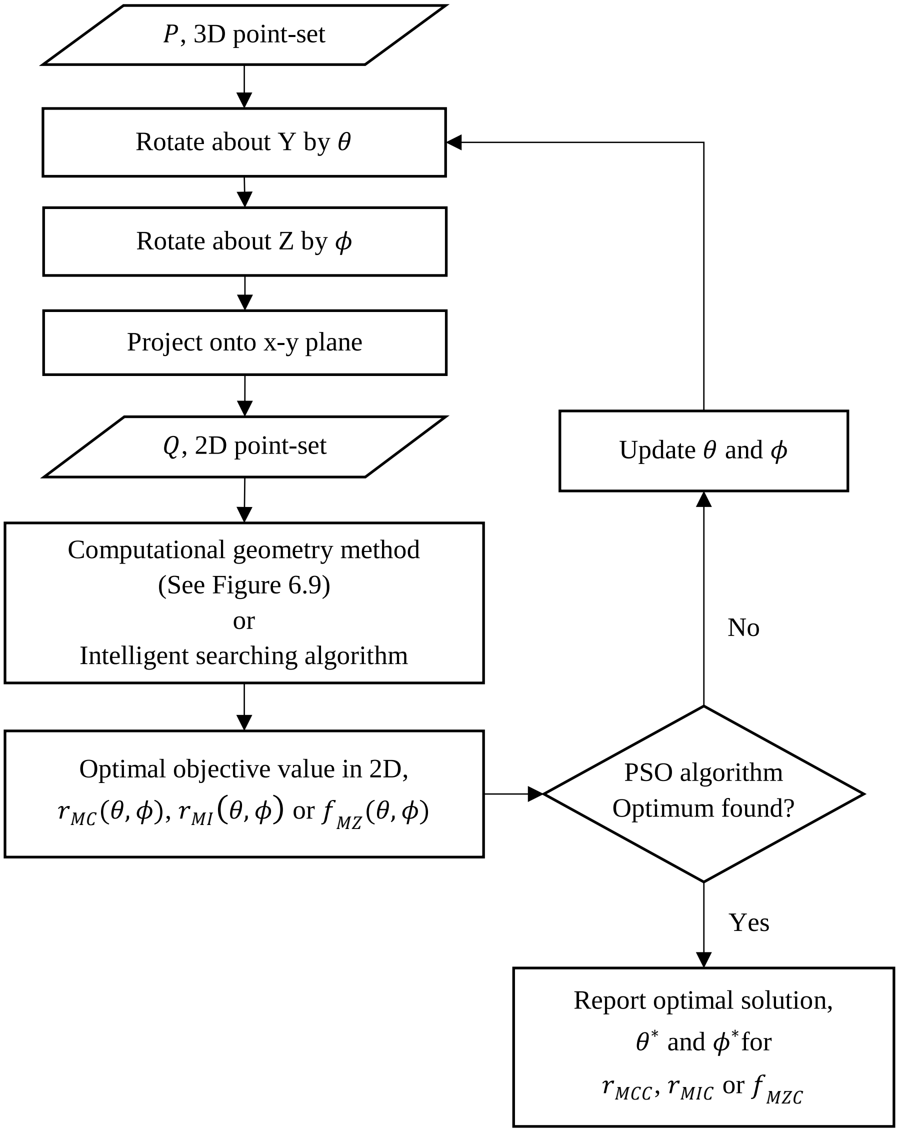

Figure 6.11: Flowchart of MCC, MIC and MZC fitting for a 3-D point-set

6.4 RESULTS

Four sets of data from literature are used for testing the proposed method [20], [95], [97]. Computational geometry methods are used in all iterations for all test results shown in this section. For the PSO solver, the particle size is selected as 20 and maximal number of iterations to be 2000. The searching regions of θ and ϕ are 0 to $\pi/2$ and $-\pi$ to π, which represent the half upper spherical surface. Table 6.2 to Table 6.5 show the results, and the used data sets are shown in

Appendices A, B, C and D, respectively. For each data set, MIC, MCC and MZC are fitted. The fitted cylinder axis direction specified by polar and azimuthal angles, the radii of MIC and MCC and the cylindricity are compared with the results reported in literature.

Table 6.2: Results of MIC, MCC and MZC fitting using data set 1 [20]

	Proposed approach			Literature		
	θ^*(rad)	ϕ^*(rad)	Radius/Cylindricity	θ^*(rad)	ϕ^*(rad)	Radius/Cylindricity
MIC	0.95531	0.78539	49.9956642	0.95531	0.78541	49.9956642
MCC	0.95539	0.78539	50.0041462	0.95539	0.78541	50.0041462
MZC	0.95534	0.78539	0.0094101	0.95540	0.78541	0.00941

Table 6.3: Results of MIC, MCC and MZC fitting using data set 2 [20]

	Proposed approach			Literature		
	θ^*(rad)	ϕ^*(rad)	Radius/Cylindricity	θ^*(rad)	ϕ^*(rad)	Radius/Cylindricity
MIC	5.9435E-03	-2.68556	59.9362261	6.0663E-03	-2.65050	59.9362147
MCC	7.0404E-03	-1.46939	60.0706331	1.29075E-02	-1.22423	60.0578814
MZC	2.9801E-03	-1.78017	0.1839574	2.9665E-03	-1.78017	0.18396

Table 6.4: Results of MIC, MCC and MZC fitting using data set 3 [95]

	Proposed approach			Literature		
	θ^*(rad)	ϕ^*(rad)	Radius/Cylindricity	θ^*(rad)	ϕ^*(rad)	Radius/Cylindricity
MIC	1.5627E-02	2.67553	12.0001738	1.5625E-02	2.67574	12.000174
MCC	1.5613E-02	2.67779	12.0015866	1.5613E-02	2.67780	12.001587
MZC	1.5632E-02	2.67489	0.0027883	1.5619E-02	2.67689	0.002788

Table 6.5: Results of MIC, MCC and MZC fitting using data set 4 [97]

	Proposed approach			Literature		
	θ^*(rad)	ϕ^*(rad)	Radius/Cylindricity	θ^*(rad)	ϕ^*(rad)	Radius/Cylindricity
MIC	2.5528E-05	1.34413	34.9529517	N/A	N/A	N/A
MCC	2.3936E-04	-0.90318	34.9823836	N/A	N/A	N/A
MZC	1.6417E-04	-1.62812	0.0318301	1.6397E-04	-1.63300	0.0319

Compared with the literature using sequential linear programming approach [20], the model developed in this paper shows similar results for data set 1 as can be seen in Table 6.2. However, for data set 2 given in Table 6.3, results of MCC fitting are slightly different from those reported in literature. The difference may be caused by the selection of the initial guess and the small displacement assumption. Table 6.4 shows the comparison between the methods presented in this paper with the genetic algorithms (GA) method [95], while two methods show high consistency on all MIC, MCC and MZC fitting performance. The Geometry Optimization Searching Algorithm (GOSA) [97] is also compared with the proposed method in Table 6.5, while the proposed method has similar estimation on cylindricity. The validity of the proposed methods on MIC, MCC and MZC fitting is thus demonstrated.

6.5 SUMMARY

In this chapter, the 2-D circular fitting problems are reviewed to verify tolerance specifications including maximum and minimum possible radii, roundness of a circular feature. Traditional fitting approaches including numerical and computational geometry-based methods define MI using at least three points in the point-set. However, the fitted inscribed circle may not be fitted within the roundness profile for some ill-shaped point-set. Hence, two more scenarios that define MI with one or two points and a tangent point on MC are considered. The corresponding solutions for these two scenarios are also developed based on Voronoi diagram.

The 3-D cylinder fitting problem for MCC, MIC and MZC can be modeled as problems of optimizing cylinder axis and solved using different algorithms. The reported approaches use four or five parameters to specify the cylinder axis’s orientation and translation, which are geometrically redundant. Only two orientation variables are required if the 3-D point-set is projected along the axis direction. The proposed approach fits MIC, MCC and MZC by finding the optimal cylinder axis direction to project the point-set so that the corresponding 2-D specifications (minimum radius, maximum radius and roundness) of the projected 3-D point-set are optimized (maximized, minimized and minimized, respectively).

CHAPTER 7. CONCLUSIONS AND FUTURE WORK

This chapter provides a summary of the work presented in this thesis. The overall objective of this dissertation is to characterize and compensate for CNC machine tool's machining inaccuracy. The proposed objective is met by machine tool quasi-static error modeling, point-set based workpiece metrology and GD&T verification. Specific conclusion drawn from the research work is given in Section 7.1. Recommendations for continued research are also presented in Section 7.2.

7.1 SUMMARY AND CONCLUSIONS

The main contributions of this thesis may be categorized into two basic areas, machine tool quasi-static error modeling and point-set based workpiece metrology.

7.1.1 MACHINE TOOL QUASI-STATIC ERROR MODELING

(1) A general modeling approach for quasi-static error for a 5-axis machine with a redundant axis is developed. This approach can be applied to model machines with different kinematic and static structures. The error model for a rotary joint modeled by Fourier sine series is proposed based on experimental data provided in literature. The error model of the 5-axis machine was found to be dependent on 32 linearly independent parameter groups, whose values could be evaluated using the least-squares fitting technique with errors observed in machine's workspace.

(2) The model is identified and verified experimentally using a laser tracker. A large set of volumetric error measurements collected by a laser tracker with 290 quasi-random observations in machine's workspace is done within 90 minutes, which shows the capability of laser tracker on collecting a large set of measurements efficiently.

(3) The average magnitude of residual error vectors in the two training sets are 27.7 and 30.7 microns, which are consistent with the repeatability of the machine and the fact that the thermal environment changed during the experiments. 90% and 83% of mean and maximum quasi-static error are captured by the proposed model. The modeling approach, along with the convenience of observing errors as a large set of randomly selected points in a machine's workspace with a laser tracker can make for an effective means of regularly updating compensation tables of machines.

(4) To reduce the number of observations for reducing measurement time but still get robust estimation on error parameters, different design observers including A, D and K-optimal design based on optimal design theory used in design of experiments (DOE) are proposed. Experiments have been conducted to assess the behavior of K-optimal (minimizing the condition number of the design matrix) observers. Compared with the condition number of 437.8 for 290 randomly-generated commands, the 80-point K-optimal observers have a better conditioned design matrix with condition number of 122.0. The constrained 80-point K-optimal observer chosen for with a condition number of 207.3 is also found to be an improvement.

(5) Over six identical data collection cycles, the constrained K-optimal observer set produces mean and maximum residuals of 30 and 100 microns, respectively, which are comparable to those (27.7 and 107.3 microns) produced by the 290 quasi-randomly generated point-set. More importantly, one data collection cycle takes only 24 minutes. This clearly demonstrates that a smaller strategically-chosen measurement set can produce estimates comparable to those produced by larger point-sets.

(6) To test the possibility of using the observer sets to track the thermal drift of a five-axis machine with 32 error parameters, a data collecting cycle consisting of 76 constrained K-optimal observers is used for each of the six thermal states including initial set and four heating and one cooling cycles. The mean and maximum modeling residuals for six thermal states are found to be 26.3 and 98.2 microns, respectively, which are close to the mean and maximum modeling residuals (27.7 and 107.3 microns, respectively) modeled by 290 quasi-random generated points. This also shows that using a smaller observer set does not corrupt the accuracy of the error model.

(7) The thermal error of the machine is observed to be significant (around 60 microns over the course of 320 minutes) during the operation of the machine. The largest mean residual error for the six measurement cycles conducted is observed to be 33.9 microns. During this period, if a static compensation model whose parameters were estimated with the machine in a cold state was used, the mean residual error (the average error one would expect after compensation) would have risen from 26.3 to 155.1 microns over the course of 320 minutes. If rudimentary workspace drift was compensated, the residual error would have grown from 26.3 to 98.1 microns. This not only demonstrates that the observer is able to consistently track thermal errors of the machine as its thermal is was continuously varying, but also serves as a reminder of the importance and magnitude the thermal component of quasi-static errors.

(8) It is observed that the error parameters correspond to W-axis vary the most because it holds a heat source, the spinning spindle. For the Y-axis, as it is the closet axis to W-axis and the second most heated axis, the variation of the error parameters associated with the Y-axis is also considerable. The other axes, on the other hand, behave relatively stable as the machine

being warmed up and cooled down. The proposed methodology on thermal error tracking is also capable of analyzing thermal stability of each axis of the machine.

7.1.2 POINT-SET BASED METROLOGY

Planar surfaces

(1) The concept of replacing traditional metrology with laser-scanned data and virtual gages is introduced. A metrology system using point-set, algorithms based on constrained optimization formulations has been developed. By fitting a coordinate system of the casting, point-set data representing the casting can be aligned with the nominal CAD model. The point-set manipulation algorithms are used to extract 13 features represented by sub-point-sets from the raw data set with 14 million points. The point-set filter based on convex hull is introduced to reduce the number of constraints and greatly improves the computational efficiency.

(2) To find a displacement such that all functional planar surfaces can have enough material for machining, the virtual gage analysis is developed to model the problem as a constrained min-max optimization. By solving the optimization using linear programing solver, the optimal displacement information that simultaneously satisfies all GD&T requirements can be obtained. The virtual gage analysis can be used for two metrology purposes:

    (a) Post-process examination: to determine if a finished part satisfies all GD&T requirements.

    (b) Pre-process examination: to adjust the machining reference coordinate before the raw casting is machined so that the finished part can be conforming.

(3) The validity is tested using a test part with eight functional features. Fixturing errors, which are introduced intentionally by placing spacers are detected by the virtual gage analysis and compensated by the displacement information given by the algorithm. The final casting is

measured and found to be conforming against all GD&T requirements. The feasibility of examining the raw casting before the machining process is thus demonstrated.

(4) The virtual gage analysis is enhanced to deal with the case that some material conditions might not be satisfied and the solution may not exist because the problem is overconstrained. Slack variables are introduced to release the constraints. The system is modeled by linear programming problem with slack variable. The enhanced model is tested using an industrial part with 29 functional planar features. Although some features may not have enough material for machining, the enhanced virtual gage analysis can still suggest an optimal offset information that satisfies all satisfiable material conditions. The analysis can also predict the machining allowance of every features with and without the suggested compensation. If some material conditions cannot be satisfied, the part cannot be properly machined. However, the compensation can still improve the conformity of the machined part.

Cylindrical surfaces

(1) The common tolerance specifications of cylindrical surfaces include minimum and maximum possible cylinder radii and the cylindricity error. Typically, the tolerance specifications of a cylinder represented by a 3-D point-set is approximated by 2-D data-set and its specifications. Minimum circumscribed circle (MC), maximum inscribed circle (MC) and minimum zone circle (MZ) of circular fitting problems are discussed.

(2) In defining MI of a 2-D point-set, typically, at least three points in the point-set are used. However, the inscribed circle determined using three points may not satisfy the condition that MI must be fitted within the roundness profile. Two cases are identified for the point-sets whose inscribed circles determined by at least three points are not within the roundness profile. These two cases provide internally tangent circles of MC, which are also inscribed circles of

the point-set. The Voronoi diagram-based approaches to find all these internal tangent circles of MC are developed, and MI can be defined as the largest internally tangent circle found in these two cases.

(3) A simplified approach for fitting MIC, MCC and MZC combining PSO solver and computational geometry models is developed. The approach has only two angle variables that describe the orientation of the cylinder axis. Compared with the approach in the literature, which uses five variables that describe cylinder axis's orientation (two angle variables) and offset (three translation variables), the developed model uses only two angle variables but still gets comparable results. The tolerance specifications of cylindrical surfaces including minimum and maximum acceptable radii and the cylindricity can be verified using the radii of MIC, MCC and the radial separation of MZC, respectively.

(4) Four data sets found in literature are used to test the proposed cylinder fitting method. Compared with the sequential linear programming method, the proposed model solved by Particle Swarm Optimization (PSO) has slightly different results on three types of cylinder fitting. It is because PSO can found global optimum for a nonlinear and discontinuous objective function without an accurate initial guess. The Genetic Algorithm (GA)-based approach, Geometry Optimization Searching Algorithm (GOSA)-based method and the proposed approach have similar performance since GA, GOSA and PSO algorithms all provide global optimum for the objective function. The accuracy and validity of the work are demonstrated.

7.2 RECOMMENDATIONS FOR FUTURE WORK

The research in this thesis was a step forward towards developing a scientific basis for enhancing machining accuracy through machine tool error modeling and compensation and workpiece metrology. The methodology was conceptually developed and demonstrated by examining raw casting before machining and machining after compensation. Some recommendations for future work in this area are listed below.

(1) For better performance for quasi-static error model, a thermally stable environment would be necessary. Additionally, tracking the thermal drift of the machine with time would yield better model performance. For this, a quicker (consisting of fewer and more strategically-chosen points) and more convenient data-collection cycle that can be easily embedded into the normal operation of the machine is needed. A higher order model that better describes machine's error characteristic is another approach to reduce the modeling residual.

(2) Chapter 4 has demonstrated the feasibility of tracking thermal errors with constrained K-optimal observers with periodic measurements taking only around 24 minutes to perform. Future work will address the evaluation of D- and A-optimal observers.

(3) Faster and less intrusive (than laser trackers) methods for implementing the observers need to be explored. The error model can be even simplified by replacing the error parameters with stable thermal behavior with constants. For the fewer parameters, the fewer observations would be required to get robust estimation.

(4) This work opens possibility of using temperature readings for tracking thermal errors. By correlating the estimated values of the error model parameter to temperatures in different parts of the machine, it should be possible to compute volumetric errors using only temperature readings, thus reducing the time required by, and invasiveness of, the thermal error tracking system.

(5) The main difficulty in implementing the point-set metrology is to setup the constraints in the linear programming problem because transferring the GD&T requirements (usually given in a 2D print) into algebraic inequalities is lengthy and unintuitive. To improve the practicability

of the whole process, a better user interface that helps the user to define virtual gages should be developed for future consideration.

(6) The proposed approach can be extended to tolerance verifications of straightness, concentricity and runout. More complicated tolerance verification of point-set data, for example, conicity and the profile of any given surface can be done using projection model and PSO optimization.

REFERENCES

[1] M. C. Lee, “Auto-Calibration of Machine Tools Exploiting Redundancy, Transitivity, and Error Projection,” Doctoral dissertation, University of Illinois at Urbana-Champaign, 2000.

[2] M. Weck, P. McKeown, R. Bonse, and U. Herbst, “Reduction and compensation of thermal errors in machine tools,” *CIRP Ann. Technol.*, vol. 44, no. 2, pp. 589–598, 1995.

[3] R. Ramesh, M. A. Mannan, and A. N. Poo, “Error compensation in machine tools—a review: part I: geometric, cutting-force induced and fixture-dependent errors,” *Int. J. Mach. Tools Manuf.*, vol. 40, no. 9, pp. 1235–1256, 2000.

[4] R. Venugopal, M. Barash, and M. C. Shaw, “Thermal effects on the accuracy of numerically controlled machine tools,” *CIRP Ann. Technol.*, vol. 35, no. 1, pp. 255–258, 1986.

[5] D. A. Krulewich, “Temperature integration model and measurement point selection for thermally induced machine tool errors,” *Mechatronics*, vol. 8, no. 4, pp. 395–412, 1998.

[6] J. Bryan, “International status of thermal error research,” *CIRP Ann. Technol.*, vol. 39, no. 2, pp. 645–656, 1990.

[7] P. M. Ferreira and C. R. Liu, “A contribution to the analysis and compensation of the geometric error of a machining center,” *CIRP Ann. Technol.*, vol. 35, no. 1, pp. 259–262, 1986.

[8] V. Kiridena and P. M. Ferreira, “Mapping the effects of positioning errors on the volumetric accuracy of five-axis CNC machine tools,” *Int. J. Mach. Tools Manuf.*, vol. 33, no. 3, pp. 417–437, 1993.

[9] H. Tajbakhsh, Z. Abadin, and P. M. Ferreira, “L∞ parameter estimates for volumetric error in models of machine tools,” *Precis. Eng.*, vol. 20, no. 3, pp. 179–187, 1997.

[10] V. S. B. Kiridena and P. M. Ferreira, “Parameter estimation and model verification of first order quasistatic error model for three-axis machining centers,” *Int. J. Mach. Tools Manuf.*, vol. 34, no. 1, pp. 101–125, 1994.

[11] Suh, S. H., S. K. Kang, D. H. Chung, and I. Stroud, *Theory and design of CNC systems*. Springer Science & Business Media, 2008.

[12] ASME Y14.5M-1994, *Dimensioning and Tolerancing*. The American Society of Mechanical Engineers, 1994.

[13] V. S. Kiridena, “Modeling, parameter estimation and compensation of quasistatic errors in multi-axis CNC machining centers,” Doctoral dissertation, University of Illinois at Urbana-Champaign, 1993.

[14] V. S. B. Kiridena and P. M. Ferreira, “Kinematic modeling of quasistatic errors of three-axis machining centers,” *Int. J. Mach. Tools Manuf.*, vol. 34, no. 1, pp. 85–100, 1994.

[15] J. Creamer, P. M. Sammons, D. A. Bristow, R. G. Landers, P. L. Freeman, and S. J. Easley, “Table-based volumetric error compensation of large five-axis machine tools,” *J. Manuf. Sci. Eng.*, vol. 139, no. 2, p. 021011, 2017.

[16] J. Mailhe, J. M. Linares, J. M. Sprauel, and P. Bourdet, “Geometrical checking by virtual gauge, including measurement uncertainties,” *CIRP Ann. - Manuf. Technol.*, vol. 57, no. 1, pp. 513–516, 2008.

[17] E. Pairel, P. Hernandez, and M. Giordano, “Virtual gauge representation for geometric tolerances in CAD-CAM systems,” *Model. Comput. Aided Toler. Des. Manuf. - Sel. Conf. Pap. from 9th CIRP Int. Semin. Comput. Toler. CAT 2005*, pp. 3–12, 2007.

[18] L. Mathieu and A. Ballu, *Virtual gauge with internal mobilities for the verification of functional specifications*. Boston: Springer, 1998.

[19] K. Carr and P. Ferreira, “Verification of form tolerances part I: Basic issues, flatness, and straightness,” *Precis. Eng.*, vol. 17, no. 2, pp. 131–143, 1995.

[20] K. Carr and P. Ferreira, “Verification of form tolerances part II: Cylindricity and straightness of a median line,” *Precis. Eng.*, vol. 17, no. 2, pp. 144–156, 1995.

[21] P. M. Ferreira and C. R. Liu, “A method for estimating and compensating quasistatic errors of machine tools,” *J. Eng. Ind.*, vol. 115, no. 1, pp. 149–159, 1993.

[22] E. Creighton, A. Honegger, A. Tulsian, and D. Mukhopadhyay, “Analysis of thermal errors in a high-speed micro-milling spindle,” *Int. J. Mach. Tools Manuf.*, vol. 50, no. 4, pp. 386–393, 2010.

[23] S. C. Veldhuis and M. A. Elbestawi, “A strategy for the compensation of errors in five-axis machining,” *CIRP Ann. Technol.*, vol. 44, no. 1, pp. 373–377, 1995.

[24] R. Ramesh, M. A. Mannan, and A. N. Poo, “Error compensation in machine tools—a review: Part II: thermal errors,” *Int. J. Mach. Tools Manuf.*, vol. 40, no. 9, pp. 1257–1284, 2000.

[25] S. Yang, J. Yuan, and J. Ni, “Accuracy enhancement of a horizontal machining center by real-time error compensation,” *J. Manuf. Syst.*, vol. 15, no. 2, pp. 113–124, 1996.

[26] M. Rahman, J. Heikkala, and K. Lappalainen, “Modeling, measurement and error compensation of multi-axis machine tools,” *Int. J. Mach. Tools Manuf.*, vol. 40, no. 10, pp. 1535–1546, 2000.

[27] J. Denavit, “A kinematic notation for lower-pair mechanisms based on matrices,” *Trans. ASME. J. Appl. Mech.*, vol. 22, pp. 215–221, 1955.

[28] M. A. Donmez, D. S. Blomquist, R. J. Hocken, C. R. Liu, and M. M. Barash, “A general methodology for machine tool accuracy enhancement by error compensation,” *Precis. Eng.*, vol. 8, no. 4, pp. 187–196, 1986.

[29] J. Yuan and J. Ni, “The real-time error compensation technique for CNC machining systems,” *Mechatronics*, vol. 8, no. 4, pp. 359–380, 1998.

[30] J. Ni, “CNC Machine Accuracy Enhancement Through Real- Time Error Compensation,” *J. Manuf. Sci. Eng.*, vol. 119, no. November 1997, pp. 717–725, 1997.

[31] P. N. Sheth and J. J. Uicker, “A generalized symbolic notation for mechanisms,” *J. Eng. Ind.*, vol. 93, no. 1, pp. 102–112, 1971.

[32] S. H. Suh, E. S. Lee, and S. Y. Jung, “Error modelling and measurement for the rotary table of five-axis machine tools,” *Int. J. Adv. Manuf. Technol.*, vol. 14, no. 9, pp. 656–663, 1998.

[33] P. M. Ferreira and C. R. Liu, “An analytical quadratic model for the geometric error of a machine tool,” *J. Manuf. Syst.*, vol. 5, no. 1, pp. 51–63, 1986.

[34] A. C. Okafor and Y. M. Ertekin, “Derivation of machine tool error models and error compensation procedure for three axes vertical machining center using rigid body kinematics,” *Int. J. Mach. Tools Manuf.*, vol. 40, no. 8, pp. 1199–1213, 2000.

[35] F. Y. Lin and P. L. Chiu, “A near-optimal sensor placement algorithm to achieve complete coverage-discrimination in sensor networks,” *IEEE Commun. Lett.*, vol. 9, no. 1, pp. 43–45, 2005.

[36] S. MartíNez and F. Bullo, “Optimal sensor placement and motion coordination for target tracking,” *Automatica*, vol. 42, no. 4, pp. 661–668, 2006.

[37] J. B. Bryan, “A simple method for testing measuring machines and machine tools Part 1: Principles and applications.,” *Precis. Eng.*, vol. 4, no. 2, pp. 61–69, 1982.

[38] J. C. Ziegert and C. D. Mize, “The laser ball bar: a new instrument for machine tool metrology,” *Precis. Eng.*, vol. 16, no. 4, pp. 259–267, 1994.

[39] Z. Du, S. Zhang, and M. Hong, “Development of a multi-step measuring method for motion accuracy of NC machine tools based on cross grid encoder,” *Int. J. Mach. Tools Manuf.*, vol. 50, no. 3, pp. 270–280, 2010.

[40] Y. Kakino, Y. Ihara, and Y. Nakatsu, “The measurement of motion errors of NC machine tools and diagnosis of their origins by using telescoping magnetic ball bar method,” *CIRP Ann. Technol.*, vol. 36, no. 1, pp. 377–380, 1987.

[41] A. B5.54–1992, *Methods for the Performance Evaluation of Computer Numerically Controlled Machining Centers*. 1992.

[42] R. Hocken, J. A. Simpson, B. Borchardt, J. Lazar, C. Reeve, and P. Stein, “Three dimensional metrology,” *Ann. CIRP*, vol. 26, no. 2, pp. 403–408, 1977.

[43] J. Ni and S. M. Wu, “An On-Line Measurement Technique for Machine Volumetric Error Compensation,” vol. 115, no. February, pp. 85–92, 1993.

[44] B. Bringmann, A. Küng, and W. Knapp, “A measuring artefact for true 3D machine testing and calibration,” *CIRP Ann. Technol.*, vol. 54, no. 1, pp. 471–474, 2005.

[45] T. L. Schmitz, J. C. Ziegert, J. S. Canning, and R. Zapata, “Case study: A comparison of error sources in high-speed milling,” vol. 32, no. 2, pp. 126–133, 2008.

[46] S. Aguado, D. Samper, J. Santolaria, and J. J. Aguilar, “Identification strategy of error parameter in volumetric error compensation of machine tool based on laser tracker measurements.,” *Int. J. Mach. Tools Manuf.*, vol. 53, no. 1, pp. 160–169, 2012.

[47] H. Schwenke, M. Franke, J. Hannaford, and H. Kunzmann, “Error mapping of CMMs and machine tools by a single tracking interferometer,” *CIRP Ann. Technol.*, vol. 54, no. 1, pp. 475–478, 2005.

[48] H. Schwenke, R. Schmitt, P. Jatzkowski, and C. Warmann, “On-the-fly calibration of linear and rotary axes of machine tools and CMMs using a tracking interferometer,” *CIRP Ann.*, vol. 58, no. 1, pp. 477–480, 2009.

[49] G. T. Smith, *Machine tool metrology: An industrial handbook*. Springer, 2016.

[50] J. Mayr *et al.*, “Thermal Issues in Machine Tools,” *CIRP Ann.*, vol. 61, no. 2, pp. 771–791, 2012.

[51] P. Blaser, F. Pavliček, K. Mori, J. Mayr, S. Weikert, and K. Wegener, “Adaptive learning control for thermal error compensation of 5-axis machine tools,” *J. Manuf. Syst. 44, 302-309.*, vol. 44, pp. 302–309, 2017.

[52] J. Jedrzejewski and W. Modrzycki, “A new approach to modelling thermal behaviour of a machine tool under service conditions,” *CIRP Ann. Technol.*, vol. 41, no. 1, pp. 455–458, 1992.

[53] H. Yang and J. Ni, “Dynamic neural network modeling for nonlinear, nonstationary machine tool thermally induced error,” *Int. J. Mach. Tools Manuf.*, vol. 45, no. 4–5, pp. 455–465, 2005.

[54] A. K. Srivastava, S. C. Veldhuis, and M. A. Elbestawit, “Modelling geometric and thermal errors in a five-axis CNC machine tool,” *Int. J. Mach. Tools Manuf.*, vol. 35, no. 9, pp. 1321–1337, 1995.

[55] J. Y. Yan and J. G. Yang, “Application of synthetic grey correlation theory on thermal point optimization for machine tool thermal error compensation,” *Int. J. Adv. Manuf. Technol.*, vol. 43, no. 11–12, pp. 1124–1132, 2009.

[56] A. Krause, A. Singh, and C. Guestrin, “Near-optimal sensor placements in Gaussian processes: Theory, efficient algorithms and empirical studies,” *J. Mach. Learn. Res.*, vol. 9, no. Feb, pp. 235–284, 2008.

[57] C. Q. Liu, Y. Ding, and Y. Chen, “Optimal coordinate sensor placements for estimating mean and variance components of variation sources,” *IIE Trans.*, vol. 37, no. 9, pp. 877–889, 2005.

[58] J. Kiefer, “Optimum experimental designs,” *J. R. Stat. Soc.*, vol. Series B, no. Methodological, pp. 272–319, 1959.

[59] V. V. Fedorov, *Theory of optimal experiments*. Elsevier, 2013.

[60] J. J. Ye and J. Zhou, “Minimizing the condition number to construct design points for polynomial regression models,” *SIAM J. Optim.*, vol. 23, no. 1, pp. 666–686, 2013.

[61] X. Zhao, T. M. Kethara Pasupathy, and R. G. Wilhelm, “Modeling and representation of geometric tolerances information in integrated measurement processes,” *Comput. Ind.*, vol. 57, no. 4, pp. 319–330, 2006.

[62] J. U. Turner, “Relative positioning of parts in assemblies using mathematical programming,” *Comput. Des.*, vol. 22, no. 7, pp. 394–400, 1990.

[63] J. U. Turner and M. J. Wozny, “Tolerances in computer-aided geometric design,” *Vis. Comput.*, vol. 3, no. 4, pp. 214–226, 1987.

[64] G. Ameta, J. K. Davidson, and J. Shah, “Influence of form on Tolerance-Map-generated frequency distributions for 1D clearance in design,” *Precis. Eng.*, vol. 34, no. 1, pp. 22–27, 2010.

[65] Z. Shen, G. Ameta, J. J. Shah, and J. K. Davidson, “A Comparative Study Of Tolerance Analysis Methods,” *J. Comput. Inf. Sci. Eng.*, vol. 5, no. 3, p. 247, 2005.

[66] J. K. Davidson, A. Mujezinović, and J. J. Shah, “A New Mathematical Model for Geometric Tolerances as Applied to Round Faces,” *J. Mech. Des.*, vol. 124, no. 4, p. 609, 2002.

[67] H. Chen, S. Jin, Z. Li, and X. Lai, “A comprehensive study of three dimensional tolerance analysis methods,” *CAD Comput. Aided Des.*, vol. 53, pp. 1–13, 2014.

[68] C. H. Menq, H. T. Yau, and G. Y. Lai, “Automated precision measurement of surface profile in CAD-directed inspection,” *IEEE Trans. Robot. Autom.*, vol. 8, no. 2, pp. 268–278, 1992.

[69] M. Marziale and W. Polini, “A review of two models for tolerance analysis of an assembly: Vector loop and matrix,” *Int. J. Adv. Manuf. Technol.*, vol. 43, no. 11–12, pp. 1106–1123, 2009.

[70] H.-W. Ko *et al.*, “Machine-Tool Error Observer Design With Application To Thermal,” pp. 1–11, 2018.

[71] H. W. Ko *et al.*, “Quasistatic error modeling and model testing for a 5-axis machine with a redundant axis,” *J. Manuf. Process.*, vol. 31, pp. 875–883, 2018.

[72] B. M. Colosimo, G. Moroni, and S. Petrò, “A tolerance interval based criterion for optimizing discrete point sampling strategies,” *Precis. Eng.*, vol. 34, no. 4, pp. 745–754, 2010.

[73] K. D. Summerhays, R. P. Henke, J. M. Baldwin, R. M. Cassou, and C. W. Brown, “Optimizing discrete point sample patterns and measurement data analysis on internal cylindrical surfaces with systematic form deviations,” *Precis. Eng.*, vol. 26, no. 1, pp. 105–121, 2002.

[74] R. P. Henke, K. D. Summerhays, J. M. Baldwin, R. M. Cassou, and C. W. Brown, “Methods for evaluation of systematic geometric deviations in machined parts and their relationships to process variables,” *Precis. Eng.*, vol. 23, no. 4, pp. 273–292, 1999.

[75] S. M. Obeidat and S. Raman, “An intelligent sampling method for inspecting free-form surfaces,” *Int. J. Adv. Manuf. Technol.*, vol. 40, no. 11–12, pp. 1125–1136, 2009.

[76] M. Affan Badar, S. Raman, and P. S. Pulat, “Intelligent Search-Based Selection of Sample Points for Straightness and Flatness Estimation,” *J. Manuf. Sci. Eng.*, vol. 125, no. 2, p. 263, 2003.

[77] K. M. Carr, “Modeling and verification methods for the inspection of geometric tolerances using point data,” Doctoral dissertation, University of Illinois at Urbana-Champaign, 1995.

[78] E. Mainsah, J. A. Greenwood, and D. G. Chetwynd, *Metrology and properties of engineering surfaces*. Springer Science & Business Media, 2001.

[79] M. Giordano, L. Mathieu, and F. Villeneuve, *Product Life-Cycle Management: Geometric Variations*. John Wiley & Sons, 2012.

[80] P. Axelsson, “Processing of laser scanner data—algorithms and applications,” *ISPRS J. Photogramm. Remote Sens.*, vol. 54, no. 2–3, pp. 138–147, 1999.

[81] G. F. Marshall and G. E. Stutz, *Handbook of Optical and Laser Scanning, Second Edition*, 2nd editio. CRC Press, 2011.

[82] R. B. Rusu and S. Cousin, “3d is here: Point cloud library (pcl),” in *IEEE International Conference on Robotics and Automation*, 2011, pp. 1–4.

[83] T. S. R. Murthy and S. Z. Abdin, “Minimum zone evaluation of surfaces,” *Int. J. Mach. Tool Des. Res.*, vol. 20, pp. 123–136, 1980.

[84] X. Wen, Q. Xia, and Y. Zhao, “An effective genetic algorithm for circularity error unified evaluation,” *Int. J. Mach. Tools Manuf. 46*, vol. 46, pp. 1770–1777, 2006.

[85] L. Xiuming and S. Zhaoyao, “Development and application of convex hull in the assessment of roundness error,” *Int. J. Mach. Tools Manuf. 48*, vol. 48, pp. 135–139, 2008.

[86] E. Welzl, “Smallest enclosing disks (balls and ellipsoids),” *New results new trends Comput. Sci.*, pp. 359–370, 1991.

[87] U. Roy and X. Zhang, “Establishment of a pair of concentric circles with the minimum radial separation for assessing roundness error,” *Comput. Des.*, vol. 24, no. 3, pp. 161–168, 1992.

[88] U. Roy and X. Zhang, “Development and application of voronoi diagrams in the assessment of roundness error in an industrial environment,” *Comput. Ind. Eng.*, vol. 26, no. 1, pp. 11–26, 1994.

[89] A. Okabe, B. Boots, K. Sugihara, and S. N. Chiu, *Spatial tessellations: concepts and applications of Voronoi diagrams*. John Wiley & Sons, 2009.

[90] K. Kim, S. Lee, and H. Jung, “Assessing Roundness Errors Using Discrete Voronoi Diagrams,” *Precis. Eng.*, vol. 16, no. 8, pp. 559–563, 2000.

[91] G. L. Samuel and M. S. Shunmugam, “Evaluation of circularity from coordinate and form data using computational geometric techniques,” *Precis. Eng.*, vol. 24, no. 3, pp. 251–263, 2000.

[92] F. Liu, G. Xu, L. Liang, Q. Zhang, and D. Liu, “Minimum Circumscribed Circle and Maximum Inscribed Circle of Roundness Deviation Chord Method,” *IEEE Trans. Instrum. Meas.*, vol. 65, no. 12, pp. 2787–2796, 2016.

[93] C. M. Shakarji and V. Srinivasan, “Computational Investigations for a New, Constrained Least-Squares Datum Definition for Circles, Cylinders, and Spheres,” *ASME 2016 Int. Mech. Eng. Congr. Expo.*, p. V011T15A007-V011T15A007, 2018.

[94] S. H. Cheraghi, G. Jiang, and J. Sheikh, “Evaluating the geometric characteristics of cylindrical features,” *Precis. Eng.*, vol. 27, no. 2, pp. 195–204, 2003.

[95] H. Lai, W. Jywe, C. Chen, and C. Liu, “Precision modeling of form errors for cylindricity evaluation using genetic algorithms,” *Precis. Eng.*, vol. 24, no. 4, pp. 310–319, 2000.

[96] X. Wen and A. Song, “An improved genetic algorithm for planar and spatial straightness error evaluation,” *Int. J. Mach. Tools Manuf.*, vol. 43, no. 11, pp. 1157–1162, 2003.

[97] L. Xianqing, S. Hongwei, X. Yujun, L. Jishun, Z. Jia, and D. Mingde, “Method for cylindricity error evaluation using Geometry Optimization Searching Algorithm,” *Measurement*, vol. 44, no. 9, pp. 1556–1563, 2011.

[98] X. Wen and A. Song, “An immune evolutionary algorithm for sphericity error evaluation,” *Int. J. Mach. Tools Manuf.*, vol. 44, pp. 1077–1084, 2004.

[99] X. Wen, J. Huang, D. Sheng, and F. Wang, “Conicity and cylindricity error evaluation using particle swarm optimization,” *Precis. Eng.*, vol. 34, no. 2, pp. 338–344, 2010.

[100] J. B. Bryan, “The Abbe principle revisited: an updated interpretation,” *Precis. Eng.*, vol. 1, no. 3, pp. 129–132, 1979.

[101] P. M. Pardalos and G. Schnitger, “Checking local optimality in constrained quadratic programming is NP-hard,” *Oper. Res. Lett.*, vol. 7, no. 1, pp. 33–35, 1988.

[102] *ISO 12181-1,2:2011 Geometrical product specifications (GPS) — Roundness — Part 1: Vocabulary and parameters of roundness*. .

[103] M. R. P. Kumar, P. P. Kumaar, and R. Kameshwaranath, “Roundness error measurement using teaching learning based optimization algorithm and comparison with particle swarm optimization algorithm,” *Int. J. Data Netw. Sci.*, vol. 2, no. 3, pp. 63–70, 2018.

[104] T. C. Chang and R. A. Wysk, *Computer-aided manufacturing*. Prentice Hall PTR, 1997.

[105] C. Du, C. Luo, Z. Han, and Y. Zhu, “Applying particle swarm optimization algorithm to roundness error evaluation based on minimum zone circle,” *Measurement*, vol. 52, pp. 12–21, 2014.

APPENDIX A. CYLINDRICITY VERIFICATION DATA SET 1

Table A.1: Data set 1 [20]

No.	X	Y	Z	No.	X	Y	Z
1	-11.820859	50.421254	-15.817382	21	33.207329	64.844079	-10.665479
2	42.403448	-6.693162	56.567707	22	34.46129	41.806234	94.623903
3	10.366902	80.249947	26.965969	23	-26.871029	3.103967	39.48246
4	18.527457	61.577469	-13.680418	24	-4.153639	67.427229	23.451422
5	23.930322	23.878386	-41.820643	25	22.371	47.845956	88.060867
6	66.363729	0.636729	49.246025	26	67.398986	16.520701	79.062822
7	-3.608026	-24.493246	39.678687	27	79.257377	49.418921	4.727043
8	75.507564	20.208045	6.298139	28	-37.543275	31.718373	8.573268
9	48.919097	55.614254	-13.266609	29	49.576671	65.965076	-6.501629
10	65.713317	2.841028	3.498858	30	96.781947	53.421231	22.908004
11	46.632786	80.517454	4.866333	31	-18.623157	23.988046	47.691608
12	13.598993	83.519129	30.375	32	58.416292	-4.557784	48.525368
13	84.570573	18.219363	28.224203	33	48.408528	15.833662	81.511728
14	2.322453	-10.802862	51.268799	34	31.694971	-2.169579	63.538387
15	82.820384	38.516367	9.148307	35	-18.366214	2.837799	46.415679
16	3.553158	75.111087	30.738097	36	81.087477	11.573666	46.319607
17	-5.898713	21.39033	60.097056	37	57.311572	-9.09605	38.123767
18	30.009532	-24.696147	35.870356	38	68.59397	33.580936	-6.118165
19	-3.793621	-14.263808	46.897322	39	89.036231	21.72231	35.086999
20	58.357492	87.161327	11.960644	40	3.141412	52.730721	67.919265

APPENDIX B. CYLINDRICITY VERIFICATION DATA SET 2

Table B.1: Data set 2 [20]

No.	X	Y	Z	No.	X	Y	Z
1	60.051121	0.002953	3.946134	11	50.684216	-32.022045	22.865941
2	-57.932024	15.399312	15.983017	12	57.318676	17.619539	22.082457
3	57.432130	17.488707	20.365942	13	-40.408130	-44.485701	22.692315
4	55.022756	-23.936632	11.505062	14	-39.838370	44.994386	7.411167
5	29.180100	-52.423113	1.037163	15	-10.261352	-59.146784	22.600675
6	-58.861558	-11.113569	20.134482	16	53.919844	26.493193	18.949042
7	-44.597179	40.113733	2.005267	17	-8.540012	59.442972	13.092342
8	-23.247383	-55.406652	17.669299	18	-59.369089	8.361285	7.133233
9	34.041568	-49.309081	15.807863	19	-38.029817	46.404843	4.995216
10	-34.084135	-49.427745	12.479981	20	47.946099	-35.925380	27.276243

APPENDIX C. CYLINDRICITY VERIFICATION DATA SET 3

Table C.1: Data set 3 [95]

No.	X	Y	Z	No.	X	Y	Z
1	11.0943	0.4522	65.2328	13	10.8150	0.5918	85.2304
2	5.0940	10.8450	65.0765	14	4.8148	10.9846	85.0740
3	-6.9063	10.8439	65.0089	15	-7.1855	10.9835	85.0641
4	-12.9065	0.4498	65.0897	16	-13.1858	0.5894	84.8952
5	-6.9063	-9.9429	65.0540	17	-7.1855	-9.8033	85.0516
6	5.0940	-9.9418	65.2216	18	4.8149	-9.8022	85.2171
7	10.9546	0.5220	75.2316	19	10.6754	0.6616	95.2291
8	4.9544	10.9148	75.0752	20	4.6752	11.0544	95.0728
9	-7.0459	10.9137	75.0770	21	-7.3253	11.0533	94.9077
10	-13.0461	0.5196	74.8964	22	-13.3254	0.6592	95.0940
11	-7.0459	-9.8731	75.0528	23	-7.3252	-9.7335	95.0504
12	4.9545	-9.8720	75.2204	24	4.6752	-9.7323	95.2170

APPENDIX D. CYLINDRICITY VERIFICATION DATA SET 4

Table D.1: Data set 4 [97]

No.	X	Y	Z	No.	X	Y	Z
1	-7.169	-34.21	30.5	41	6.174	34.448	60.8
2	-14.324	-31.887	30.5	42	14.594	31.802	60.8
3	-25.509	-23.877	30.5	43	28.945	19.636	60.8
4	-33.683	-9.283	30.5	44	34.887	2.506	60.8
5	-34.167	7.378	30.5	45	31.458	-15.279	60.8
6	-25.051	24.377	30.5	46	20.538	-28.299	60.8
7	-7.1	34.238	30.5	47	6.489	-34.352	60.8
8	8.071	34.034	30.5	48	-7.37	-34.173	60.8
9	19.797	28.86	30.5	49	-15.817	-31.168	81
10	28.945	19.647	30.5	50	-28.573	-20.125	81
11	34.891	2.595	30.5	51	-33.084	-11.238	81
12	34.168	-7.484	30.5	52	-34.611	4.838	81
13	28.525	-20.224	30.5	53	-25.488	23.924	81
14	18.113	-29.908	30.5	54	-16.38	30.884	81
15	7.971	-34.044	30.5	55	-0.788	34.969	81
16	-6.778	-34.29	30.5	56	10.206	33.457	81
17	-11.322	-33.073	43.3	57	21.025	27.971	81
18	-20.262	-28.475	43.3	58	30.269	17.524	81
19	-24.016	-25.389	43.3	59	33.93	8.513	81
20	-31.014	-16.098	43.3	60	34.787	-3.557	81
21	-34.584	-4.991	43.3	61	31.626	-14.895	81
22	-30.614	16.878	43.3	62	18.362	-29.755	81
23	-19.46	29.045	43.3	63	3.4	-34.794	81
24	-5.802	34.484	43.3	64	-5.318	-34.548	81
25	10.248	33.444	43.3	65	-8.484	-33.909	97.5
26	23.395	26.026	43.3	66	-18.607	-29.583	97.5
27	32.027	14.043	43.3	67	-26.163	-23.175	97.5
28	34.985	0.374	43.3	68	-33.973	-8.156	97.5
29	33.364	-10.449	43.3	69	-34.49	5.71	97.5
30	27.418	-21.716	43.3	70	-29.015	19.502	97.5
31	17.678	-30.164	43.3	71	-17.564	30.231	97.5
32	-0.706	-34.942	43.3	72	-3.842	34.765	97.5
33	-14.539	-31.801	60.8	73	15.389	31.424	97.5
34	-22.717	-26.563	60.8	74	29.428	18.935	97.5
35	-26.655	-22.607	60.8	75	34.054	8.042	97.5
36	-31.59	-14.943	60.8	76	34.931	-2.026	97.5
37	-34.892	1.967	60.8	77	33.204	-11.008	97.5
38	-29.347	18.995	60.8	78	29.818	-18.273	97.5
39	-19.04	29.322	60.8	79	21.451	-27.629	97.5
40	-6.241	34.407	60.8	80	15.791	-31.216	97.5