HIGH-SPEED MILLING OF TITANIUM ALLOYS:

MODELING AND OPTIMIZATION

WANG ZHIGANG

NATIONAL UNIVERSITY OF SINGAPORE

2005

HIGH-SPEED MILLING OF TITANIUM ALLOYS: MODELING AND OPTIMIZATION

WANG ZHIGANG

(B. Eng, M. Eng)

A THESIS SUBMITTED

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF MECHANICAL ENGINEERING

NATIONAL UNIVERSITY OF SINGAPORE

2005

ACKNOWLEDGEMENTS

I would like to express my deepest and heartfelt gratitude and appreciation to my supervisors, Professor Mustafizur Rahman and Associate Professor Wong Yoke-San, for their valuable guidance, continuous support and encouragement throughout the entire research work. I also want to take this opportunity to show my sincere thank to National University of Singapore (NUS) for providing me a research scholarship and to Advance Manufacturing Lab (AML) for the excellent facilities without which the present work would not have been done.

I would like to thank Assoc. Prof. Li Xiaoping for his precious advice about the cutting force model. I would also like to thank the following staffs for their help without which this project would not be successfully completed: Mr. Tan Choon Huat, Lim Soon Cheong and Wong Chian Long from Advanced Manufacturing Lab (AML), who provided technical assistance in performing the machining operations and Mr. Kwa Lam Koon from CITA who helped to configure parallel computation environments.

Special thanks come to my family members for their continuous support and understanding that help me complete this work successfully.

At various stages of this research work, a lot of encouraging supports and help were delivered by my friends. Thanks also come to my friends, Dr. Liu Kui, Mr. Yong Dong, Dr. Sun Jie, Mr. Fan Liqing, Mr. Wu Yifeng, Li Lingling, Li Tao, Wang Yue, Reza, Tauhid, Ibrahim, Majharul, Tabassumul and Sonti.

CONTENTS

SUMMARY

With the advent of high-performance CAD/CAM systems and CNC machines, high-speed machining (HSM) has established its dominant position among other rapid manufacturing techniques. High-speed milling of aluminum has been applied successfully for more than a decade; however, high-speed applications on the difficult-to-cut materials, such as titanium alloys, are still relatively new. Titanium alloys have been widely used in the aerospace, biomedical, automotive and petroleum industries because of their good strength-to-weight ratio and superior corrosion resistance. However, it is very difficult to machine them due to their poor machinability. Among all titanium alloys, Ti-6Al-4V is most widely used. Due to the poor machinability of Ti-6Al-4V, selecting the optimal machining conditions and parameters is crucial.

In this study, a new type of tool, which is binder-less cubic boron nitride (BCBN), has been used for high-speed milling of Ti-6Al-4V. Firstly, the effects of cutting speed, feed rate per tooth and depth of cut on cutting forces and tool life are investigated based on the experimental results at different cutting conditions. The wear mechanism is also analyzed. Then, a new approach for theoretical modeling of the milling process geometry is presented, which ensures the analytical solution to accurate undeformed chip thickness. Since the axial depth of cut in this study is smaller than the nose radius of the cutter, the effect of tooth radius is considered in the calculation of the uncut chip area. Moreover, the non-uniform chip area is represented with an equivalent element. The Johnson-Cook (JC) flow stress model is used to describe the deformation behavior of Ti-6Al-4V. After obtaining the JC constitutive model of flow stress and the equivalent element representation, a finite element method (FEM) is used to simulate

the high-speed milling of Ti-6Al-4V. Then, a new cutting force model is proposed based on FEM-simulation results and Oxley's cutting force model. Experimental verification is also provided to justify the accuracy of the developed cutting force model. Based on the cutting force model and the analytical solution to the true cutting path trajectory in milling, the constraints about surface roughness, cutting forces and machining power have been determined for the optimization model.

In this study, a new advanced searching method genetic simulated annealing (GSA), which is a hybrid of GA and SA, is developed and used to determine optimal HSM cutting strategies for milling operations. In order to improve its efficiency further, GSA has been parallelized with hierarchical parallel GA model. In the optimization model, two objectives are considered: minimum production time and production cost. For this multi-objective optimization problem, the fitness assignment is based on the concept of non-dominated sorting genetic algorithm (NSGA). For each simulation of parallel GSA (PGSA), a Pareto-optimal front has been found, which is composed of many Pareto-optimal solutions. Along the Pareto-optimal front, the optimal cutting parameters have been found with a weighted average strategy. Then, based on the concept of dynamic programming, the optimal cutting strategy has been obtained. Two case studies are given for the verification of the simulation results. Based on the experimental results and comparison with other algorithms, PGSA together with non-dominated sorting methodology is found to be much more suitable for multi-objective optimization of the cutting parameters for milling operation.

LIST OF TABLES

LIST OF FIGURES

NOMENCLATURE

Symbol	Units	Description
a	*mm*	depth of cut
AB	*mm*	shear plane near the centre of the chip formation zone
C, C'	–	strain rate constants for chip formation zone
C_l, C_o	\$/*min*	labor and overhead cost
C_{mat}, C_t	\$	cost of work material and the cutting tools
C_{pr}	\$/*min*	production cost per component
D, R	*mm*	nominal diameter and radius of the cutter
d_a	mm	diameter of the arbor
E	MPa	modulus of elasticity of arbor material
f_z	*mm/r*	feed rate of the milling cutter per tooth per revolution
f_{zmax}, f_{zmin}	*mm/r*	maximum and minimum feed rate per tooth per revolution
F_d	*N*	permissible force with regard to deflection
F_s	*N*	permissible force with regard to strength
F_C	*N*	force component in the cutting direction
F_f	*N*	frictional force at tool-chip interface
F_n	*N*	normal force at tool-chip interface
F_N	*N*	force normal to the shear plane
F_r, F_t	*N*	force components in radial and tangential directions
F_R	*N*	resultant cutting force
F_S	*N*	shear force on the shear plane
F_T	*N*	force components in the direction of cutting
F_X, F_Y, F_Z	*N*	forces in X, Y and Z directions
h	*mm*	undeformed chip thickness
h_e	*mm*	chip thickness of the equivalent element
I, J	mm^4	moment of inertia of stub arbor
K	$Wm^{-1}K^{-1}$	thermal conductivity
k_{AB}	*Pa*	shear flow stress on the shear plane
k_{chip}	*Pa*	average shear flow stress at chip-tool interface
K_t	*Pa*	specific cutting force in the mechanistic model

l	*mm*	length of the shear plane
L	*N*	length of cut (mm)
l_c	*mm*	tool-chip contact length
L_s	*mm*	length of the holder
n	–	strain-hardening index in the empirical stress-strain relation
n	*r/min*	rotational speed of the cutter
N_b	–	total number of components in the batch
N_p	–	number of passes
NP	–	number of processors
P_c	kW	cutting power
P_m	kW	nominal motor power
r	*mm*	radius of the tool tip
R_a, $Ra_{\max}$	*mm*	surface roughness and its maximum permissible value
R_{max}	*mm*	maximum height of the machined surface irregularities
R_T	–	thermal number
S	*J/kg K*	specific heat
S_i	mm^2	uncut chip area
T	*min*	tool life
t_1	*mm*	undeformed chip thickness
t_2	*mm*	chip thickness
T_a	*min*	process adjusting time and quick return time
T_{AB}	*°C*	temperature at the shear plane
T_{int}	*°C*	average temperature along tool-chip interface
T_L	*min*	loading and unloading time
T_m	*min*	machine time
T_{mod}	*°C*	velocity-modified temperature
T_p	*min*	machine preparation time per component
T_{pr}	*min*	tool production time per component
T_s	*min*	set up time of the machine for a new batch
T_{tc}	*min*	tool changing time per component
T_{tool}	*°C*	tool temperature
T_w	*°C*	the initial temperature of the workpiece
ΔT	*°C*	temperature rise

ΔT_C	$^{\circ}C$	average temperature rise in chip
ΔT_M	$^{\circ}C$	maximum temperature rise in chip
ΔT_{SZ}	$^{\circ}C$	temperature rise in chip formation zone
v, V_C	*m/min*	cutting speed
V_{chip}	*m/min*	chip velocity in orthogonal cutting
v_{max},v_{min}	*m/min*	maximum and minimum available cutting speed
V_S	*m/min*	shear velocity
z	–	number of teeth on the cutter
α	*deg*	rake angle
β	–	Proportion of heat conducted into workpiece material
β	*deg*	clearance angle
ϕ	*deg*	shear angle
γ_{AB}	–	shear strain at the shear plane *AB*
$\dot{\gamma}_{AB}$	s^{-1}	maximum shear strain rate at the shear plane *AB*
$\dot{\gamma}_{\text{int}}$	s^{-1}	shear strain rate at the tool-chip interface
δ	–	chip compression ration (t_1/t_2)
δ, δ_{max}	*mm*	deflection and maximum deflection of the tool holder
ε	–	effective strain
ε_{AB}	–	effective strain at *AB*
$\dot{\varepsilon}$	s^{-1}	effective strain-rate
$\dot{\varepsilon}_0$, v	–	constants in velocity modified temperature equation
ς	*deg*	angle depending on h and φ
η	–	overall efficiency
φ	*deg*	instantaneous angular position of the tooth of the cutter
φ_0	*deg*	equivalent angle of φ and ξ
θ	*deg*	angle made by F_R and *AB*
$\theta_0,\theta_1,\theta_2,\theta_3$	*deg*	Limits of integration
θ_e	*deg*	the engagement angle between the nose radius and the chip
λ	*deg*	mean friction angle
ξ	–	the ratio between tangential and radial cutting force
ρ	kg/m^3	density of the workpiece material

σ_1	*Pa*	value of σ at $\varepsilon = 1$
σ_N	*Pa*	normal stress at the tool-chip interface
τ	*Pa*	frictional shear stress
τ_{int}	*Pa*	resolved shear stress at tool-chip interface
τ_s	*Pa*	permissible torsional stress of the arbor material
Ψ	–	tool-chip interface temperature factor

Chapter 1
Introduction

This chapter introduces high-speed milling of titanium alloys, and presents a brief overview of the optimization of machining processes, the main research objectives, and the general structure of this dissertation. Sections 1.1 and 1.2 describe high-speed machining in general and high-speed machining of titanium alloys, respectively. Section 1.3 presents a brief overview of the optimization of machining processes. The research objectives are stated in Section 1.4, followed by Section 1.5 which outlines the organization of this dissertation.

1.1 High-speed machining

With the wide use of CNC machines together with high-performance CAD/CAM systems, high-speed machining (HSM) has demonstrated its superior advantages to other rapid manufacturing techniques. In addition to the increased productivity, HSM is capable of generating high-quality surfaces, burr-free edges and a virtually stress-free component after machining, and it can be used to machine thin-wall workpieces, because the cutting forces in HSM are lower. Another significant advantage of high-speed machining is minimization of effects of heat on machined parts. Most of the cutting heat is removed, reducing thermal warping and increasing the life of the cutting tool. In many cases, the need for a cooling fluid is eliminated. Also, the elimination of cutting fluids reduces subsequent contributions to pollution and aids in the recovery and recycling of such expensive materials as aluminum-lithium alloys. Since HSM has

so many advantages, it is widely used in aircraft and aerospace industry, automotive industry, household equipments, precision mechanics and optical industry and so on. For example, in the aerospace industry the wing spars are machined from expensive forged aluminum billets while the stringers are machined from milled bars. The final geometry of the spar and the stringers requires up to 90 percent of the original material to be removed. In addition, the generation of stress-free surfaces on the component parts is of vital importance. Consequently the economics of the process largely depends on the metal removal rates. The ability to maintain high removal rates also depends on the ability to control the chips in the machining area. This implies that applying HSM techniques to large aerospace components is economically attractive.

1.2 HSM of titanium alloys – Ti-6Al-4V

Although high-speed milling of aluminum has been applied in industries successfully for more than a decade, high-speed applications on the difficult-to-cut materials such as titanium alloys are still relatively new.

Titanium alloys have been widely used in the aerospace, biomedical, automotive and petroleum industries because of their good strength-to-weight ratio and superior corrosion resistance. However, it is very difficult to machine them due to their poor machinability. During the machining of titanium alloys with conventional tools, tool wear progresses rapidly because their low thermal conductivity and high chemical reactivity result in higher cutting temperature and strong adhesion between the tool and the work material (Zoya and Krishnamurthy, 2000). In 1955, Siekmann (1955) pointed out that “machining of titanium and its alloys would always be a problem, no matter what techniques are employed to transform this metal into chips”. The poor

machinability of titanium and its alloys have led many large companies (for example Rolls-Royce and General Electrics) to invest much in developing techniques to minimize machining cost (Ezugwu and Wang, 1997).

Among all titanium alloys, Ti-6Al-4V is most widely used. When high-speed milling Ti-6Al-4V, the wear on the conventional tools, such as high-speed steel (HSS) and tungsten carbide (WC), progresses rapidly, because the poor thermal conductivity of Ti-6Al-4V results in the higher temperature closer to the cutting edge during machining. And there exists strong affinity between the tool and workpiece material. In addition, titanium alloys are generally difficult to machine at cutting speed over 30m/min with HSS tools, and over 60m/min with cemented WC tools, which results in the very low productivity. Other types of tool materials, including ceramic, diamond, and cubic boron nitride (CBN), are highly reactive with titanium alloys at higher temperature, and consequently they are not suitable to be used in high-speed milling of Ti-6Al-4V (Lopez de lacalle et al., 2000).

1.3 Optimization of machining process

Due to the poor machinability of Ti-6Al-4V, selecting the optimal machining conditions and parameters is crucial. The determination of efficient machining parameters has been a problem that has confronted manufacturing industries for nearly a century, and is still the subject of many studies. Optimal machining parameters are of great concern in manufacturing environments, where economy of machining operation plays a key role in the competitive market. Economic machining has gained great importance as CNC machines are extensively employed worldwide. Although CNC machines can noticeably eliminate auxiliary tooling and reduce set-up time, the

machining time is almost the same as in conventional machining when machining parameters are selected from machining databases or handbooks. Because the CNC machines are more expensive than the conventional ones, there is a greater need to operate CNC machines as efficiently as possible in order to obtain the required pay-back. Optimal values of cutting parameters have to be determined before loading workpieces on the CNC machines, since the cost of machining on CNC machines is sensitive to the cutting parameters.

1.4 Main objectives of this study

This study focuses on the theoretical modeling and optimization of cutting parameters for high-speed milling of Ti-6Al-4V with binder-less CBN (BCBN) cutting tools. Firstly, cutting performance of BCBN tools will be investigated and the wear mechanism is also analyzed. Then, a new approach for theoretical modeling of the undeformed chip thickness will be proposed. A cutting force model will be developed based on Oxley's cutting force theory for more accurate prediction of the cutting forces and application to high-speed milling of Ti-6Al-4V. Based on the analytical solution to the cutting pass trajectory and cutting force model, the constraints related to surface roughness, cutting forces and the machining power will be determined for a subsequently developed optimization model. Following these studies, an advanced search algorithm, parallel genetic simulated annealing, is proposed to find the optimal cutting parameters for high speed milling of Ti-6Al-4V using two objective functions, minimum production time and minimum production cost.

In order to achieve the above objectives, the following necessary sub-objectives need to be accomplished:

- Investigation of cutting performance of BCBN tools in terms of cutting forces and tool life when used for high-speed milling of Ti-6Al-4V, and analysis of the wear mechanism.
- Building a tool life model for high-speed milling of Ti-6Al-4V with BCBN tools to determine production cost.
- Building an accurate tooth trajectory model and a cutting force model for high-speed milling of Ti-6Al-4V with BCBN tools. Based on these two models, identifying and studying the cutting forces and machining power constraints for the optimization model, such that practically feasible optimal cutting strategy can be obtained.
- Analysis and development of an efficient genetic simulated annealing (GSA) search algorithm and its parallelized version, parallel genetic simulated annealing (PGSA), and development of a toolkit to implement these two algorithms.
- Comparing the proposed algorithm PGSA with other optimization approaches to validate the efficiency of PGSA based on a set of test functions.
- Using the algorithm PGSA to obtain optimal cutting parameters for high-speed milling of Ti-6Al-4V with BCBN tools according to two objective functions: minimum production time and minimum production cost.

1.5 Organization of this dissertation

There are eight chapters in this dissertation. In this chapter, the problem of high-speed milling of titanium is first described. Then, a brief overview of the optimization of machining processes is presented. Finally, the main research objectives are described. This section outlines the organization of this dissertation.

In Chapter 2, a comprehensive review is given, which includes five parts. Firstly, the previous work on machining of titanium alloy Ti-6Al-4V is presented. Then, a brief overview of geometrical modeling of milling process is given. The review of cutting force models is next presented, with focus on the analytical and numerical models. In addition, four types of often used optimization algorithms and the hybrid of genetic algorithm and simulated annealing are described. Finally, the optimization of machining processes done by other researchers is reported.

Chapter 3 describes the experimental setup. In this chapter, experimental details and the equipment used in this study are first given. Then, the corresponding experimental methods for measuring cutting forces and tool life are explained.

Chapter 4 presents the investigations of the cutting performance when slot milling titanium alloy Ti-6Al-4V in terms of cutting forces, tool life and wear mechanism. A new tool material, which is binder-less cubic boron nitride (BCBN), is used for high-speed milling of Ti-6Al-4V. The effects of cutting speed, feed rate per tooth and depth of cut on cutting forces and tool life are investigated based on the experimental results at different cutting conditions. From the observations on the flank and rake faces of the worn-out cutting tools with SEM, EDX and XRD, the main wear mechanisms of the BCBN tools are analyzed.

Chapter 5 describes the development of a cutting force model for high-speed milling of Ti-6Al-4V. In this chapter, a brief review of Oxley's cutting force model is given, and the theoretical modeling of milling process geometry is presented. Then, the non-uniform chip area has been represented with a 2-D equivalent element. The Johnson-

Cook (JC) flow stress model is used to describe flow stress properties of Ti-6Al-4V. After obtaining the JC constitutive model of flow stress and the equivalent element representation, the chip flow during high-speed milling of Ti-6Al-4V is simulated with finite element method (FEM). Based on FEM simulation results and Oxley's predictive cutting force theory, a new cutting force model is proposed. Finally, experimental verification is also provided to justify the accuracy of the cutting force model.

Chapter 6 presents a parallel genetic simulated annealing (PGSA) algorithm that has been developed and benchmarked with a set of test functions. Firstly, the hybrid of genetic algorithm and simulated annealing, referred as GSA, is described. Then, the hierarchical parallel GA model was used to realize the parallelization of GSA. The full details about the implementation of PGSA, such as the representation of individuals, selection, crossover and mutation of operator of GSA, and migration strategy, are given. Finally, the performance of PGSA is evaluated against a standard set of test functions in comparison to other advanced search algorithms.

In Chapter 7, the optimization of multi-pass milling is investigated with regard to optimal cutting passes, and three corresponding cutting parameters: depth of cut, cutting speed and feed rate. Two objectives, minimum production time and production cost, are considered. For this multi-objective optimization problem, the fitness assignment is based on the concept of non-dominated sorting genetic algorithm (NSGA). Then, the optimization algorithm PGSA described in Chapter 6 is used to find the optimal cutting parameters under certain constraints. For each simulation, PGSA can find a Pareto-optimal front which is composed of many Pareto-optimal solutions. Along the Pareto-optimal front, the optimal cutting parameters can be found

with the weighted average strategy. Then, based on the concept of dynamic programming, the optimal cutting strategy can be obtained. Finally, two case studies are given for the verification of the simulation results.

Chapter 8 concludes the thesis with a summary of contributions, and the directions for future work are also suggested.

Chapter 2
Literature review

Although high-speed milling of aluminum is widely used in aerospace industry, high-speed applications on difficult-to-cut materials such as titanium alloys are still relatively new. There is still a need to investigate the cutting mechanism for high-speed milling of titanium alloys. This chapter introduces an overview of high-speed machining of titanium alloys; then a brief overview of milling process modeling and conventional optimization algorithms provides a theoretical base for the remainder of the work. Section 2.1 describes the previous work done on the machining and high-speed machining of titanium alloys. The review of the geometrical models and cutting force models of milling processes is provided in Sections 2.2 and 2.3, respectively. Section 2.4 presents a brief overview of the often used optimization algorithms which is the foundation for the development of the advanced search algorithm, parallel genetic simulated annealing (PGSA). Section 2.5 presents an overview of the optimization of machining processes and the ways to handle constraints for optimization problems, followed by the conclusions of this chapter in the last section.

2.1 Previous work about high-speed machining of titanium alloys

A literature review reveals that the machining of titanium and its alloys have not received much attention in recent years. This may result from the difficulties associated with machining of titanium and its alloys.

Titanium is a poor conductor of heat. Heat, generated by the cutting action, does not dissipate quickly. Therefore, most of the heat is concentrated on the cutting edge and the tool face, and thus tool life is adversely affected.

In addition, titanium has a strong alloying tendency or chemical reactivity with the cutting tool materials at high cutting temperatures. This causes welding to the tool during the machining operation and consequent galling, smearing, and chipping of the machined surface, along with rapid destruction of the cutting tool.

Another property of titanium is that it has a lower elastic modulus than steel and super-alloys and thus it has higher "springiness" than these metals. The higher springiness results in greater deflections of a workpiece during practical machining process. Thus, proper measures may be taken to improve stiffness during machining of titanium and its alloys. Normally, sharp and properly shaped cutting tools are preferred. In addition, greater clearances of cutting tools are also required due to these deflections.

As described in Section 1.2 of Chapter 1, the cutting performance of conventional HSS and WC cutting tools is very poor when machining titanium alloys. Some researchers tried coated carbide tools, because the coating can reduce the friction between tool/chip interface and it can also provide a good thermal barrier for the tool. López de lacalle et al. (2000) and Jawaid et al. (1999) have investigated the cutting performance of milling titanium alloys with coated carbide tools. López de lacalle et al. (2000) recommended that the cutting speed range is between 50 and 100m/min when machining titanium alloys with coated carbide. Actually, this speed range still belongs to conventional cutting speed.

King and Vaughn (1984) stated that as the cutting speed increases above the conventional speed range, new dynamic effects are encountered in the cutting process, and the cutter configuration must be considered a function of the cutting speed and other process parameters. Therefore, Taylor's empirical equations are no longer adequate since they only considered cutting speed. There is a necessity to investigate the mechanism of HSM of titanium and its alloys.

Recently, other researchers have tried advanced tool materials such as cubic boron nitride (CBN) and polycrystalline diamond (PCD), to achieve HSM. Zoya and Krishnamurthy (2000) used CBN tools for high-speed turning of titanium alloys and evaluated the machining performance. They found that a good surface finish can be achieved with a cutting speed of 185m/min, and a cutting speed range of 185-220m/min can be recommended for the machining of titanium alloys with CBN tools. In their study, deformation at the cutting nose of CBN tools was observed during the machining of titanium alloys, and they claimed that wear of CBN tools can also be due to diffusion wear. Nabhani (2001) compared the performance of PCD and polycrystalline CBN (PCBN) with that of coated tungsten carbide tool when machining titanium alloys. Diffusion and dissolution exacerbated by higher cutting temperature was observed as the predominant tool wear mechanism, and the tool wear probably resulted from a combination of dissolution/diffusion and attrition processes. The author also found that failure of carbide tools was the result of plastic deformation under compressive stress in the presence of high temperatures generated close to the cutting edge. For PCBN tools, they found that wear mechanisms were the same as those for carbide tools except that PCBN tools had lower wear rate and better machined surfaces had been achieved. Based on the experimental results, Nabhani

(2001) concluded that PCD tools had the lowest wear rate and produced the best workpiece surface quality. The PCD cutter was also used to investigate the possibility of finishing milling of titanium of titanium alloy Ti-6Al-4V by Kuljanic et al. (1998). . In their study, no crater or flank wear land was observed during machining, and the tool life of PCD cutters is very long. Their explanation for such a long tool life is that the formation of a titanium carbide film from a reaction between the work material and tool material on the diamond tool surface. This titanium carbide film can protect the tool, particularly the crater wear of the cutter. Bhaumik et al. (1995) used wBN-cBN composite tools to machine Ti-6Al-4V and investigated the wear mechanism of this type of tool. Based on X-ray dot mapping of compositional analysis, they indicated that titanium existed on the crater area, and when the adherent materials were taken away, accelerated attrition was observed on both the rake and flank faces. Some tests were also carried out with K20 grade cemented carbide tools, and the tool life of cemented carbide tools was found to be limited by rapid cratering on the rake face and deformation at the tool nose. König and Neises (1993) used PCD tools for turning Ti-6Al-4V, and they found that the diffusion and dissolution processes are exacerbated by the high local temperature resulting from the poor thermal conductivity of the workpiece material.

Both PCD and CBN are currently very expensive. In addition, PCD are highly reactive with titanium alloys at high temperature and CBN cannot maintain its hardness and strength at higher temperature, and consequently they are not suitable for machining titanium alloys at higher speed. Therefore, there is still a need to investigate the cutting mechanism for high-speed milling of titanium alloys with new type of tools.

Zareena (2002) had done many experiments to show that the BCBN (binder-less CBN) inserts have a remarkably longer tool life than conventional CBN inserts under all cutting conditions (up to 400m/min). The prominent reason for tool wear of CBN was found to be the result of the chemical reactivity between titanium and the binder element cobalt from the CBN tools. Similar to CBN tools, PCD tools were also observed to undergo diffusion-dissolution wear. The high thermal conductivity of BCBN tools and the absence of binder element were found to be the reason for longer tool life for these tools. In this thesis, BCBN inserts are also selected as cutting tools.

2.2 Geometrical models for milling processes

In milling process, each tooth of the cutter produce chips with variable thickness. Unlike turning processes, in milling the instantaneous chip thickness (h) varies periodically as a function of time-varying immersion angle φ. One of the key issues in the geometrical models for milling processes is to model the undeformed chip thickness. The often used circular geometry of chip formation in milling is shown in Figure 2.1, where f_z is the feed rate (mm/rev-tooth), φ is the instantaneous angle of immersion and $R = D/2$, where D is the diameter of the cutter.

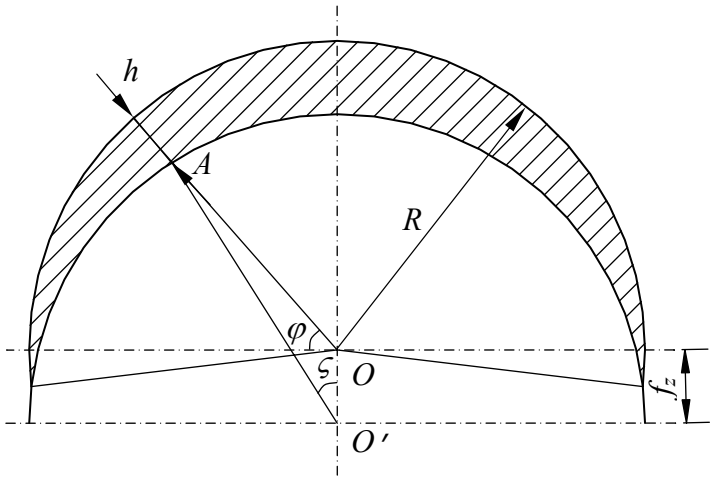


Figure 2.1 Geometry of chip thickness of the milling process

The instantaneous chip thickness $h(\varphi)$ at the position defined by angle φ is given by:

$$h(\varphi) = R - \overline{AO} = R - \sqrt{(R - f_z \sin\varphi)^2 - f_z^2 \cos^2\varphi} \qquad (2.1)$$

if f_z/R is very small, say less than 0.01, the undeformed chip thickness can be approximated as:

$$h(\varphi) = f_z \sin\varphi \qquad (2.2)$$

Although the circular tooth-path approximation makes it easy to establish a model of the milling process, it also affects the accuracy of the model. For high speed milling or micro-milling, there is a great need for higher accuracy, so the circular tooth-path could not meet the requirement.

More than half century ago, Martellotti (1941, 1945) derived more accurate equations for the cutting point trajectories in up and down milling. At that time without any help of modern computing facilities, it was very difficult for Martellotti to carry out this type of comprehensive work. But because complicated mathematical analysis was involved, Martellotti's undeformed chip thickness models have rarely been used for establishing the model of undeformed chip thickness. Recently, further research works have been carried out on modeling the milling process geometry. Montgomery and Altintas (1991) proposed a method to determine cutting forces in five distinct regions where the cutting edge travels during dynamic milling. Trochoidal motion of the milling cutter was used in determining uncut chip thickness. Spiewak (1995) proposed comprehensive models to integrate and expand essential features of previously developed analytical and numerical models. Some other researchers also investigated the undeformed chip thickness of milling processes; a detailed review is referred to the

recent work of Li (2001). However, an analytical solution to the undeformed chip thickness with higher accuracy is still a necessity with the help of modern software.

2.3 Cutting force models for machining processes

The implementation of high-speed milling is affected by various factors, such as the cutting force, the cutting temperature and the cutting power etc. Among them, the cutting force is an important parameter, as it relates to the deflection of the cutter, tool breakage as well as basic data for estimation of chatter vibration and machining error. The generation of mathematical models of high-speed milling is essential for deeper understanding of this advanced process. Therefore, an analytical model is needed to be established to predict the cutting forces for high-speed milling of Ti-6Al-4V.

According to the comprehensive survey conducted by the CIRP's working group (van Luttervelt et al., 1998), over 55 major groups are involved in modeling effort. Over 43% of the research groups were active in experimental/empirical modeling followed by 32% involved in analytical modeling and 18% involved in numerical modeling. Analytical and numerical models are more accurate and more applicable, and many researchers have developed numerous models of these two types for the machining operations.

2.3.1 Analytical models

According to the methods used, analytical models can be divided into two main categories: mechanistic models and shear plane/zone models.

Mechanistic models are based on the relationship between cutting forces and the undeformed area of cut, cutting tool geometry, cutting conditions and workpiece geometry. The cutting force is usually assumed to be proportional to the undeformed area of cut. For example, the tangential cutting force of milling process can be predicted using the following equation:

$$F_t = K_t\, b\, h \tag{2.1}$$

where K_t is the specific cutting pressure, b is the width of cut and h is the thickness of the cut. Basically based on the experimental results, K_t can be obtained; then it can be used to predict the cutting force very quickly for a set of fixed combination of cutting tools and the workpiece. This mechanistic approach works without knowing cutting forces mechanics parameters such as shear angle, shear stress and friction angle, so it is widely used. However, these models are commonly computer-based and depend heavily on empirical cutting data for their modeling capacity (van Luttervelt et al., 1998). So strictly speaking, this type of mechanistic models is not purely an analytical model, because it heavily depends on empirical cutting data.

The shear plane model was developed based on Merchant's shear plane theory (Merchant, 1944, 1945). This model is based on the assumption of continuous chip formation in a narrow zone which is idealized as a plane with uniform distributed shear stress. This method has been used quite successfully in the prediction of forces in several practical machining operations (Armarego, 1995; Altintas, 1995).

For the shear plane model stated above, a major problem is regarding the uncertain magnitude of tool-chip friction and the shear stress at the shear plane (van Luttervelt et al., 1998). In addition, this model is based on the assumption that workpiece material

deforms at constant flow stress. Considering the dependence of flow stress of metal on strain, strain rate and temperature, Oxley (1989) developed a more effective model, which considered the variation of flow stress properties in terms of the strain, strain-rate and temperature. This model assumes a thin shear zone, chip equilibrium and uniform shear stress in the secondary deformation zone at the tool-chip interface. Oxley's predictive machining theory is widely used for predicting many machining characteristic factors, such as the shear angle, cutting forces, flow stress etc. However, one of the major assumptions in Oxley's predictive machining theory is that the tool is perfectly sharp. In practice, it is impossible for the cutting tools to be perfectly sharp. In practice, the nose radius of the tool can improve the quality of finished surface; in addition, it can also improve the strength and wear characteristics of the tool. In this study, the axial depth of cut is smaller than the nose radius of the cutting tool, so Oxely's machining theory might need some revision before it is applied to this new specific case.

Recently, some other researchers also considered the temperature effects on the cutting forces, and then developed analytical models for the machining processes. Moufki et al. (1998) established a model for orthogonal cutting, which combined a thermo-mechanical analysis of the material flow within the primary shear zone and a modeling of friction at the tool-chip interface. In their model, they considered the primary shear zone is considered as a thin straight band. In addition, they assumed that tool-chip contact was described by use of a mean friction coefficient, which was a decreasing function of the temperature. Jacobus et al. (2001) proposed a thermo-mechanical model to predict the full in-plane biaxial residual stress profiles at and beneath the newly-generated surface from the turning process. They claimed that thermal effects

lead to the development of tensile stresses in the surface/near-surface layer and supports the presence of thermal and mechanical effects on residual stress that are consistent with the experimental data. Huang and Liang (2003) established a model to investigate the effect of tool thermal property on cutting forces. By thermal modeling of both the primary and secondary heat sources, they modified Oxley's predictive machining theory to analyze the metal cutting behaviors. Temperature distributions along the primary and secondary shear zones are modeled with the moving heat source method. To generalize the modeling approach, the modified Johnson–Cook equation is applied in the modified Oxley's approach to represent the workpiece material properties as the function of strain, strain rate, and temperature. Prediction results from their study showed that the Johnson–Cook equation works well as the material constitutive equation.

2.3.2 Numerical models

In numerical modeling, finite element method (FEM) techniques were found to be the most dominant tool (van Luttervelt et al., 1998). In this approach, the solution region is first divided into many smaller elements, so that various tool geometry, cutting conditions, and more sophisticated material and friction models can be incorporated (Altintas, 2000). Then, element equations are formulated. Based on the interconnected relation of elements, element equations can be assembled into the global equations. Finally, after solving the global equations, numerical solution to the problem domain is obtained. Two basic approaches are often used to solve the global equations, which are the Newton Raphson method and the direct iteration. Thus, the main advantage of FEM is its ability to predict chip flow, cutting forces, and especially a distribution of tool temperatures and stresses for various cutting conditions by simply changing the

input data. Ozel and Altan (1998; 2000) have done much definitive works in this field. They developed a predictive model for high-speed milling based on FEM simulations. Using their model, the resultant cutting forces, flow stresses and temperatures in turning and flat end milling were predicted primarily. More importantly, with less number of experiments, this method is able to estimate the variations of flow stress and friction conditions of high speed machining. In their model, the tooth-path was assumed to be circular; however, this approximation will cause some error for high-speed milling.

Using the commercial code FORGE2, Ng et al (1999) presented an FE model to simulate orthogonal machining of hardened die steel with advanced ceramic tools. Unfortunately, their model underestimated the magnitude of cutting force due to limited data on the sensitivity of the workpiece material to strain hardening and the strain-rate sensitivity at elevated temperature, and oversimplification of the frictional conditions at the tool/chip interface. Based on Oxley's theory, Carrino et al. (2003) used a coupled thermo-mechanical finite element model to simulate orthogonal cutting of carbon steel C40. In their model, the tangential force applied along the tool/chip interface was assumed to be a fraction of the shear stress of the material. Good agreement between experimental and numerical results was found based on cutting forces measurements.

Altintas (2000) and his research group at UBC developed an Arbitrary Lagrangian-Eulerian (ALE) formulation, which has been applied for the prediction of cutting variables in machining. In the developed ALE code, the effects of edge radius on the cutting edge on the cutting forces were considered.

Although rapid progress and better results of FEM have been achieved recently, there are also some problems for FEM which need to be considered. The most significant problem is to obtain the material properties under the practical cutting conditions. Nowadays, the data of material properties used in the simulation of FEM are obtained in tensile or compression tests. Obviously the real cutting conditions are different from those of tensile or compression tests. In addition, the numerical model requires significant amounts of computation time. The computational burden is almost unbearable for 3-D modeling. It is still a long way to go before the finite element method can be used to simulate practical machining operations with an acceptable degree of accuracy and reliability and an acceptable amount of effort for daily use (von Luttervelt et al., 1998).

2.4 An overview of often used optimization methods

In this section, four types of often used optimization methods have been reviewed, the details are given as follows.

2.4.1 Dynamic programming

Dynamic programming was developed by Bellman (1957), who described the way of solving problems where the best decisions need to be found one after another. Since its first development, dynamic programming has been used for solving larger number of sequential, or multi-stage, decision problems.

For the dynamic programming approach, difficult to solve or complicated problems are decomposed into equivalent smaller problems that are much easier to solve. For example, when dynamic programming is used to solve a multi-variable problem, firstly this multi-variable problem is decomposed into a series of single variable problems;

then it can be solved as a series of single variable problems. The decomposition is achieved by tandem projection onto the space of each of the variables. In other words, the variables are projected first onto a subset of the variables, then onto a subset of these, and so on. One of the key shortcomings in the existing dynamic programming is the large requirement of high computational effort for large systems over a wide range of values.

2.4.2 Geometric programming

Clarence M. Zener (1961), Director of Science at Westinghouse, published the first of several papers on a new optimization technique he had discovered while working on the optimal design of transformers. Zener used the result called Cauchy's arithmetic-geometric inequality to show that the arithmetic mean of a group of terms is always greater than or equal to the geometric mean of the group. He was able to convert the optimization of a nonlinear economic model for transformer design to one of solving a set of linear algebraic equations to obtain the optimum. The use of Cauchy's arithmetic-geometric inequality led to the name of geometric programming for the technique.

GP is a useful method to be used for solving nonlinear problems subject to nonlinear constraints. It is more powerful than other mathematical optimization techniques when the problem is restricted by one or two constraints. Petropoulos (1973), Root and Ragsdell (1976), and Somlo and Nagy (1977) have successfully used GP to optimize the machining problems. However, GP can only handle continuous variables. In addition, as the number of constraints increases in large-scale problems, the optimized

problem might be more complicated than the original one; so other optimization techniques should be employed in conjunction with GP (Somlo and Nagy, 1977).

2.4.3 Genetic algorithms

Genetic algorithm (GA) was developed by Holland (1975) and is a stochastic global search method that mimics the metaphor of natural biological evolution. GA has been applied successfully to enormous number of problems in business, engineering and science. Good solutions can be found within reasonable amounts of time for many practical problems.

GA operates on a population of potential solutions applying the principle of survival of the fittest to produce (hopefully) better and better approximations to the final optimal solution. At each generation, a population of candidate solutions is maintained. To generate a new population, candidate solutions are randomly paired. For each pair of solutions, a crossover operator is first applied with a moderate probability (crossover rate) to generate two new solutions. Each new solution is then modified using a mutation operator with a small probability (mutation rate). The resulting two new solutions replace their parents in the old population to form a temporary new population. Each solution in the temporary population is ranked against other solutions based on their fitness function values. A lottery process is then used to determine a new population identical in size to the previous population, such that higher ranked candidates are allowed to have higher probability to exist in the new population. GA iterates over a large number of generations and, in general, as the algorithm executes, solutions in the population become more fit, resulting in better candidate solutions. The pseudo code in Fig. 2.2 gives an overview of a traditional GA.

```
begin
    t = 0
    initialize P(t)
    evaluate P(t)
    while (not termination-condition) do
    begin
        t = t + 1
        select P(t) from P(t-1)
        reproduce P(t)
        mutate P(t)
        evaluate P(t)
    end
end
```

Figure 2.2 Pseudo code of the simple genetic algorithm

It is important to note that GA provides a number of potential solutions to a given problem and the choice of final solution is left to the user. At this point, GA differs substantially from other traditional search and optimization methods. The most significant differences are (Chipperfield and Fleming, 1996):

- GA searches a population of solutions in parallel, not a single solution
- GA uses probabilistic transition rules, not deterministic ones.
- GA works on an encoding of the parameter set rather than the parameter set itself (expected in where real-valued individuals are used).
- GA does not require derivative information or other auxiliary knowledge; only the objective function and corresponding fitness levels influence the directions of search.

In addition, GA is a search strategy ideally suited to parallel computing. Parallelism means that the search effort is distributed simultaneously along many hyper-planes (regions) of the space.

Although GA has been used to solve numerous practical problems, the successful application of GA depends on the population size or the diversity of individual solutions in the search space. If GA cannot hold its diversity well before the global optimum is reached, it may be difficult for it to find the global optimum, and sometimes it even results in the premature convergence to the local optimum. Although maintaining diversity is the predominant concern, it also results in deterioration in the performance of GA. There is a great need to find a trade-off between the population diversity and performance of GA (exploration and exploitation).

In addition, GA is also sensitive to parameters (Ingber and Rosen, 1992). Adjustments in population size, crossover probability, mutation rate, and other parameters can influence GA's behavior and lead to significantly superior (or inferior) results. Such adjustments are problem dependent and are best accomplished through trial and error. The difficulty and uncertainty of finding optimal parameters is a significant problem for GA.

2.4.4 Simulated annealing

Simulated annealing is a very general optimization method, which stochastically simulates the physical annealing process. About half century ago, Metropolis et al. (1953) first introduced an algorithm to simulate a collection of atoms at a given

temperature, which is also called Metropolis algorithm. The term simulated annealing derives from the roughly analogous physical process of heating and then slowly cooling a substance to obtain a strong crystalline structure. The Metropolis algorithm could be used for optimization problems, where the minima of the objective function correspond to the ground state of the substance. The details of Metropolis algorithm can be described briefly as follows.

To apply simulated annealing, the system is initialized with a particular configuration. A new configuration is constructed by imposing a random displacement. SA algorithm approaches optimization problems by randomly generating a candidate solution and, then making successive random modifications. If the new generated solution is found to have a better fitness than its predecessor, it is retained and the previous solution is discarded. If the modified solution is a little fitter than its predecessor, it is still retained with probability directly related to the current temperature. As the execution of the algorithm continues and the temperature becomes cooler, it becomes less likely that unfavorable solutions are accepted. By using this approach, it is possible for an SA algorithm to move out of local minima during early execution due to the probabilistic acceptance of some upward moves. During the algorithm's execution, when the temperature is decreased to a small value, it is more likely that good solutions will not be discarded so that an optimal solution can be reached. The execution diagram of SA is shown as Fig. 2.3.

With sufficiently slow cooling and stationary distribution of solutions, SA possesses a formal process of global convergence (Mahfoud and Goldberg, 1995). SA algorithms are conceptually simple, however, finding optimal parameters for SA, i.e., initial

temperature T, cooling rate α (which is used to update the temperature at fixed intervals), etc., is by no means simple or straightforward. Firstly, setting parameters for SA is problem dependent, and it is best accomplished by trial and error. Furthermore, previous studies (Ingber and Rosen, 1992; Chen et al., 1998) with SA have demonstrated that SA algorithms are very sensitive to parameters, and their performances are largely dependent on fine tuning of the parameters. The problem dependent nature of setting parameters and sensitivity to parameters limit the effectiveness and robustness of SA. There is a single solution that is modified over time, so the SA algorithm is naturally serial and difficult to implement on parallel systems with appreciable speed-up.

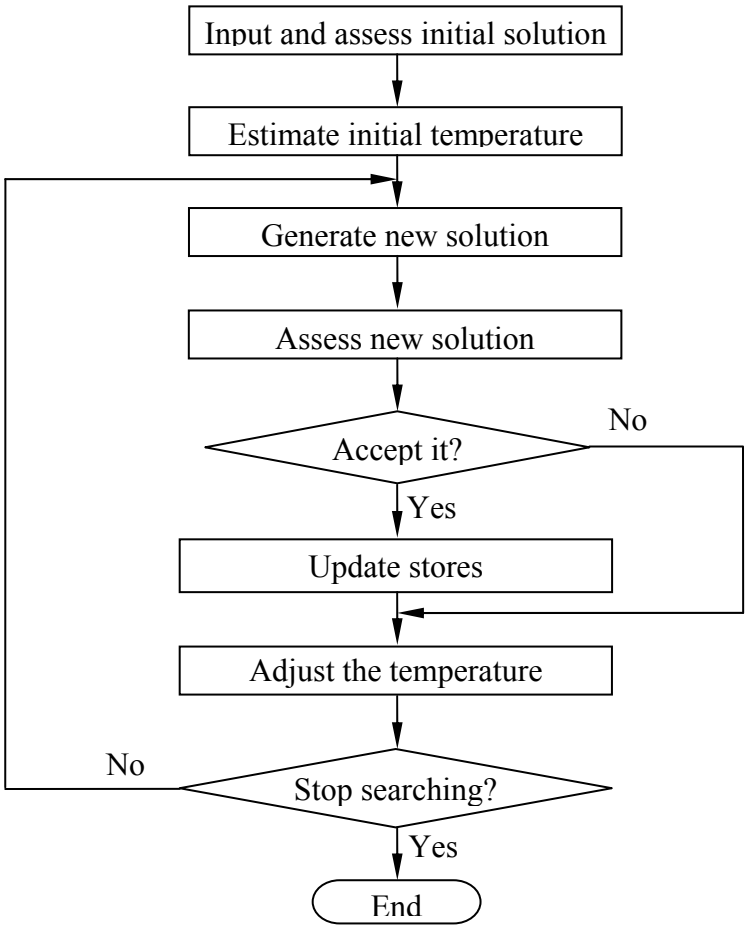


Figure 2.3 Diagrammatic structure of the algorithm simulated annealing

2.4.5 Overview of hybrid of GA and SA

As stated above, GA and SA are both independently valid approaches toward problem solving with certain strengths and weaknesses. GA can begin with a population of solutions in parallel. However, in order to find the global solution, higher population diversity is needed, but higher population diversity reduces the performance of GA significantly. On the other side, SA has better convergence properties, but it cannot easily exploit parallelization. Therefore, there is a need to use strengths of these two algorithms and to overcome their weaknesses.

An alternative approach is to combine GA with SA to develop a new method. Recently many researchers tried to combine GA and SA to provide a more powerful optimization method that has both good global convergence control and efficient parallelization.

Chen and Flan (1994) had shown that the hybrid of GA and SA can perform better for nine difficult optimization problems than either GA or SA independently. Sirag and Weisser (1987) proposed a unified thermodynamic genetic operator to solve ordering problems. The unified operator introduced the temperature cooling concept of SA to the conventional GA operation of crossover and mutation to yield offspring. This operator can ensure greater population diversity at high temperature and less population diversity at low temperature. The greater convergence control from the addition of a SA-like temperature is empirically shown to increase performance of GA. Brown et al. (1989) developed a parallel GA/SA hybrid for solving the quadratic assignment problem (QAP). In their hybrid method, GA used low-temperature annealing as a selection procedure. For each generation of GA, selection and crossover

are followed by a full schedule of low-temperature SA. Mutation was then used to modify the elements of individuals in the temporary population. Empirical results indicated that their algorithm outperformed other common QAP methods. Mahfoud and Goldberg (1995) also introduced a GA and SA hybrid. Their hybrid runs SA procedures in parallel, which uses mutation as the SA neighborhood operator and incorporates crossover to reconcile solutions across the processors. Two variations of this approach have shown asymptotical convergence as in SA, and the algorithm is empirically shown to regularly converge to the global optimum for two GA-deceptive problems. Varanelli and Cohoon (1995) proposed a similar hybrid method of GA and SA, and their proposed parallel version more closely resembled SA controlling a number of distinct GAs running in parallel. In addition, Chen et al. (1998) also proposed a hybrid method, which maintains one solution per Processing Element (PE). Each PE accepts a visiting solution from other PEs for crossover and mutation. For the selection process, the SA cooling schedule and system temperature were used to decide whether the new generated individual was accepted or not. In this method, they used the local selection of SA to replace the conventional selection process of GA. Recently, Hiroyasu et al. (2000) proposed an algorithm involving several processes. In each process SA is employed. The genetic crossover is used to exchange information between solutions at fixed intervals. Based on parallel simulated annealing in (Mahfoud and Goldberg, 1995; Hiroyasu et al., 2000), Baydar (2002) developed a parallel simulated annealing algorithm which also uses the survival of the fittest method, and acceptable results were found with this algorithm.

Each of the above approaches has its own strengths, because some good characteristics of GA and SA are maintained when combining GA and SA together. In this study, a new tightly coupled hybrid of GA and SA is presented.

2.4.6 Overview of parallelization of GA

GA has been applied successfully to solve numerous problems in reasonable amount of computation time. However, for problems with a large search space and costly function evaluation, it may require GA to take days or even months to find an acceptable solution. Therefore, a faster GA is needed. There have been several efforts to make GA run faster. And a most promising alternative is to parallelize GA.

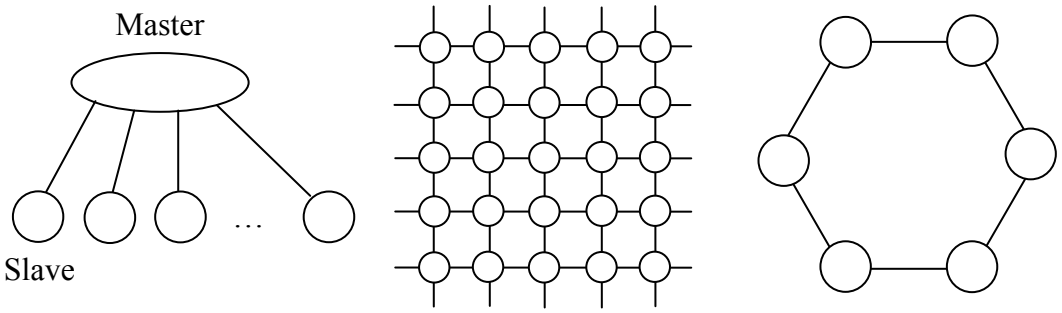
There are several ways to parallelize GA (Gordon and Whitley, 1993), and the parallel GA (PGA) has been developed and successfully applied to optimize practical problems (Alba and Troya, 1999). According to the nature of the population structure and recombination mechanisms used, PGA can be classified into four categories: single-population master-slave PGA, coarse-grained PGA, fine-grained PGA and hierarchical hybrids of PGA (Cantú-Paz, 2000).

In single-population master-slave GA as shown in Fig. 2.4 (a), there is a single population. The operators of GA (selection, crossover and mutation) are implemented on the master processor, and the evaluation of fitness function is distributed among several slave processors.

The fine-grained parallel GA treats each individual as a separate breeding unit, and the individuals may mate with those selected from a small local neighborhood, as shown in Fig. 2.4 (b). Since the neighborhoods overlap, fit individuals will migrate through the population.

The coarse-grained parallel GA is very popular and widely used. In a coarse-grained PGA as shown in Fig. 2.4 (c), the entire population is divided into several

subpopulations. Each subpopulation runs a conventional GA independently and concurrently on its own subpopulation. After several epochs, best individuals migrate from one subpopulation to another according to a migration topology.



(a) Master-slave model (b) Fine-grained model (c) Coarse-grained model

Figure 2.4 Different models of parallel genetic algorithms

The hierarchical PGA combines coarse-grained PGA with master-slave or fine-grained PGA, so that it has the benefits of its components (Cantú-Paz, 2000). In this thesis, a master-slave/coarse-grained PGA, which combines master-slave PGA and coarse-grained PGA together, has been used to parallelize the serial algorithm.

With the use of LINUX clusters and the workstation network, the cost of parallel computing hardware has significantly reduced. And now it has been feasible to move complicated engineering optimization problems to high performance parallel computing environment. In this study, the parallel algorithms are presented and implemented on the SUN workstation network.

2.5 An overview of optimization of milling process

Many efforts have been made to optimize machining parameters. However, most work done on the optimization of cutting conditions in machining is mainly restricted to turning operations (rotational parts), although milling operations play important roles

in the present manufacturing industry with considerably greater number of machining centers employed.

Although the importance of using optimal cutting parameters was identified in the early 1900s, advances in the development of optimization strategies have been very slow (Sönmez et al., 1999), since the problem is too complex due to the nonlinear dependence of machining variables. In the efforts made by researchers, several types of methods have been used for the optimization of cutting parameters.

Direct search method is one of the most popular mathematical optimization methods. Direct techniques are based on the derivative method in which the gradient of an objective function is calculated and set to zero to obtain points at which its value becomes maximum or minimum (Gray et al., 1997; Duffuaa et al., 1993). The second derivative of the function is then used to determine the maxima or minima. Clearly, therefore, the objective function in classical and direct optimization must be continuous and twice differentiable. However, this requirement is generally not met in real-world problems. Tolouei-Rad and Bidhendi (1997), Wang (1998), Wang and Armarego (2001), Wang et al (2002), Armarego et al. (1993), and Kilic et al. (1993) used graphical techniques for the optimization of machining conditions by mapping the relevant constraints and objective functions in the planes. Armarego et al. (1993) and Kee (1996) used the constrained optimization strategies for selecting the optimal cutting conditions. Kayacan et al., (1996) and Agapiou (1992) used knowledge base and machining topology; Jang (1992) presented a unified optimization approach for the selection of the machining parameters. Wang (1993) used a neural network based approach to optimize cutting parameters. However optimization by these methods

often ends in local optimum or fails to converge on a result. Sönmez et al., (1999), Ermer (1971), Agapiou (1992) and Shin and Joo (1992) used the dynamic programming (DP) optimization method. DP can solve both continuous and discrete variables and yield a global optimal solution. However, if the optimization problem involves a large amount of independent parameters with a wide range of values such as the cutting parameters, the use of DP is limited. In addition, the geometric programming (GP) method was used for optimization (Sönmez et al., 1999; Petropoulos, 1973; and Jha, 1990). Jha (1990) has attempted to optimize cutting parameters in milling by GP and concluded that the GP-based program was very slow for producing good results. Dereli et al. (2001) indicated that it needed long execution times for good scores of the objective function by Sönmez et al. (1999), who developed a system for the constrained optimization of cutting parameters to be used in the multi-pass plain and face milling operations using DP plus GP. The above mentioned optimization techniques either tend to result in the local minima or take a long time to converge on a reasonable result. In order to overcome the long execution time using DP plus GP, Dereli et al. (2001) used genetic algorithm (GA) to develop an optimization system that had better performance.

Recently, two new approaches to optimization have been developed independently: GA and SA. Both of these algorithms are probabilistic search algorithms that are capable of finding globally optimal results to complicated optimization problems (Lee and Choi 1995). For application in machining processes, besides Dereli et al. (2001), Shunmugam et al. (2000) used GA to optimize the multi-pass face-milling and obtained optimal cutting parameters. Liu and Wang (1999) improved the convergence speed of traditional GA and obtained good results by defining and changing the

operating domain of GA. However, for above optimization studies with GA, the two-stage methodology is needed for multi-pass machining processes. Firstly the number of passes or depth of cut are given; then in the second stage GA was used to find the optimal cutting parameters for a given depth of cut. A hybrid procedure based on the simulated annealing algorithm and the Hooke-Jeeves pattern search (SA/PS) has been used to optimize the machining process by Chen and Tsai (1995). They found that the SA/PS algorithm can obtain a near-optimal solution within a reasonable computation time. Juan et al. (2003) applied SA to search optimal cutting parameters of high-speed milling of SKD61 tool steels. Their optimization is for the rough machining of HSM, and only one cutting passes was considered. So when SA or SA/PS is used to optimize multi-pass milling processes, the two-stage methodology as that for GA discussed above is also needed.

The main issue for optimization of engineering problems is how to handle the constraints. Although it is easy for GA to construct feasible solutions for some engineering optimization problems, for some problems with complicated constraints, generation of feasible solutions that fulfill the constraints is itself a difficult problem. Normally, three approaches for GA have been used to handle problem constraints: feasible individuals, repair algorithm approach and the penalty function (Michalewicz, 1996). In the first approach, solutions that violate a constraint(s) are infeasible and therefore are discarded. This approach is impractical because many problems are tightly constrained and finding a feasible solution may be almost as difficult as finding the optimal one. In addition, infeasible solutions often contain valuable information and should not be discarded outright. In the second one, special repair algorithms are used to “correct” any infeasible solutions. For most cases, such repair algorithms might

be computationally intensive to run and the resulting algorithm must be tailored to the particular application (Michalewicz, 1996). In the third method, a penalty term is incorporated into the fitness function to penalize solutions that violate the constraints. The idea is to degrade the fitness of the infeasible solutions but not throw away valuable information contained in the cost term of the fitness function. In this study, this penalty function approach has been used to handle the constraints.

2.6 Concluding marks

As mentioned above, there has been much research on the cutting performance of high-speed machining of titanium alloys. Many analytical and numerical models have been proposed by researchers to predict the process parameters, such as cutting forces, shear stress, etc. Although titanium alloys are widely used in aerospace industry and automobile industry, predictive cutting force models about high-speed milling of titanium received little attention. Therefore, a model is needed to predict the cutting forces for high-speed milling of Ti-6Al-4V with advanced cutting tools. At the same time, the circular tooth path does not accurately describe the geometry of chip formation for high-speed milling. Thus, a new method for modeling of the milling process geometry is required to get accurate undeformed chip thickness.

GA and SA have been used to find globally optimal solutions to optimization problems, but they also have certain weaknesses. Many researchers tried to hybridize GA and SA to maintain the strengths of GA and SA, and avoid their weaknesses. For practical engineering optimization problems, faster and powerful optimization methods are needed to ensure the global convergence to optimal solutions. Parallelization of the serial algorithm is one of the most promising methods to improve its efficiency. There

are four models to parallelize GA. The hierarchical PGA model is a good choice to parallelize, because it has the benefits of its components.

Due to the poor machinability of Ti-6Al-4V, it is also very crucial to choose the optimal machining conditions and parameters. For a meaningful optimization of machining process, it is necessary to satisfy certain number of constraints for the optimization model, such as power constraint, cutting forces constraint and finishing surface roughness, etc. For the optimization of engineering problems, penalty function is an effective approach to handle the constraints. Under the practical constraints, the optimization methods can be used to obtain the optimal and applicable solutions.

Chapter 3
Experiment details

3.1 Introduction

In this study, many experiments have been carried out to investigate the cutting performance of high-speed milling of Ti-6Al-4V. And all experiments were conducted in the Advanced Manufacturing Lab (AML), National University of Singapore. Details of all experiments and the experiment methods are described in this chapter.

3.2 Experimental setup

The general description of the experimental setup is shown in Fig. 3.1. The main parts of the setup include the machine tool and cutting tools, workpiece material and the measurement system. The bar-shaped workpiece was mounted on a KISTLER 9265B dynamometer to measure the cutting forces in three directions. The force signals were subsequently amplified by a KISTLER charge amplifier and then captured and stored in a computer via an analogue/digital interface. In the following sections, the full details of this four components are given.

3.2.1 Machine tool

All the experimental tests were carried out on a 3-axis Makino vertical machining center, and its specifications are as follows:

Cutting feed rate 1~50,000mm/min

Least input increment of feed 0.001mm

Spindle speed range 15~30,000rpm.

Resolution in X, Y, Z axis 0.001mm

Spindle power (25%ED/continuous) 22/18.5kW

Power supply 55kVA (maximum power consumption)

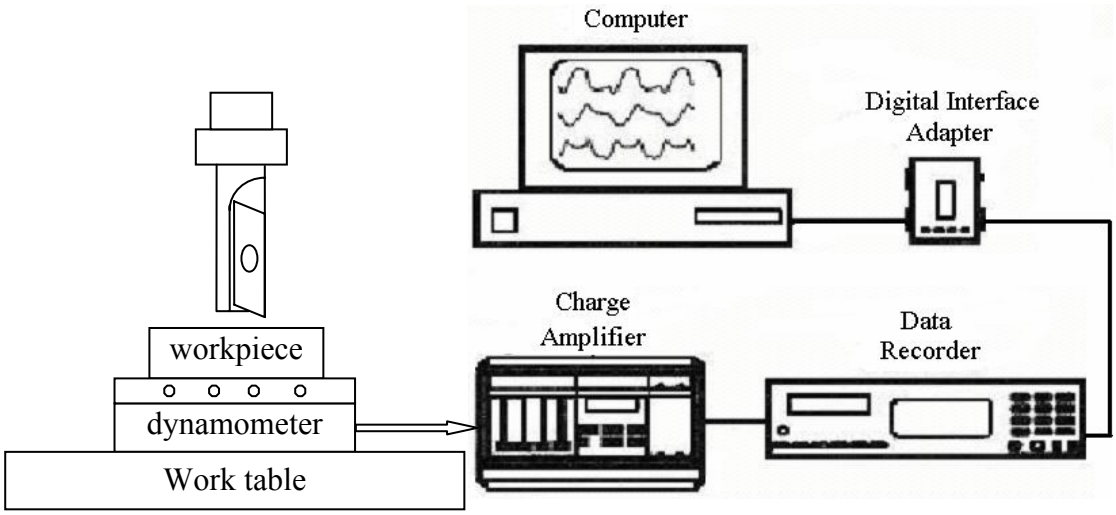


Figure 3.1 Block diagram of the experimental setup

3.2.2 Cutter material

The HSK A-63 tool shank was used to hold the cutter CHE 2020R. The cutter was made by Sumitomo Electrical Industries, Japan. The specifications of the cutter are as follows:

Material of the holder: SCM Japanese standard;

Modulus of elasticity: 211GPa;

Modulus of rigidity: 82.4GPa;

Length of the holder: 100mm;

Diameter of the holder: 18mm;

3.2.3 Insert material

The high-purity binderless CBN (BCBN) inserts were used in this study, whose ISO Catalog No is TEGN 110308 (Sumitomo Electric Industries, 2000). Throughout the experiments only one active insert was used in the two teeth cutter and a dummy insert was used for balancing. The binderless cubic boron nitride sintered product contains neither a binder nor a sintering agent or catalyst. During the production process of BCBN, the raw material of hBN is completely converted to a cubic phase at a high temperature and under an extremely high pressure. Then, the converted particles of CBN are bonded together, keeping the particle size extremely fine. The mechanical properties of BCBN and CBN are shown in Table 3.1 (Uesaka and Sumiya, 1999).

Table 3.1 Mechanical and thermal properties of CBN and BCBN

Properties		BCBN	CBN	Cemented carbide 6% Co	Ceramic Si_3N_4 Hot-press
CBN contents (vol%)		>99.9	85 – 90	—	—
CBN grain size (micron)		<0.5	1 – 3	—	—
Other constituents		-(comp. hBN)	Binder (Co etc.)	—	—
Process		Direct conv.	CBN + Binder	—	—
Hardness (GPa)	room temperature	50 – 55	35 – 40	15.69	21.58
	1000°C	20	12	—	—
TRS* (GPa)	room temperature	1.35	1.40	2.20	0.793
	1000°C	1.60	0.55	—	—
Thermal conductivity (W/mK)		360 – 400	100 – 130	95	29
Thermal stability		1620	1270	—	—

* Transverse rupture strength

The properties of cemented carbide and ceramic materials (Sandvik hard materials, 2001) were also cited for highlighting the hardness of BCBN material as shown in Table 3.1. From Table 3.1, it can be seen that cement carbide (6% of Co) has higher

TRS value, but its hardness value is smaller. On the contrary, BCBN has the higher hardness and less TRS values. Normally, making hardness of a material higher also reduces its TRS. Although the TRS value of the binderless CBN material is slightly less than that of CBN containing binder material at room temperature, it has much higher hardness and thermal conductivity than CBN. More importantly, at 1000^{o}C, the BCBN material has higher hardness and larger TRS values than the corresponding CBN tool material.

The BCBN tools are thought to exhibit excellent mechanical properties and superior thermal stability because the sintered body contains no secondary phases and consists of extremely fine CBN particles (Sumiya et al., 2000). Therefore, this type of tools shows great promise for high-speed milling of difficult-to-cut materials. The main advantageous features of the binderless CBN sintered product are as follows (SEI news, 2000):

- It has high thermal conductivity as well as outstanding resistance to heat and thermal shock because it is a sintered product made of single-phase CBN. Thus, it minimizes the possibility of thermal cracks and chipping occurring at the edge.
- It is suitable for interrupted cutting, having superior mechanical characteristics, such as hardness and strength, because it is made of fine particles not larger than 0.5 microns that are solidly bonded to one another without a binder or a sintering agent/catalyst in the grain boundaries.
- It can be cut and brazed to create desired shapes for turning and milling, as well as the manufacturing of special brazing tools.

The cutting tool's geometry is given as following:

Clearance angle (deg):	20
Axial Rake (deg):	1
Radial Rake (deg):	10
Nose radius of the cutter (mm):	0.80
Cutting edge radius (mm):	0.015
Diameter of the cutter (mm):	20

3.2.4 Workpiece materials

Titanium alloys have been widely used in the aerospace, biomedical, automotive and petroleum industries because of their good strength-to-weight ratio and superior corrosion resistance. However, it is very difficult to machine these alloys, because they have poor conduction and active alloying tendency.

Normally, titanium alloy structures can be classified according to the following scheme: α-phase structure, β-phase structure and (α+β)-phase structure. Commercial titanium alloys are divided into three classes (α-alloys, β-alloys and (α+β)-alloys) depending on whether they consist entirely of the α-phase, entirely of the β-phase or of a mixture of the two phases.

α-alloys contain α-phase structure, which is stabilized in titanium by the addition of the interstitial solid-solution forming elements oxygen, nitrogen and carbon and by a number of non-transitional metallic elements. The presence of these elements in titanium, even in small amounts, produces a marked increase in the strength of the metal (McQuillan and McQuillan, 1956). This type of alloys has excellent tensile

properties and creep stability at room temperature up to 300^{o}C, which is used mainly for corrosion resistance and cryogenic applications.

β-alloys contain significant qualities of β-stabilizers and are characterized by hard hardenability, improved forgeability and cold formability, as well as high density. Basically these alloys offer ambient temperature strength equivalent to that of (α+β)-alloys, but their elevated temperature properties are inferior to that of the (α+β)-alloys (Ezugwu and Wang, 1997).

(α+β)-alloys contain addition α and β stabilizers and they possess microstructure of the mixture of α and β structure. This type of alloys has an excellent combination of strength and ductility, therefore, it accounts for about 70% of all titanium alloys used.

Among all titanium alloys, Ti-6Al-4V is most widely used. Thus, it has been chosen as the workpiece material in this study, which specially is Ti-6Al-4V bar (ASTM B265 Grade 5). The dimensions of the Ti-6Al-4V test pieces are 40×50×100mm, and its nominal composition (wt.%) and mechanical properties are shown in Tables 3.2 and 3.3 (Zareena, 2002), respectively.

Table 3.2 Composition of Ti-6Al-4V

Content	C	Fe	N	O	Al	V	H	Ti
Composition(wt.%)	0.05	0.09	0.01	-	6.15	4.40	0.005	Balance

3.2.5 Measurement system

A Kistler quantz 3-component dynamometer (Type 9265B) was used to measure the cutting forces. The workpiece was mounted on the dynamometer, as shown in Fig. 3.2.

The measurement ranges are: -15~15kN for X and Y direction and 10~20kN for Z direction.

Table 3.3 Mechanical properties of Ti-6Al-4V

Tensile strength (MPa)	993
Yield strength(MPa)	830
Elongation (%)	14
Modulus of elasticity(GPa) Young's modulus	114
Poisson's ration	0.342
Emissivity (-)	0.65
Hardness (HRC)	36
Linear thermal expansion coefficient ($10^{-6}K^{-1}$)	8.6
Thermal Conductivity ($Wm^{-1}K^{-1}$)	7.2
Specific Heat Capacity ($J\ kg^{-1}K^{-1}$)	560
Electrical Resistivity (μΩm)	1.67

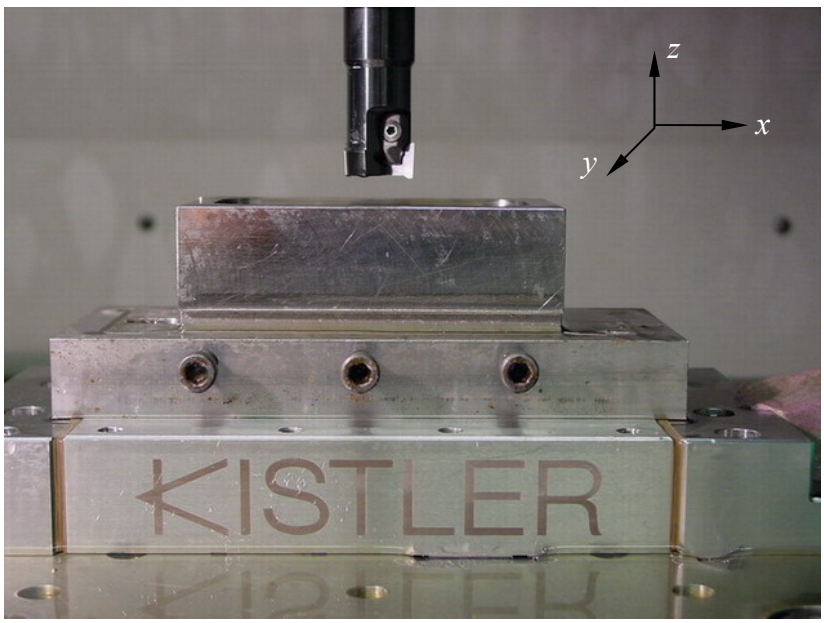


Figure 3.2 Kistler quartz 3-component platform dynamometer Type 9265B

The output signals from the dynamometer are amplified by a Kistler charge amplifier; and then passed through an analogue/data interface. Finally the digital cutting forces are displayed with a computer real-timely. At the same time, a Sony recorder (type PC 208 Ax) was used to record the cutting force signals in a Sony data cartridge of 1.3GB capacity. The sampling frequency of digitizing the analogue cutting force signals is 24 kHz. The digital data stored in the cartridge can be analyzed further with the support of PC Scan MK II data acquisition software.

3.2.6 Cutting fluids used in this study

The coolant used during experiments is 5% fully synthetic, transparent water based FUCHS coolant of grade ECOCOOL S/C05. The coolant was injected at a pressure of 40 bar into the cutting zone at a flow rate of 50 liters per minute through 8 nozzles arranged around the cutter.

3.3 Experimental design

Slot milling passes over a length of 100 mm were performed across the workpiece throughout the tests. After each pass, the conventional CBN tools were used to clear the slot and the edges of the workpiece to ensure a new surface for the next pass. Apart from measuring cutting forces, the tool flank and rake wear were recorded simultaneously, after each pass.

To ensure effective results, the design of experimental techniques was used to plan the machining experiments efficiently. Then, the data collected during the experiments was analyzed with the statistical approach. In recent years, much powerful software has been developed and used in data statistical analysis. With the help of such software,

it becomes easier and more convenient to analyze experimental data. So the main concern for experiments is the experimental design.

3.3.1 Experimental methods

Before conducting any experiments, the experimental method must be carefully chosen to accomplish experimental objectives. A well-designed experiment can substantially reduce the number of experiments. There are several classes of experimental designs. Fully crossed factorial designs are widely used for experimental design, which are efficient for studying effects of every control factor on the response variable. With the full factorial designs, in each complete trial or replication of the experiment all possible combinations of levels of the factors are investigated (Montgomery, 1997). Apart from factorial designs, response surface methodology is another widely used method for experimental designs.

Response surface methodology (RSM) is a combination of mathematical and statistical techniques used in an empirical study of model building and optimization. In many real applications, two-order response surface is required to be fitted. The central composite design is one of the best choices for fitting a second-order surface (Myers and Montgomery, 1995). The central composite design is built-up from the factorial designs by adding several center points. Therefore, by choosing the proper number of center points, the efficient estimation of the pure quadratic terms can be controlled. Based on the regression analysis of experimental results, a model of the response surface with independent input variables can be obtained. The commonly used mathematical models can be represented by:

$$y = f(x_1, x_2, x_3) + \varepsilon \tag{3.1}$$

where y is the response, f is the response function, x_1, x_2, x_3 are independent input variables and ε is error that is normally distributed about the observed response y with zero mean.

In this study, the responses are cutting forces or tool life. The experimental response Y (tool life or cutting forces) can be expressed in terms of depth of cut a, feed rate f and cutting speed v:

$$Y = Ca^{\alpha} f^{\beta} v^{\gamma} \varepsilon' \tag{3.2}$$

where C is a constant, α, β and γ are exponents and ε' is the experimental error.

Above non-linear model can be linearized by taking a logarithmic transformation on both sides of Eq.(3.2). Then the following equation is obtained:

$$\ln Y = \ln C + \alpha \ln a + \beta \ln f + \gamma \ln v + \ln \varepsilon' \tag{3.3}$$

The constants and exponents C, α, β and γ can be determined by the method of least squares. For Eq. (3.3), the first order linear model of RSM can be written as follows:

$$Y_1 = \ln Y - \ln \varepsilon' = b_0 x_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 \tag{3.4}$$

where Y_1 is the estimated response based on first-order equation, x_0 = 1(a dummy variable); x_1, x_2 and x_3 are the logarithmic transformation of a, f and v, respectively. b_0 = ln$C$, $b_1 = \alpha$, $b_2 = \beta$, $b_2 = \gamma$. So b_i (i = 1, 2, 3, 4) are the values to be estimated using experimental data.

The general second order polynomial response can be extended from Eq. (3.4) as:

$$\begin{aligned} Y_2 = \ln Y - \ln \varepsilon' &= b_0 x_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 \\ &+ b_4 x_1 x_2 + b_5 x_2 x_3 + b_6 x_1 x_3 + b_7 x_1^{\,2} + b_8 x_2^{\,2} + b_9 x_3^{\,2} \end{aligned} \tag{3.5}$$

where Y_2 is the estimated response based on second order equation. The parameters i.e., b_0, b_1, b_2, …, b_9 are the values to be estimated using experimental data by least-square

method. Validity of the model will have to be tested with the help of statistical test, such as F-test, Chi square test (Montgomery, 2001), etc. before it is used in the optimization process.

3.3.2 Experimental design for measuring cutting forces

For the cutting force measuring experiments, three dominant factors (depth of cut, feed rate and cutting speed) are considered in the planning of the experiments. The range of values of each factor was set at three levels, namely low, medium and high levels, as shown in Table 3.4, selected on the basis of present day manufacturing requirements. Based on this setting, a 3^3 full factorial design was used so that all the interactions between the independent variables could be investigated. Based on this, a total number of 27 experiments in high-speed milling were carried out, and each experiment has a combination of different levels of factors as shown in Table 3.5.

Table 3.4 Cutting parameters and their values

Variables		Values of different levels		
Designation	Description	Low(−)	Medium(0)	High(+)
a	Depth of cut (mm)	0.050	0.075	0.100
f	Feed rate (mm/r)	0.050	0.075	0.100
v	Cutting speed (m/min)	300	350	400

3.3.3 Experimental design for measuring tool life

In order to determine the tool life model, the orthogonal design has been developed which attempts to approximate the equation using the smallest number of experiments possible. The orthogonal first-order design (with three factors) consists of 8 (2^3) corner points located at the vertices of a cube and a centre point that is repeated four times, as

illustrated in Fig. 3.3. The details of the cutting parameters are described in the next section. The 12 tests will be divided into two blocks, and each block consists of six tests. The first block of experiments includes the trial numbers of 1, 4, 6, 7, 9 and 10. The four parameters (C, α, β, γ) in the tool-life equation $vT^{\alpha}f^{\beta}a^{\gamma}=C$ can be estimated using these six experiments, where v is cutting speed, T is the tool life, f is the feed rate and a is the depth of cut (Armarego, 1969). The second block of six experiments (2, 3, 5, 8, 11 and 12) has been used to provide a more precise estimate of the parameters. This design provides three levels for each independent variable.

Table 3.5 Design matrix of 3^3 factorial designs

Test No.	a	f	V	Test No.	a	f	v
1	−	−	−	15	+	0	0
2	0	−	−	16	−	+	0
3	+	−	−	17	0	+	0
4	−	0	−	18	+	+	0
5	0	0	−	19	−	−	+
6	+	0	−	20	0	−	+
7	−	+	−	21	+	−	+
8	0	+	−	22	−	0	+
9	+	+	−	23	0	0	+
10	−	−	0	24	+	0	+
11	0	−	0	25	−	+	+
12	+	−	0	26	0	+	+
13	−	0	0	27	+	+	+
14	0	0	0				

The tools were examined after each cutting pass across the workpiece until they failed. During the machining process, the adhered titanium alloy material at the flank was continuously observed, which sometimes made measurement of the wear very difficult

to obtain. The average and maximum flank wear lands, and notch wear at the nose and at the depth of cut, were measured with a toolmaker's microscope.

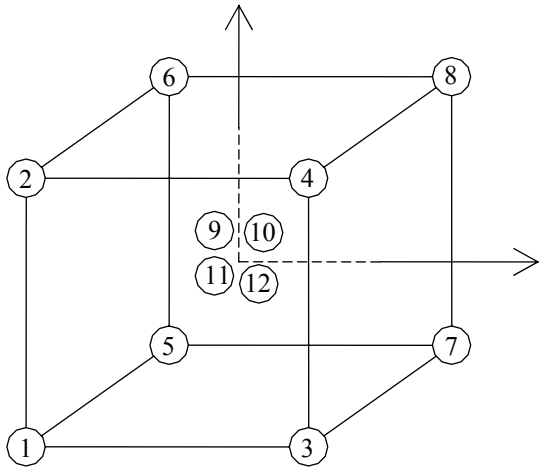


Figure 3.3 Orthogonal experiment design for three factors

The tool failure for rejecting a tool is often based on the following criteria: (i) the average flank wear or the maximum flank wear, (ii) the crater wear depth and (iii) flaking or fracture (ISO 8688-2, 1989). In this study, the average or maximum flank wear at the nose has been found to be the limiting factor that controlled the tool life in all cases. The following tool life end points based on the flank wear land are recommended (ISO 8688-2, 1989): (1) uniform wear: 0.3 mm averaged over all teeth; (2) localized wear: maximum value of 0.5mm on any individual tooth. Only one tooth is used for the experimental study. Hence, the mean value of the uniform wear value and localized wear value of 0.4 mm is chosen as the tool life criterion. The inserts were discarded when the average flank wear reached 0.4 mm, or when catastrophic fracture of the edge was observed.

Chapter 4

Analysis of cutting forces, tool life and tool wear mechanism

The performance of conventional tools is poor when used to machine titanium alloys. In this chapter, a new tool material, which is binder-less cubic boron nitride (BCBN), is used for high-speed milling of a widely used titanium alloy Ti-6Al-4V. The performance and the wear mechanism of the BCBN tool have been investigated when slot milling the titanium alloy in terms of cutting forces and tool life. Based on observations of flank and rake faces of cutting tools with SEM and EDX, main wear mechanism of the BCBN tool is analyzed when used in high-speed milling of Ti-6Al-4V. In addition, X-ray diffraction detector is also used to detect the material composite on the flank face after etching with hydrofluoric acid.

4.1 Introduction

When machining Ti-6Al-4V, conventional tools wear rapidly, because the poor thermal conductivity of titanium alloys results in higher temperature closer to the cutting edge during machining, and there exists strong adhesion between the tool and workpiece material (Zoya and Krishnamurthy, 2000). In addition, titanium alloys are generally difficult to machine at cutting speeds of over 30m/min with high-speed steel (HSS) tools, and over 60m/min with cemented tungsten carbide (WC) tools, resulting in very low productivity (López de lacalle et al., 2000). The performance of conventional tools is poor when machining Ti-6Al-4V. With the evolution of a number

of new cutting tool materials, advanced tool materials, such as cubic boron nitride (CBN) and polycrystalline diamond (PCD), have the good potential for use in high-speed milling. Some of the ultra-hard materials, such as PCD and CBN, have been used in the machining of titanium alloys (Zoya and Krishnamurthy, 2000; Nabhani, 2001; Kuljanic et al., 1998; Bhaumik et al., 1995; König and Neises, 1993). Both polycrystalline diamond and CBN are currently very expensive. In addition, PCD is highly reactive with titanium alloys and CBN cannot retain its higher hardness and strength at higher temperature, and consequently are not suitable for machining titanium alloys at higher speed.

CBN has high hardness that is second only to diamond. As for chemical and thermal stability, CBN is superior to diamond. Because of its excellent properties, CBN is widely used for the machining and polishing of hardened steel, ferrous metals and other materials.

Conventional CBN sintered tools are composed of CBN powder and metal or ceramics binder materials, which determine the bonding strength of the CBN particles. If the binder materials soften at higher cutting temperature, CBN tools cannot maintain their hardness and strength. Therefore, the mechanical and thermal properties of these conventional CBN tools strongly depend on the type and quantity of these binder materials. Recently, some single-phase CBN sintered tools without any binder materials are available. The binderless CBN (BCBN) tool used in this study is one of this type of tools, which is made by direct sintering method. To synthesize the BCBN tool, a high-purity hexagonal boron nitride (hBN) particle is first chosen as the starting material, which is 1.75×10^3 Kg/m^3 in density containing less than 0.03wt% of impurity

(B_2O_3). The hBN material is then placed in a Ta capsule and sintered at a pressure of about 7.7 GPa at various temperature settings (1,900°C to 2,700°C) using ultra high-pressure equipment (Sumiya et al., 2000). Finally, the high-purity polycrystalline CBN with superior mechanical characteristics is obtained. Excellent results had been achieved when this type of tool was used for high-speed milling of gray cast iron (Uesaka et al., 2000). In this experiment, the BCBN tools are used for high-speed milling of Ti-6Al-4V. Experimental studies by Zareena (2002) had indicated that the BCBN inserts have a remarkably longer tool life than conventional CBN inserts for the cutting conditions investigated (up to 400m/min). However, the cutting performance over a wide range of the cutting conditions and the wear mechanism of BCBN tool has not been investigated. In this chapter, BCBN inserts are selected to investigate their cutting performance in terms of cutting forces and tool when high-speed milling Ti-6Al-4V. In addition, tool wear characteristics are also analyzed.

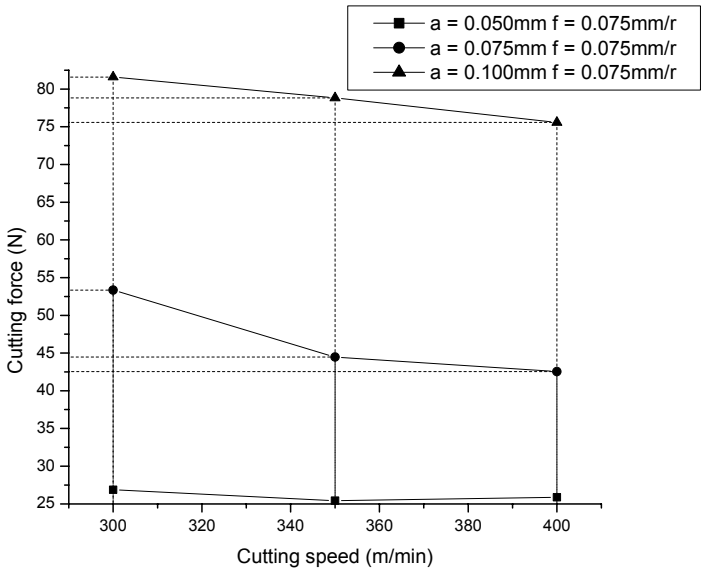
In order to investigate the cutting performance, the mathematical model (tool life) was developed on the basis of experimental results of high-speed milling. Then, the useful statistical techniques were used to test the validity of the models.

The analysis of variance (ANOVA) is probably the most useful procedure in the field of statistical inference to justify the quality of the estimated mathematical model. In this work, the commercial software SAS is used for performing ANOVA.

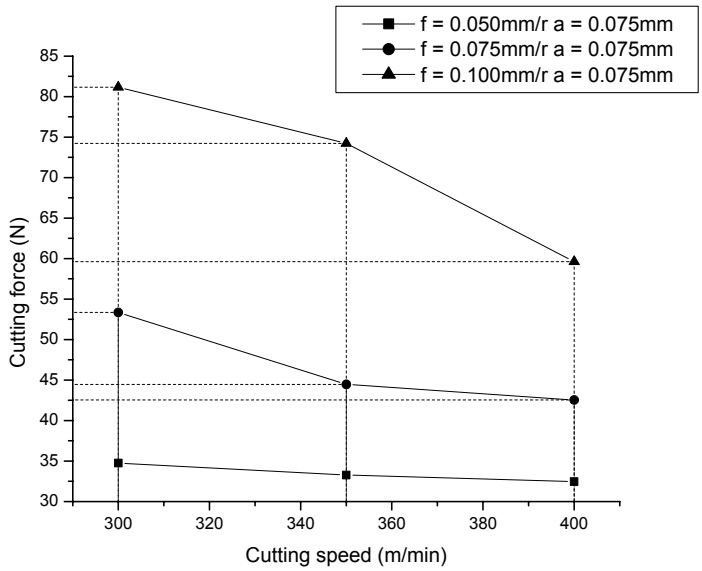
4.2 Analysis of cutting forces

Average resultant forces for different cutting speed, feed rate per tooth and axial depth of cut were investigated. Fig. 4.1 shows the average resultant forces versus cutting

speed at different conditions investigated. These were calculated as the square root of the sum of the squares for the three cutting force components F_x, F_y and F_z within one tooth period.



(a) Effects of depth of cut and cutting speed on cutting forces



(b) Effects of feed rate and cutting speed on cutting forces

Figure 4.1 Cutting forces at different cutting conditions

With the increase in the depth of cut and feed rate, the cutting forces also increase; while increasing the cutting speed decreases the observed average resultant forces as shown in Fig. 4.1. However, the effect of cutting speed is smaller than that of feed rate and depth of cut. With the increase in the cutting speed, the temperature on the tool-chip interface rises to the softening temperature of the workpiece materials, which makes the chips deform more easily, and causes the value of chip thickness to decrease. Because the chip thickness ratio r_c is inversely proportional to the value of chip thickness, the ratio r_c increases with the decrease in chip thickness. For the shear angle ϕ, there exists an equation (Trent and Wright, 2000):

$$\tan\phi = \frac{r_c \cos\alpha}{1 - r_c \sin\alpha} \tag{4.1}$$

where α is the rake angle. From Eq. (4.1), it can be seen that the shear angle increases with the increase of r_c. Therefore the shear angle increases with the cutting speed. In addition, the temperature rise at the tool-chip interface can also cause the friction force to decrease to some extent, such that the cutting forces decrease with the increase in the cutting speed for a given rake angle, feed rate and depth of cut. Similar conclusion has been drawn in (Tanaka, 1967).

4.3 Tool wear and its mechanism

4.3.1 Tool life analysis

The tool life data and volume of removed material at each cutting condition are given in Table 4.1, where *a* means depth of cut, *f* means feed rate and *v* means cutting speed.

The increase in the depth of cut and feed rate causes a decrease in the tool life. The effect of cutting speed on the tool life depends on the feed rate. At the higher feed rate

of 0.125mm/r and same depth of cut, the tool life decreases with the increase in the cutting speed. In the milling process, the undeformed chip thickness varies between zero and *f* (feed rate per revolution per tooth) during one cycle of the cutter revolution. The cutting forces also vary with the changing chip thickness, such that the stresses imposed on the cutting edge also fluctuate correspondingly. The fluctuating stresses at the cutting edge encourage flank wear and attrition wear. Oxley (1989) has pointed out that when the undeformed chip thickness is less than 0.05mm, the size effect exists. Under such a condition, the rate of increase of the cutting force is less than that of the undeformed chip thickness. Full details of the size effect are described in (Oxley, 1989; Shaw, 2003). For all experiments in this study, the size effect also exists when the undeformed chip thickness is less than 0.05mm. When the instantaneous chip thickness is greater than 0.05mm, the specific cutting pressure acting on the tool-chip interface approaches a constant value (Oxley, 1989). Therefore, the effect of the specific cutting pressure on the tool-chip interface does not vary significantly with the feed rates. However, at higher feed rate, the average undeformed chip thickness is larger, resulting in the increase in the radial cutting force, and thus causes higher normal stress on the cutting edge, which is an important factor that determines if the high speed machining is feasible. At higher feed rate and higher cutting speed, more workpiece materials are removed than those at higher feed rate and lower cutting speed within the same machining time. This causes the tool to undergo larger stress at higher frequency, which can cause the BCBN tool to wear quickly. Therefore, the tool life in Test 8 (using a lower cutting speed and a higher feed rate) is longer than that in Test 4 (using a higher cutting speed and a higher feed rate).

However, at the lower feed rate of 0.075mm/r and same depth of cut, the tool life at higher cutting speed is longer than that at lower cutting speed. For the lower tooth feed,

the cutting loads are not so high, and the cutting edge of the BCBN tool can withstand these loads. An increase in the cutting speed caused a rise in cutting temperature at the cutting edge. Because of its good thermal stability, the BCBN tool material can retain its strength at the higher temperature. On the other hand, the higher temperature at the tool-chip interface causes the workpiece material to soften, and facilitates the removal of the chips. Finally there would be an increase in tool life at higher cutting speed and a lower feed rate. Similar conclusions have been drawn by other researchers in (King and Vaughn, 1984). The best performance in terms of tool life and the amount of material removed was observed at cutting speed of 400m/min, feed rate of 0.075mm/r and depth of cut 0.075mm. Hence in practical machining, the higher cutting speed and lower feed rate are more optimal cutting conditions.

Table 4.1 Experimental results of factorial design

Test No.	*a(mm)*	*f(mm/r)*	*v(m/min)*	Tool life (min)	Material removal(cm^3)
1	0.075	0.075	400	39.759	247.95
2	0.125	0.075	400	7.769	80.75
3	0.075	0.125	400	4.11	42.75
4	0.125	0.125	400	1.37	23.75
5	0.075	0.075	300	29.841	139.65
6	0.125	0.075	300	6.09	47.5
7	0.075	0.125	300	8.76	68.4
8	0.125	0.125	300	1.825	23.75
9	0.100	0.100	350	3.92	38
10	0.100	0.100	350	4.074	45.6
11	0.100	0.100	350	4.074	45.6
12	0.100	0.100	350	4.312	41.8

To develop the mathematical model of tool life, both the data for the machining responses (tool life/material removal) and factors (cutting speed *v*, depth of cut *a* and feed rate *f*) were logarithmically transformed. Then the parameters for the mathematical model were determined using the multiple regression method, at the same time the significance of the model and parameters were then analyzed using SAS procedure GLM (SAS Institute, 1985). For the RSM models, the effect of factors and their interaction on the responses can be found with ANOVA. The results of ANOVA for the tool life were shown in Table 4.2.

Table 4.2 ANVOA of cutting performance in terms of tool life

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	3	9.99546	3.33182	39.203	0.0001
Error	8	0.67992	0.08499		
C Total	11	10.67537			
a (depth of cut)	1	4.57284419	4.57284419	53.80	0.0001
f (feed-rate)	1	5.38140083	5.38140083	63.32	0.0001
v (speed)	1	0.04121163	0.04121163	0.48	0.5059
Root MSE	Dep Mean	R-square	Adj R-sq	C.V.	
0.29153	1.75602	0.9363	0.9124	16.60172	

Table 4.3 Parameter estimates of cutting performance in terms of tool life

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
INTERCEP	1	-9.588266	4.39811508	-2.180	0.0609
a (depth of cut)	1	-2.934462	0.40248621	-7.291	0.0001
f (feed-rate)	1	-3.201843	0.40248621	-7.955	0.0001
v (speed)	1	-0.498557	0.71595802	-0.696	0.5059

F-test is used for testing the significance of regression. The regression output of F-values is shown in the Table 4.2. From Table 4.2, it can be seen that the feed rate is most dominant factor of all the considered factors for representing tool life. For milling

process, the instantaneous chip thickness depends on feed rate per tooth. Apart from feed rate, the depth of cut is another important factor. In comparison, the cutting speed has less effect on the tool life among these three factors. And the multiple regression coefficient (R^2) of the first order model was found to be 0.9363. This shows that the first order model can explain the variation in tool life to the extent of 93.63%. By looking up tables of the F-distribution, the tabulated value of F-Value $F_{3,11}(0.01)$ can be obtained as 6.22. So the calculated F-value for this model is greater than the tabulated one. Therefore, the results of ANOVA indicate that the main effects of cutting conditions (a, f, v) are significant at 99% confidence level and that their interaction effects are insignificant. As the first order model has high predictability, it is accurate enough to represent the response better. Hence, the first order model was considered as one term of the objective function for optimization in Chapter 7.

The estimates of the parameters and the corresponding T-test values are given in Table 4.3. The T-test tests the hypothesis that the mean value of the estimated variables is equal to the given number specified using the option H0 =. By default, the value for H0 is 0 in SAS. Regression procedures of SAS calculate the t-statistic and its significance probability for the specified hypothesis under the assumption that the sample comes from an approximately normal distribution (SAS Institute, 1985). If the p-value or the probability associated with the T-test is smaller than some small level (usually set at 0.05 for 95% confidence level), the hypothesis is almost unlikely. Then, the hypothesis is rejected. If the p-value associated with the t-test is not small (say $p >$ 0.05), then the hypothesis cannot be rejected, and the mean is not different from the hypothesized value (Draper and Smith, 1981). From p-values listed in Table 4.3, it is impossible for the mean values of feed rate and depth of cut to be 0. But the hypothesis

about the intercept and cutting speed cannot be rejected. Finally, the tool life in terms of cutting speed, feed rate and depth of cut is obtained as follows:

$$T = 6.8528 \times 10^{-5} a^{-2.934462} f^{-3.201843} v^{-0.498557} \quad (4.2)$$

The tool performance measured by the volume removed before tool failure was also investigated. With this metric, the performance of the tool increases with the cutting speed at lower feed rate and same depth of cut. However, at higher feed rate and lower depth of cut, the cutting performance of the tool decreases with the increase in cutting speed, as shown in Test 3 and Test 7 in Table 4.1. This trend is same as that for the tool life.

The results of ANOVA for volume of removed materials were shown in Tables 4.4 and 4.5, where DF means degree of freedom. Multiple regression coefficient (R^2) is equal to 0.8841, so first order model can explain the variation in tool life to the extent of 88.41%. The calculated F-value (20.344) for this model is also greater than the tabulated one (6.22). Therefore, the main effects of cutting conditions (a, f, v) are significant at 99% confidence level and that their interaction effects are insignificant.

Table 4.4 ANVOA of cutting performance in terms of removal volume

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	3	4.51610	1.50537	20.334	0.0004
Error	8	0.59225	0.07403		
C Total	11	5.10835			
a (depth of cut)	1	1.96184563	1.96184563	26.50	0.0009
f (feed-rate)	1	2.51105236	2.51105236	33.92	0.0004
v (speed)	1	0.04320249	0.04320249	0.58	0.4668
Root MSE	Dep Mean	R-square	Adj R-sq	C.V.	
0.27209	4.00255	0.8841	0.8406	6.79781	

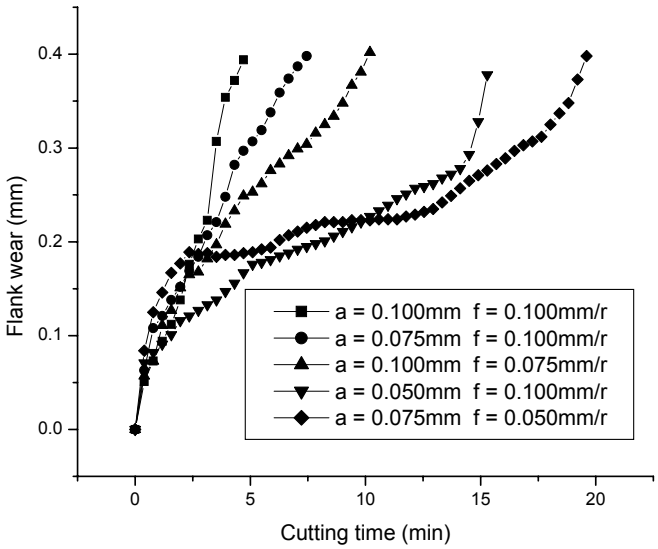
The regression output of F-values shown in the Table 4.4 indicates that the feed rate is also most dominant factor of all the considered factors for representing removal volume. From Table 4.5, the tool life in terms of cutting speed, feed rate and depth of cut is obtained as follows:

$$V = 1.9549 \times 10^{-4} a^{-1.922995} f^{-2.18859} v^{0.510457} \quad (4.3)$$

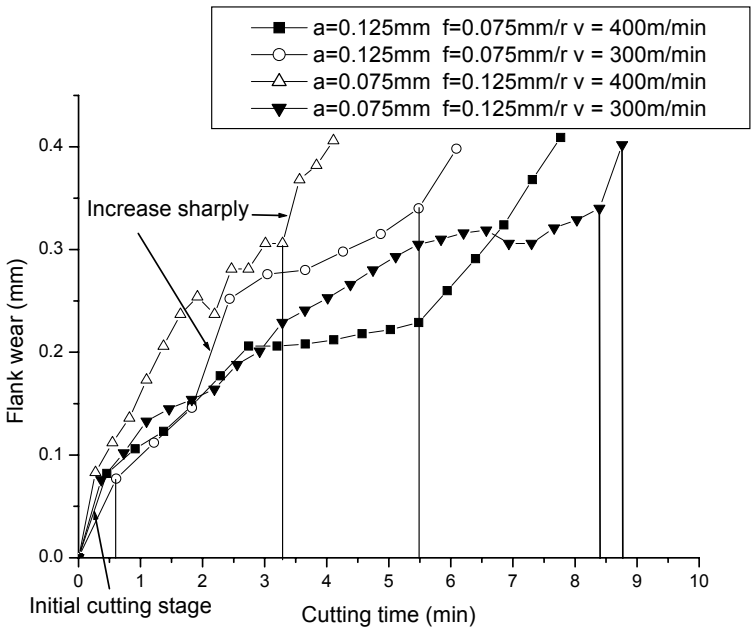
Table 4.5 Parameter estimates of cutting performance in terms of removal volume

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
INTERCEP	1	-8.539884	4.10478099	-2.080	0.0711
a (depth of cut)	1	-1.922995	0.37564223	-5.119	0.0009
f (feed-rate)	1	-2.188590	0.37564223	-5.826	0.0004
v (speed)	1	0.510457	0.66820691	0.764	0.4668

The progress of flank wear of the BCBN tool under four different cutting conditions is shown in Fig. 4.2, where *a* means depth of cut, *f* is the feed rate and *v* is cutting speed. The results show that the flank surface of the tool wears quickly at the initial cutting stage. At the later stage, the wear rate shows a tendency to increase. When milling Ti-6Al-4V at depth of cut 0.075mm, a higher feed rate of 0.125mm/r, and cutting speed of 400m/min, it was observed that workpiece material adhered to the flank (as shown in Fig. 4.5 (c)), such that the values of VB appeared to decrease, as shown in Fig. 4.2. However, when the adhered material was removed, the average flank wear increased quickly. Under four other machining conditions: (1) a = 0.125mm, f = 0.075mm/r and v = 400m/min; (2) a = 0.075mm, f = 0.125mm/r and v = 300m/min; (3) a = 0.050mm, f = 0.100mm/r and v = 350m/min; (4) a = 0.075mm, f = 0.050mm/r and v = 350m/min, the flank wear rate reached a steady state after a rapid initial wear. Subsequently the wear rate showed an obvious tendency to increase again.



(a)



(b)

Figure 4.2 Average flank wear vs. cutting time at different cutting conditions

(a) at cutting speed of 350m/min; (b) at cutting speed of 300m/min and 400m/min

As reported in (Zareena, 2002; Wang et al., 2005), the tool life of the BCBN tool is much longer than that of the CBN tool under the same cutting condition. This is because in high-speed milling, the binder materials of sintered CBN tools remarkably

affect their mechanical and thermal properties (Sumiya et al., 2000). Therefore the cutting performance of these sintered CBN tools depends strongly on the content and kind of binder materials. The binderless CBN tool does not have this problem. Compared to conventional CBN, it has been found to be more effective for the ultra-high-speed milling of cast iron (Uesaka et al., 2000). From Table 3.1 in Chapter 3, BCBN tools have higher hardness and thermal conductivity than CBN tools. Hardness and thermal conductivity are very important properties for a cutting tool since they are closely related to the resistance to wear and chipping of the cutting tool. So BCBN tools are the most suitable cutting tools for high-speed milling of titanium alloys both economically and functionally.

4.3.2 Tool wear mechanism

At the initial cutting stage (shown in Fig. 4.3), SEM analysis of the worn surface of the tool suggests that there are some micro-grooves on the flank face, which result from the abrasion between the tool and workpiece.

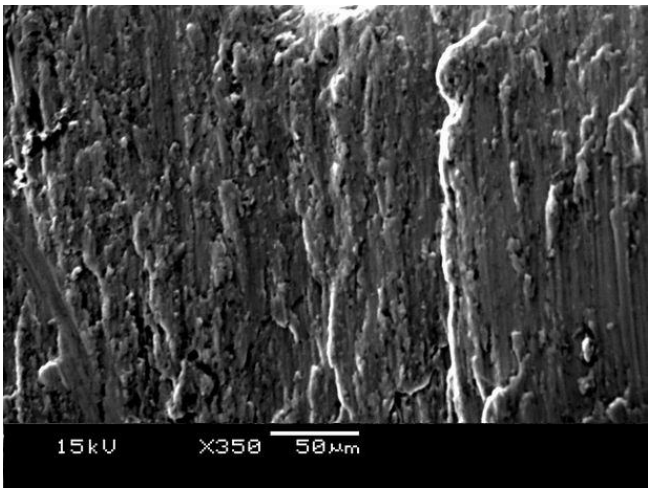
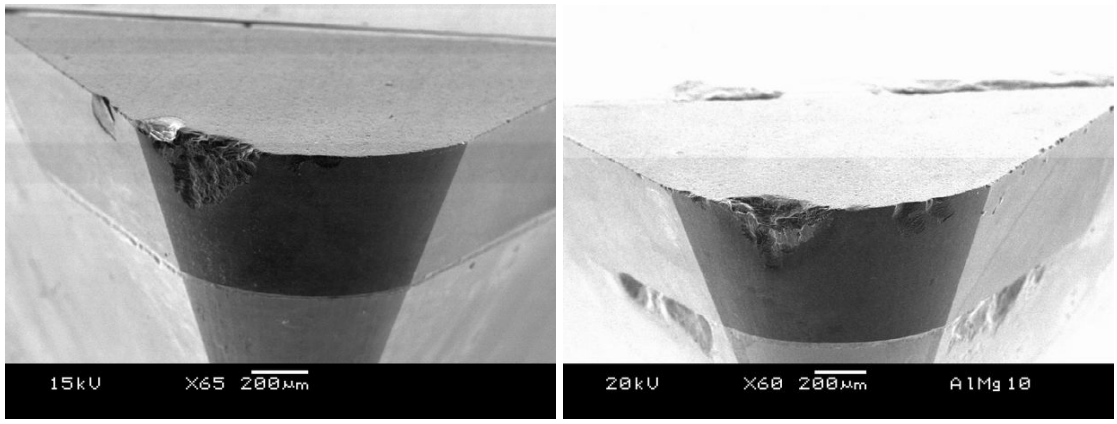


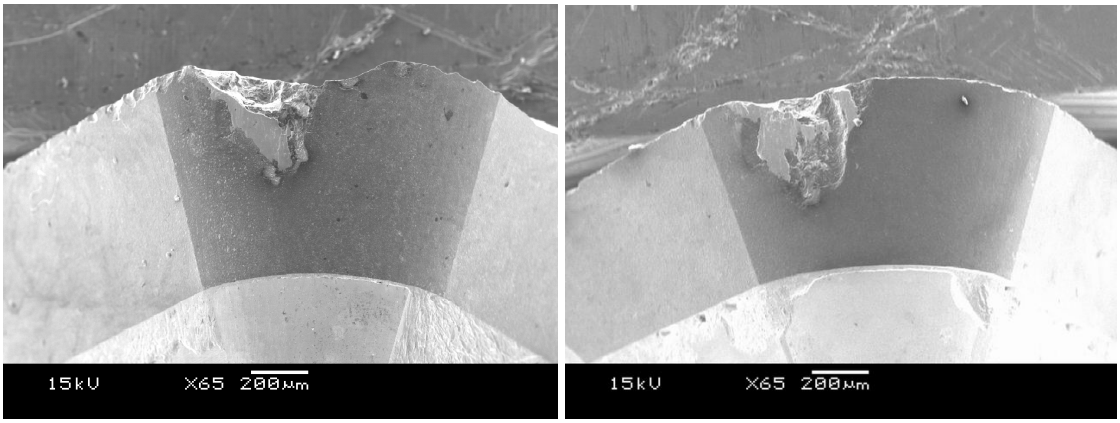
Figure 4.3 SEM of the flank face at the initial cutting stage

When tool wear reaches the failure criterion, there are no clear micro-grooves on the flank face (as shown in Fig. 4.4). This indicates that abrasion wear is not the main wear

mechanism for BCBN tools although it exists at the initial cutting stage. Non-uniform wear at the flank was found to dominate under all cutting conditions, while wear on the minor cutting edge was so small that it could not have significant effect on the tool during machining, as shown in Fig. 4.4.



(a)a=0.05mm, f=0.05mm/r, v=350m/min (b)a = 0.10mm, f = 0.075mm/r, v=350m/min

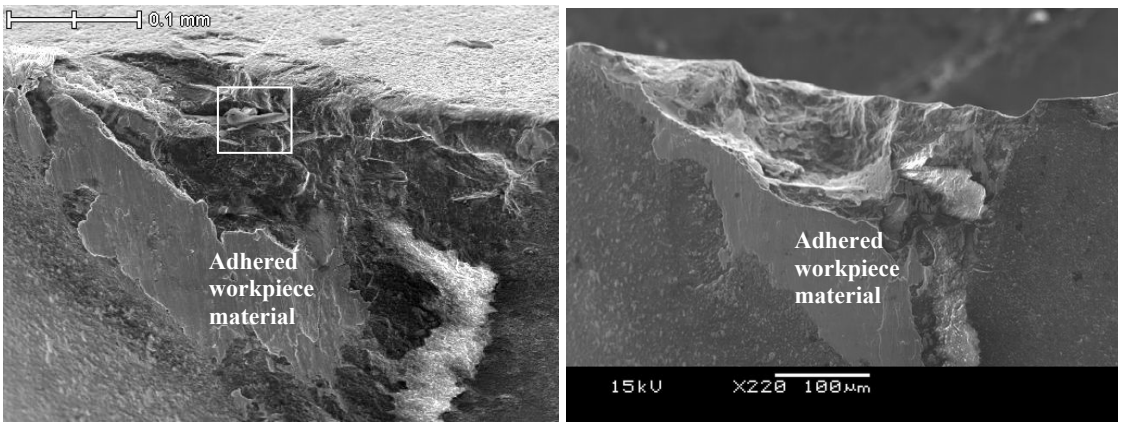


(c) a=0.125mm,f=0.075mm/r,v=400m/min (d) a=0.075mm,f=0.075mm/r,v=400m/min

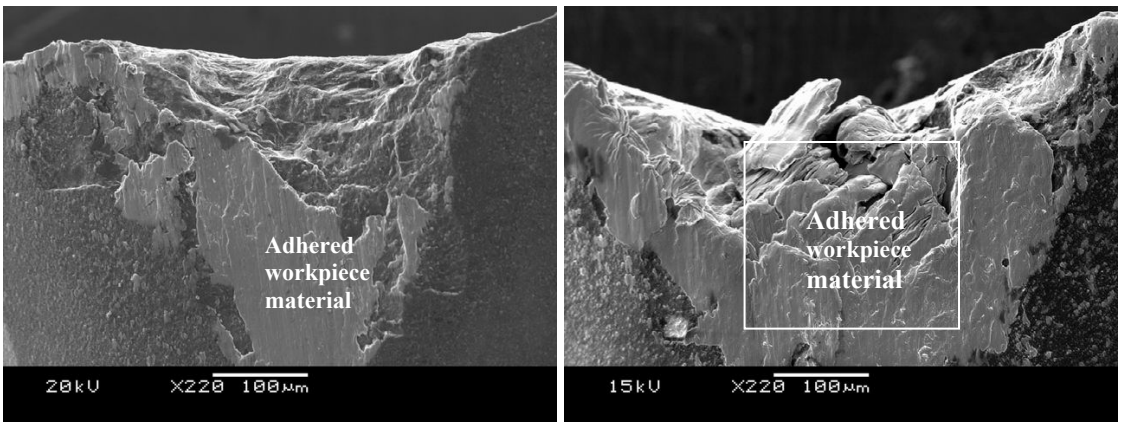
Figure 4.4 SEM of the flank of BCBN tools at four different conditions where the non-uniform flank wear is the dominant wear for these four cases

Adhesion of the workpiece materials onto the flank face of the tool was also observed, which can be seen in Fig. 4.5, and it indicates that there is a strong bond at the tool/workpiece interface.

In the entire machining process, the titanium chips are very thin because the depth of cut is very small. The combination of a small contact area and the low thermal conductivity of Ti-6Al-4V results in very high cutting temperatures at the nose of the tool. There is also strong chemical reactivity of titanium at the cutting temperature (>500ºC) with almost all tool materials available (Machado and Wallbank, 1990). The close contact between the tool and workpiece at the high temperature provides an ideal environment for inter-diffusion of the tool and workpiece material (Bhaumik et al., 1995). The high affinity of titanium alloys gives rise to the pick up of the heated outer surface layer of the workpiece during machining (Yang and Liu, 1999).



(a) a = 0.100mm, f = 0.10mm/r, v = 350m/min (b) a = 0.125mm, f = 0.075mm/r, v = 400m/min



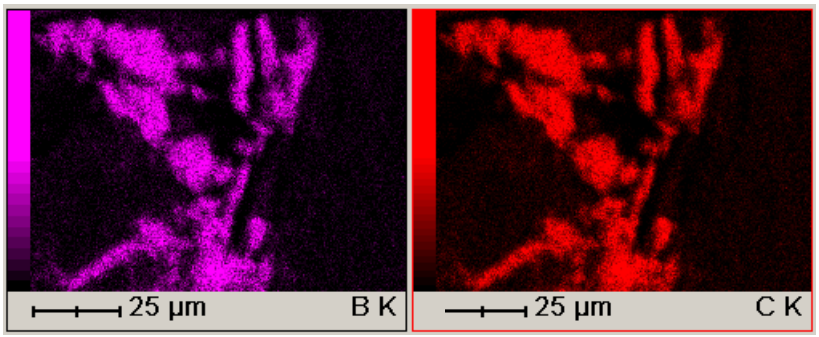
(c) a = 0.075mm, f = 0.125mm/r, v = 400m/min (d) a =0.125mm, f =0.125mm/r, v = 400m/min

Figure 4.5 SEM of the flank faces of BCBN tools at four different conditions; workpiece material is adhered to the flank for these four cases

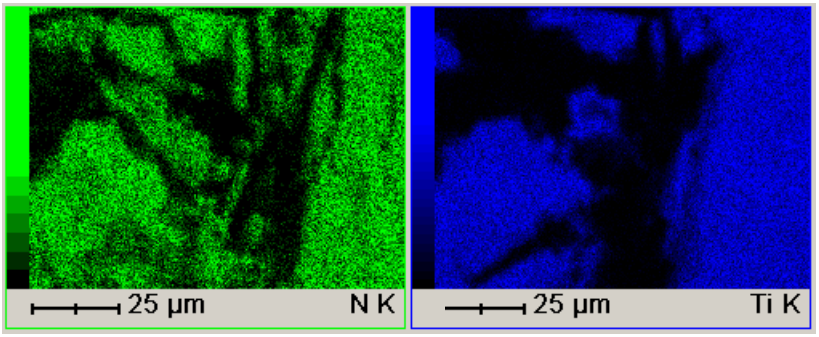
When the depth of cut and feed rate are lower, there is less adhered workpiece material to the flank (as shown in Fig. 4.5 (b) and (c)) than at larger depth of cut and higher feed rate (as shown in Fig. 4.5 (d)). Since the heat and forces generated during the cutting process are higher at larger depth of cut and higher feed rate, it can be assumed that the higher temperature is the main reason that causes the adhesion of workpiece material onto the tool flank and rake faces. At the increased depth of cut and higher feed rate, cutting forces generated during the cutting process are also higher, resulting in higher stresses in the small contact region. The greater cutting temperature and stresses generated on the flank face close to the nose probably reduced the yield strength of the tools, leading to higher wear.

After the tool was worn out, hydrofluoric acid (HF) was used to remove the bonded Ti-6Al-4V on the rake and flank surfaces. The tools treated with HF acid were then examined with EDX. The EDX results shown in Fig. 4.6 (a), (b), (c) and (d) indicate the presence of the boron (B), carbon (C), titanium (Ti) and nitride (N) elements, respectively. Fig. 4.6 (e) is derived by overlaying the four EDX figures (a), (b), (c) and (d) together. Interestingly, the carbon element was also observed on the flank, although BCBN tools and Ti-6Al-4V do not contain any carbon element. This carbon element may have come from the ECOCOOL S/So 50 coolant. The higher temperature at the tool/workpiece interface could have resulted in the decomposition or reaction between the titanium and the coolant to produce some carbon element on the flank. It can be seen that some region contains only the element Ti and N from Fig. 6 (e), so it can be assumed that in this region, there exists titanium nitride (TiN). TiN generally provides low friction against steels, carbides, ceramics, platings, etc. In other regions, there exists boron carbon nitride (BCN), which is expected to create another series of novel

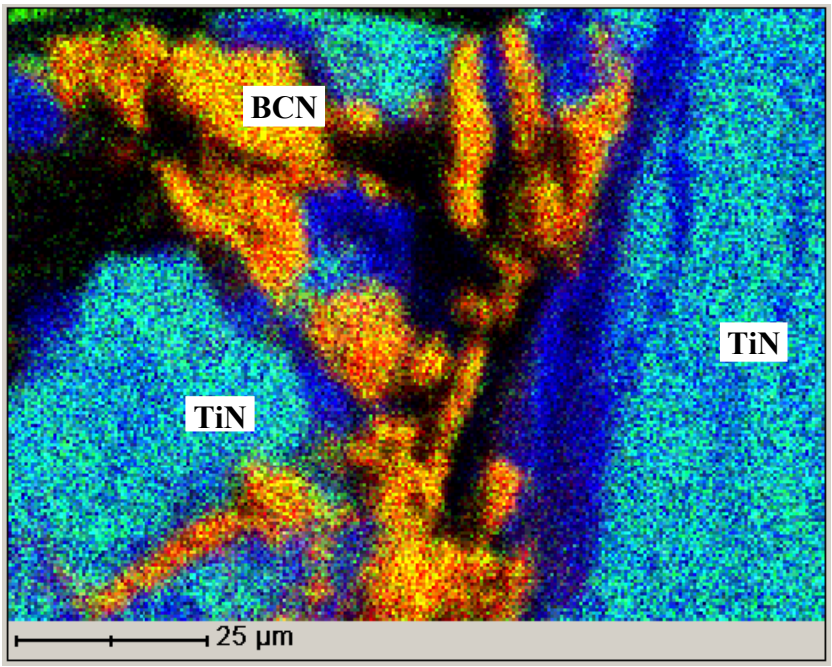
materials with physical properties intermediate to that of the pure carbon and CBN phases.



(a) Boron (b) Carbon



(c) Nitride (d) Titanium



(e) Overlay of (a)-(d) indicating TiN (i.e. presence of both Ti & N)

Figure 4.6 EDX analysis of the flank to detect the element and the four figures over each other at a = 0.125mm, f = 0.125mm/r and v = 400m/min

In order to detect the composites on the surface, an X-ray detector (XRD) was used to examine the flank. The X-ray diffraction in Fig. 4.7 shows that there are some other materials apart from CBN. After verifying the three strongest peaks in Fig. 4.7 (a) and (b), the possible main composites on the flank are titanium carbide nitride (TiCN) and TiN, as shown in Fig. 4.7 (b). The mechanism of formation of the layer is similar to that reported on PCD tools (Brookes et al., 1991). It is possible that diffusion across the interface results in the formation of a titanium nitride layer which would then protect the tool, forming a barrier to further diffusion and loss of tool material in the chips. Some support for this contention can be obtained from the observation that graphite crucibles are able to withstand attack by liquid titanium metal when a stable layer of titanium nitride is formed on the surface of the crucible (McQuillan and McQuillan, 1956).

The adhered workpiece material has some adverse effect on the quality of the machined surface. However, it could have positive impact on the machinability as well, which is helpful to reduce the wear rate (Yang and Liu, 1999). The adhered workpiece material forms a layer which becomes saturated with tool particles and serves as a diffusion boundary layer. This reduces the rate of transport of tool material into the workpiece or chips and consequently, the wear rate (Ezugwu et al., 2003). Thus, the tool wear rate is lower at the steady cutting stage, as can be seen in Fig. 4.2. As the cutting process progresses, the adhered layer remains attached to the tool until the stress reaches a sufficiently high level to detach it from the reaction layer, as will inevitably happen during the cutting process. The strong bonding strength between tool materials and the reaction layer will probably cause some hard particles of the tool material to be carried away on the flank (Rectangular regions shown in Fig. 4.5 (a)).

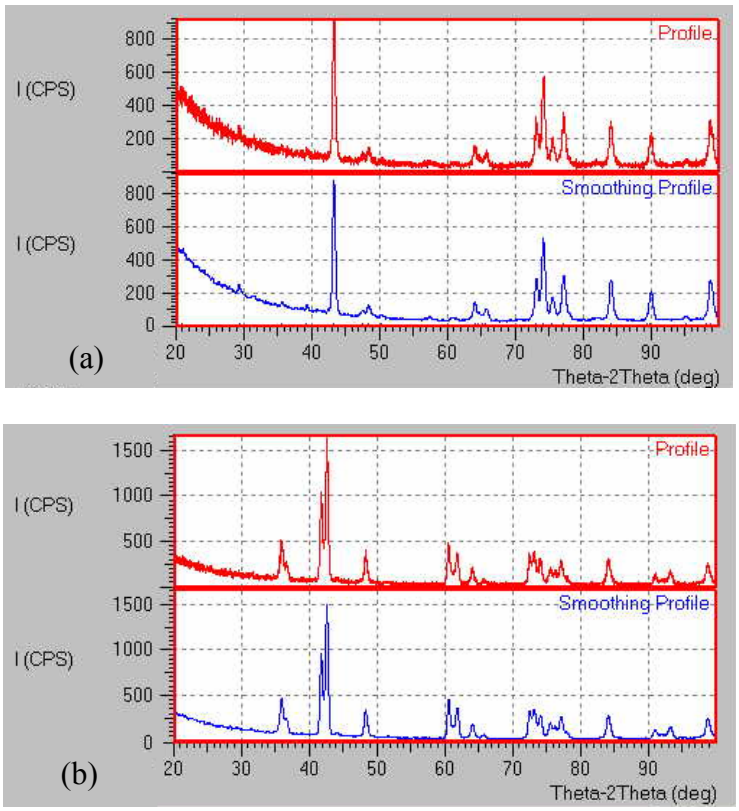


Figure 4.7 (a) X-ray diffraction of Fig. 4.6, (b) X-ray diffraction of TiN and TiCN

The removal of grains of the tool material by the adherent chip or workpiece was observed on the flank face after milling titanium Ti-6Al-4V. In Fig. 4.8 (a) (which is the close-up view of the rectangular region of Fig. 4.5 (a) at higher magnification), there is a fragment that shows the aggregate particles of tool material based on EDX analysis (shown in Fig. 4.8 (b)). In addition, there are some workpiece materials that are adhered to the unevenly worn surfaces on the flank, as shown in Fig. 4.5 (a). The above evidence indicates that fragments of tool material have been taken away by the adhered workpiece material. This causes accelerated attrition wear on both the rake and flank faces. This is reflected in the flank wear versus time plot, where the wear rate tends to increase quickly as indicated in Fig. 4.2 (b). Therefore, flank wear of the BCBN tool results not only from attrition, but also from the removal of aggregate of tool materials.

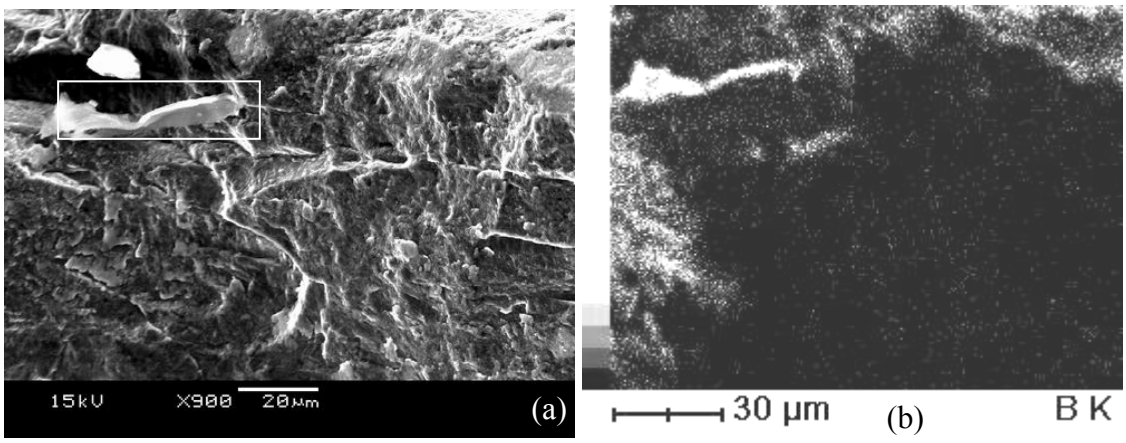
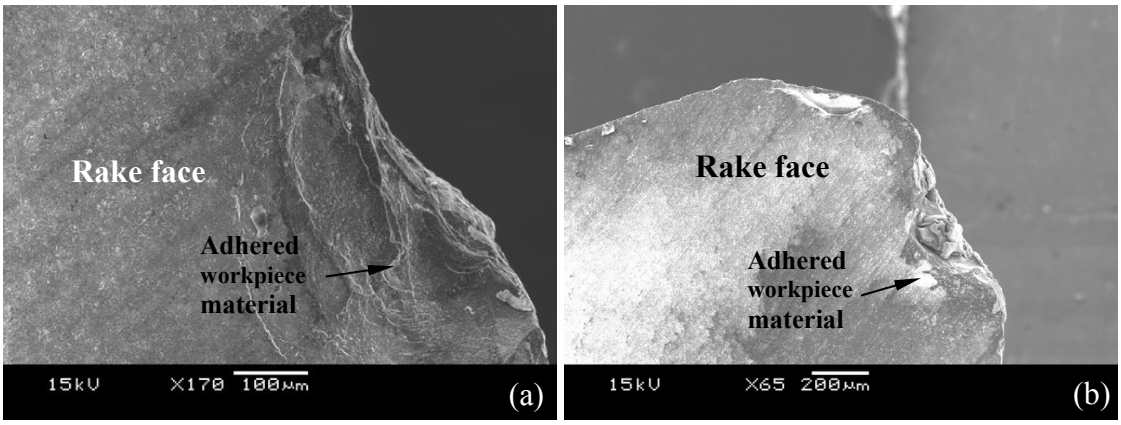
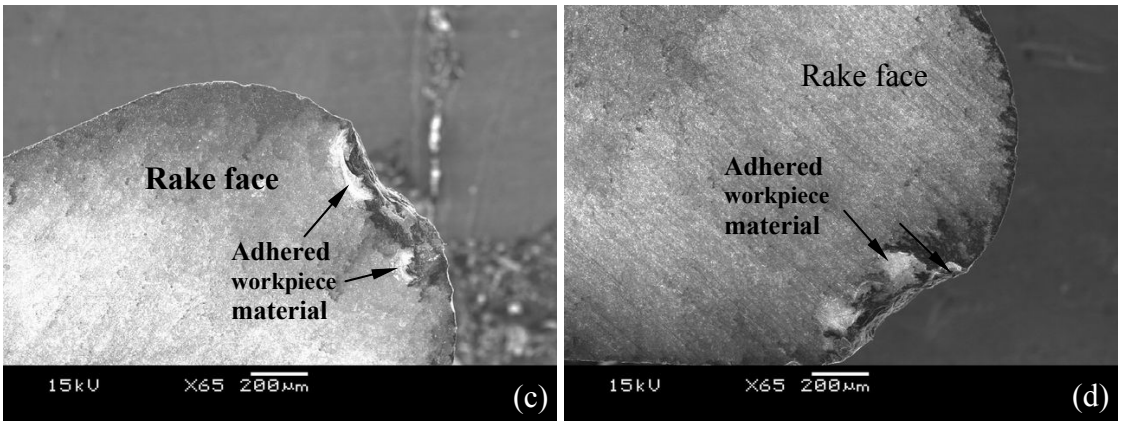


Figure 4.8 SEM and EDX of the flank face (a) Enlarged rectangular region indicated in Fig. 4.5 (a); (b) EDX of (a) shows the fragment in the rectangular region of Fig. 4.5 (a) coming from the tool material

SEM figures of the rake faces are shown in Fig. 4.9. Some workpiece materials adhered to the rake when milling Ti-6Al-4V at the larger depth of cut or higher feed rate, which can be seen in Fig. 4.9 (a), (b), (c) and (d). However, at the lower depth of cut and lower feed rate, there are only some adhered materials that are mainly found on the flank face rather than on the rake face after failure of the tools, as shown in Fig. 4.10 (a), (b), (c) and (d). There is no obvious crater wear observed on the rake face for all cases. In this study, the chip thickness was relatively thin; hence, the generated heat during cutting existed in a very small chip-tool contact area. The BCBN tool material has better thermal stability, which enables the tool to sustain higher cutting temperatures. Even if the temperature in the small heat-affected zone is very high, the BCBN tool material can still maintain its properties. In addition, with higher hardness and thermal conductivity than CBN tools, the BCBN tool has better resistance to wear, and chipping of the tool seldom happens. Therefore no obvious crater wear was observed.



(a) a=0.10mm, f=0.10mm/r, v=350m/min (b)a =0.125mm, f=0.125mm/r,v=400m/min

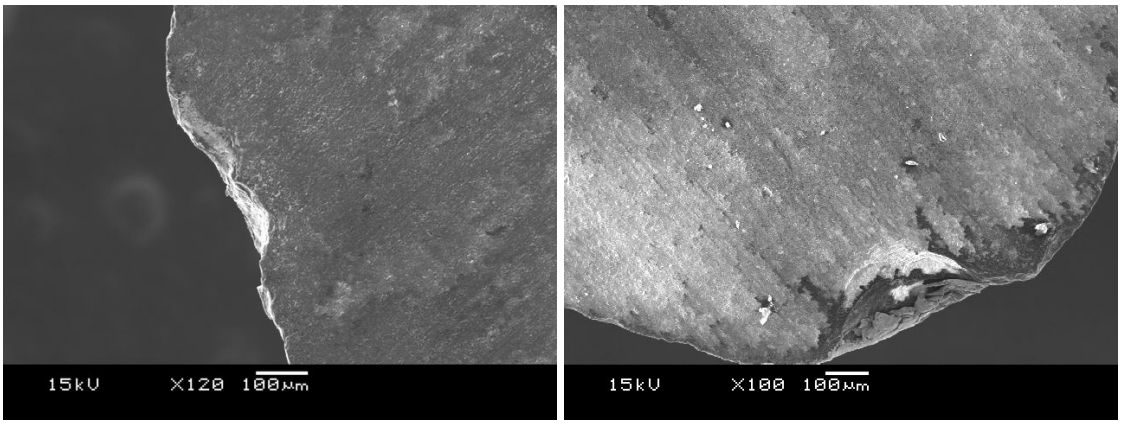


(c)a = 0.125mm, f = 0.125mm/r v = 300m/min (d) a = 0.125mm, f = 0.075mm/r, v = 300m/min

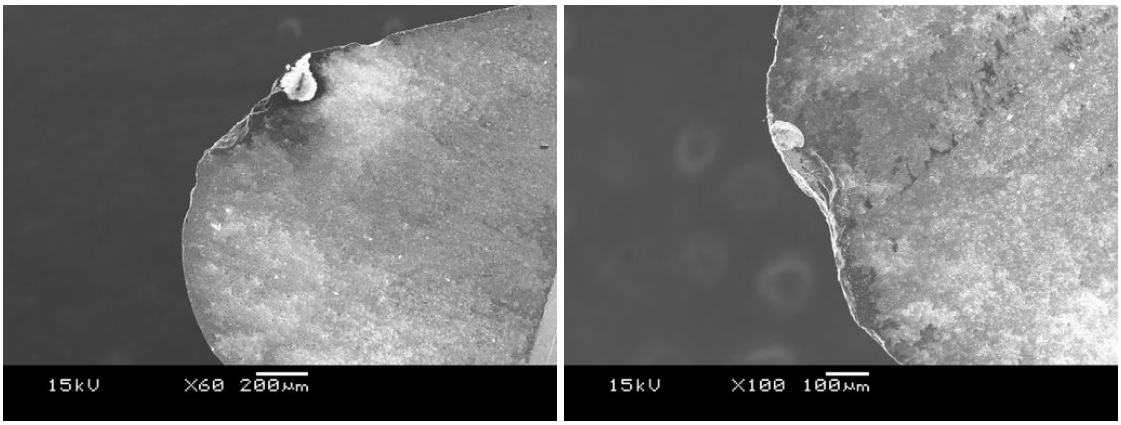
Figure 4.9 SEM of the rake faces of BCBN tools at four different cutting conditions

4.3.3 EDX observation of undersurface of chips

There are two main stages in which segmented chips are generated (López de lacalle et al., 2000), and the segmented (or cyclical kind of) swarf produced in the machining tests is shown in Fig. 4.11 (a), which results from the instability in the chip forming process. Fig. 4.11 (b) is rectangular region of Fig. 4.11 (a) and Fig. 4.11 (c) is the EDX output of Fig. 4.11 (b). From Fig. 11 (c), apart from titanium, aluminum and vanadium, which are the main elements of Ti-6Al-4V, there are also some tool material elements, such as boron and nitride, on the underside of the chips.



(a) a=0.075mm, f=0.075mm/r,v=400m/min (b) a=0.10mm, f=0.075mm/r, v=350m/min

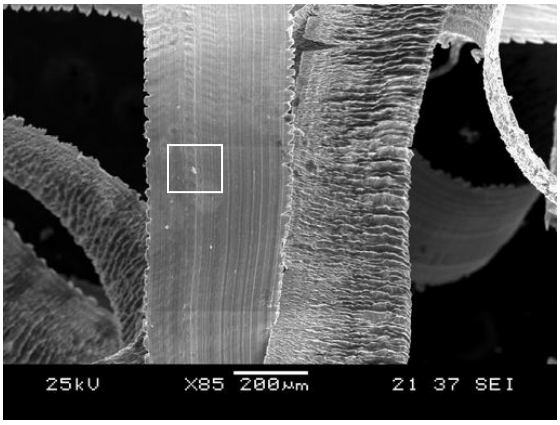


(c) a=0.05mm, f =0.05mm/r, v=350m/min (d) a=0.075mm, f =0.05mm/r, v=350m/min

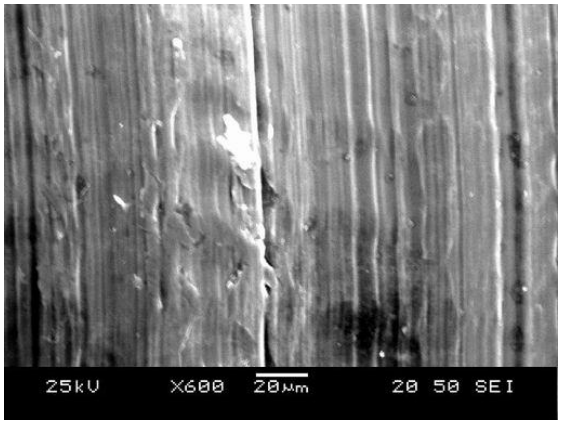
Figure 4.10 SEM of the rake faces of BCBN tools at four different cutting conditions; no obvious crater wear is observed on the rake for these four cases

The further quantitative analysis of Fig. 4.11 (b) is shown in Fig. 4.11 (d). Fig. 4.11 (d) shows that the light element boron is not sensitive enough to be detected and quantified using EDX. However, by quantifying the elements detected with EDX analysis, the weight percentage of element N is 5.98%, which is greater than that of element Al or V. As shown in Table 3.2 in Chapter 3, the element N also exists in Ti6Al4V, but its weight percentage is very low (only 0.01%). Therefore the transfer of cutting tool material to the chips has occurred, which involves dissolution of material from the tool by diffusion into the adjacent zones of the chip. Similar results were

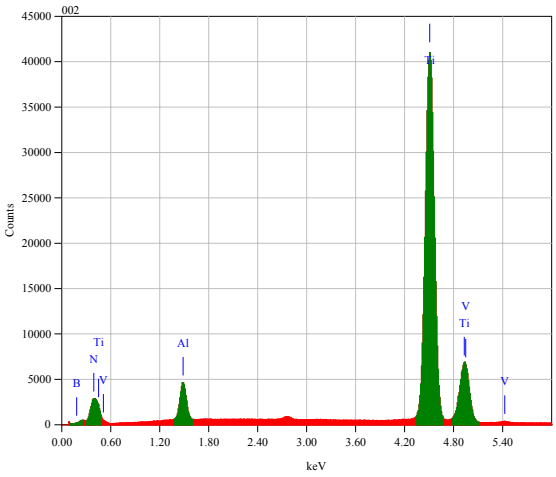
observed by other researcher (Dearnley and Grearson, 1986; Jawaid et al, 1999). The chip/workpiece at temperature above 700ºC provides an ideal environment for diffusion of the tool material atoms across the tool/chip and the tool/material interfaces (Dearnley and Grearson, 1986).



(a) Segmented swarf



(b) Rectangle region of (a)



(c) EDX of (b)

Element	(keV)	mass%	Error%	At%
B K*				
N K	0.392	5.98	0.15	17.32
Al K	1.486	4.90	0.06	7.38
Ti K	4.508	85.47	0.13	72.40
V K	4.949	3.65	0.16	2.91
Total		100.00		100.00

(d) Quantitative analysis of (b)

Figure 4.11 SEM and EDX of flank face (a = 0.1mm, f = 0.1mm/r and v = 350m/min)

There were only very small amount of tool material elements on the underside of the chips. So diffusion-dissolution wear also exists for BCBN tools, but it is not the dominant wear.

4.4 Concluding remarks

From the observed performance of BCBN tools in the machining of Ti-6Al-4V, the following conclusions can be drawn:

1. Non-uniform flank wear is the dominant wear pattern of BCBN tools when high-speed milling Ti-6Al-4V.
2. Based on the comparison with tool life of CBN and PCD tools obtained by Zareena (2002), BCBN tools have longer tool life. With the excellent mechanical properties (such as hardness and strength at higher temperature), the binderless CBN appears to become a new cutting tool material for high-speed machining of titanium alloys with superior cutting performance.
3. At lower feed rate (0.075mm/r) and the same depth of cut, the tool life at higher cutting speed is longer than that at lower cutting speed. The higher cutting speed and lower feed rate are more optimal cutting conditions.
4. There is a strong bond at the tool/workpiece interface, and the adhered workpiece material to the flank face is helpful in reducing wear rate. However, when the adhered workpiece material is subsequently removed, it will not only cause some removal of aggregate of tool materials, but also result in accelerated attrition wear on the flank face.
5. There are some workpiece materials adhered to the rake face when milling Ti-6Al-4V at higher depth of cut, feed rate and cutting speed. However, at other cutting conditions, there are only some adhered materials mainly found on the flank face, rather than on the rake face after failure of the tool.
6. Based on the EDX output of chips, dissolution of material from the tool by diffusion into the adjacent zones of the chip happens. This may cause diffusion-dissolution wear for BCBN tools.

Chapter 5
Modeling of cutting forces in milling

In this chapter, a new cutting force model for slot milling based on Oxley's predictive machining theory is developed. Firstly, a brief review of Oxley's cutting force model is given. Then, a new approach to describe the milling geometry is presented. The effect of nose radius of the tooth is considered on the calculation of uneven uncut chip area. Then, the non-uniform uncut chip area has been represented with its equivalent element, which is suitable for 2D simulation. Based on the published data, parameters of the Johnson-Cook (JC) constitutive model of flow stress are estimated with the Gauss-Newton algorithm. With the constitutive equation for Ti-6Al-4V and equivalent element representation, the milling process has been simulated with finite element method (FEM) to estimate the shear angle and strain-rate constant and further the cutting forces. Finally, some experiments were also carried out to verify the method.

5.1 Conventional orthogonal cutting theory

The widely used shear plane model was first proposed by Ernst and Merchant (1941), based on the assumption of continuous chip formation in a narrow zone which is idealized as a plane with uniformly distributed shear stress. The forces in chip formation in machining of Ti-6Al-4V are shown in Fig. 5.1. The area between the boundary *CD* and *EF* is the chip formation zone or the shear zone. Based on the assumption that the boundary *CD* and *EF* are parallel and equidistant from *AB*, Oxley (1989) developed a parallel-sided shear zone theory to predict the cutting forces.

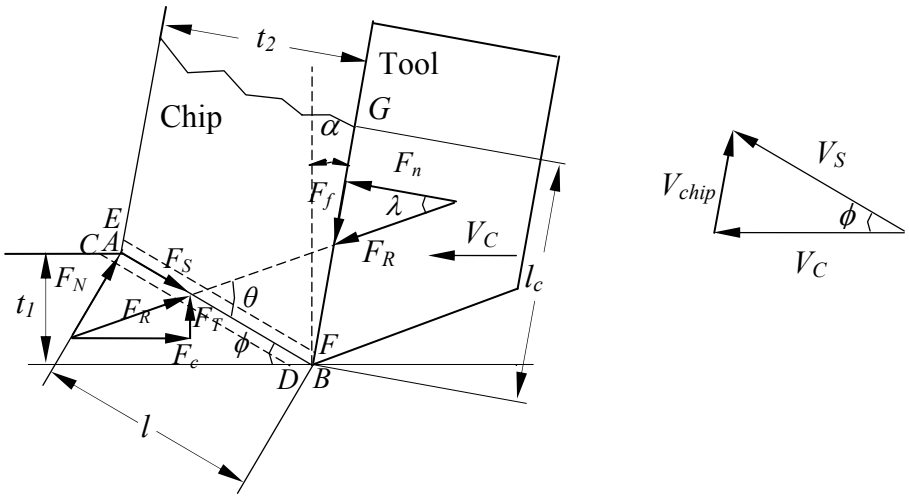


Figure 5.1 Cutting forces diagram based on the shear plane model

He assumed that the shear flow stress k_{AB} along the shear plane is constant; the shear force can be written as:

$$F_S = k_{AB} w l = \frac{k_{AB} t_1 w}{\sin\phi} \tag{5.1}$$

where w is the width of cut, l is the length of the shear plane, t_1 is the undeformed chip thickness, and ϕ is the shear angle.

According to the geometric relations as shown in Fig. 5.1, the following equations can be obtained:

$$\left.\begin{aligned}
V_{chip} &= \frac{V_C \sin\phi}{\cos(\phi-\alpha)} \\
V_S &= \frac{V_C \cos\alpha}{\cos(\phi-\alpha)} \\
t_2 &= t_1 \cos(\phi-\alpha)/\sin\phi \\
\theta &= \phi+\lambda-\alpha \\
F_R &= \frac{F_S}{\cos\theta} \\
F_C &= F_R \cos(\lambda-\alpha) = \frac{k_{AB} t_1 w \cos(\lambda-\alpha)}{\sin\phi\cos\theta} \\
F_T &= F_R \sin(\lambda-\alpha) = \frac{k_{AB} t_1 w \sin(\lambda-\alpha)}{\sin\phi\cos\theta}
\end{aligned}\right\} \tag{5.2}$$

From Eq. (5.2), the cutting forces can be predicted if the shear stress k_{AB}, shear angle ϕ, and the friction angle λ are known. Oxley (1989) proposed a theory to estimate k_{AB}, ϕ and λ. In this study, a new cutting force predictive model is based on this method, and a brief overview of it is given as follows. In the shear plane zone, k_{AB} is calculated according to the following equation:

$$k_{AB} = \frac{\sigma_1 \varepsilon_{AB}^{\ n}}{\sqrt{3}} \tag{5.3}$$

where σ_1 and n are constants, and n is also called the strain-hardening index. In reality, σ_1 and n vary with strain rate and temperature (Usui, 1982). For a combination of specific strain-rate and temperature, σ_1 and n give a good fit in the following empirical stress-strain relation:

$$\sigma = \sigma_1 \varepsilon^n \tag{5.4}$$

where σ and ε are the effective flow stress and strain. In order to find the values of σ_1 and n, it is necessary to know the value of the velocity-modified temperature T_{mod}, which is defined as a function of strain rate $\dot{\varepsilon}$ and temperature T_{AB} as:

$$T_{\text{mod}} = T_{AB}[1 - v \lg(\dot{\varepsilon} / \dot{\varepsilon}_0)] \tag{5.5}$$

the constants v and $\dot{\varepsilon}_0$ are taken as 0.09 and 1/sec, respectively.

From Eqs. (5.3) to (5.5), in order to determine k_{AB}, three parameters are needed, which are the temperature at AB T_{AB}, together with the strain ε_{AB} and the strain-rate $\dot{\varepsilon}_{AB}$ at AB. In (Oxley, 1989), the strain and strain-rate are calculated as:

$$\varepsilon_{AB} = \gamma_{AB} / \sqrt{3} \tag{5.6}$$

$$\dot{\varepsilon}_{AB} = \dot{\gamma}_{AB} / \sqrt{3} \tag{5.7}$$

where γ_{AB} and $\dot{\gamma}_{AB}$ are maximum strain and strain-rate at AB, and they are assumed to be given by:

$$\gamma_{AB} = \frac{1}{2}\frac{\cos\alpha}{\sin\phi\cos(\phi-\alpha)} \tag{5.8}$$

$$\dot{\gamma}_{AB} = C'\frac{V_S}{l} \tag{5.9}$$

The temperature at the shear plane T_{AB} is calculated from the shear forces, shear velocities, and tool and workpiece material thermal properties (thermal conductivity and specific heat), shown as the following equation:

$$T_{AB} = T_W + \eta\Delta T_{SZ} \tag{5.10}$$

where T_W is the initial workpiece temperature and ΔT_{SZ} is calculated from the equation:

$$\Delta T_{SZ} = \frac{1-\beta}{\rho S t_1 w}\frac{F_S\cos\alpha}{\cos(\phi-\alpha)} \tag{5.11}$$

where β is obtained from the following equation:

$$\begin{aligned} &\beta = 0.5 - 0.35\lg(R_T\tan\phi) \quad for \quad 0.04 \le R_T\tan\phi \le 10.0 \\ &\beta = 0.3 - 0.15\lg(R_T\tan\phi) \quad for \quad R_T\tan\phi > 10.0 \end{aligned} \tag{5.12}$$

where R_T is non-dimensional thermal number, which is given by:

$$R_T = \rho S V_C t_1 / K \tag{5.13}$$

Because the cutting forces, thermal conductivity and specific heat of workpiece material are temperature-dependent, an initial temperature is assumed to start the calculation. For the given assumed temperature, the cutting forces and thermal properties of workpiece material are determined, which can be used to calculate the temperature at AB. Then, the temperature at AB is updated with the calculated one. With the replacement of calculated temperature as the starting temperature, a new calculation starts. This process is repeated until the difference between the starting temperature and the calculated one is less than a given value, say, 0.1K.

For calculation of chip temperatures given by the following equation, a similar iterative procedure is also necessary.

$$T_{int} = T_W + \Delta T_{SZ} + \psi \Delta T_M \tag{5.14}$$

where ΔT_M is calculated from the equation:

$$\lg(\frac{\Delta T_M}{\Delta T_C}) = 0.06 - 0.195\delta(\frac{R_T t_2}{l_c})^{1/2} + 0.5\lg(\frac{R_T t_2}{l_c}) \tag{5.15}$$

where the average temperature rise in the chip ΔT_C is given by

$$\Delta T_C = \frac{F_f \sin\phi}{\rho S t_1 w \cos(\phi - \alpha)} \tag{5.16}$$

At first, the mean chip temperature is estimated for the initial calculation. Then, it is replaced with the new calculated value. This process continues until the difference between the estimated and calculated values is less than a certain value. Then, Eq. (5.5) is used to calculate the value of T_{mod} at the tool-chip interface with the temperature T_{int} and the average strain-rate $\dot{\varepsilon}_{\text{int}}$ at the tool-chip interface. The value of $\dot{\varepsilon}_{\text{int}}$ is calculated as:

$$\dot{\varepsilon}_{\text{int}} = \dot{\gamma}_{\text{int}} / \sqrt{3} = \frac{V_{chip}}{\sqrt{3}\delta t_2} \tag{5.17}$$

Finally the shear flow stress in the chip at the tool-chip interface k_{chip} is given by:

$$k_{chip} = \frac{\sigma_1}{\sqrt{3}} \tag{5.18}$$

where σ_1 is determined by the value of T_{mod} at the interface. In addition, the resolved shear stress at the tool-chip interface is found from the following equation:

$$\tau_{\text{int}} = \frac{F_f}{l_c w} \tag{5.19}$$

where l_c is the tool-chip contact length given as:

$$l_c = \frac{t_1 \sin\theta}{\cos\lambda \sin\phi}\left\{1 + \frac{C'n}{3[1+2(\pi/4-\phi)-C'n]}\right\} \tag{5.20}$$

which is derived by considering moments about B of the normal stresses on AB to find the position of F_R. Assuming that normal stress on the tool-chip interface is distributed uniformly, the resultant force F_R intersects the tool-chip interface with a distance $l_c/2$ from B. And the angle θ is given by the following equation:

$$\tan\theta = 1 + 2(\pi/4 - \phi) - C'n \tag{5.21}$$

By applying the appropriate stress equation along AB, it can be shown that ϕ should be in the range of $(0, \pi/4)$.

Based on above description, for given values of tool rake angle α, the cutting speed V_C, the thickness t_1 and width of cut w of the undeformed chip, together with the thermal and flow stress properties of the workpiece material and the initial temperature of the work T_w (say, 20°C in all calculations), the following procedure is used. For a given δ and C', the equilibrium (when τ_{int} is equal to the value of k_{chip}) values of ϕ are found, as shown in Fig. 5.2. If there is more than one shear angle ϕ which satisfied the equilibrium condition, the higher or highest value of ϕ is chosen as the solution. Then, the required value of C' is determined from the stress boundary condition at B of Fig. 5.1. According to the stress equilibrium equation (Oxley, 1989) and the theory in which the tool-chip interface is assumed to be a direction of maximum shear stress, the normal stress σ_N' at B is given by:

$$\frac{\sigma_N'}{k_{AB}} = 1 + \frac{\pi}{2} - 2\alpha - 2C'n \tag{5.22}$$

Based on the assumption of uniform normal stress along the tool-chip interface, the normal stress σ_N at B is also given by:

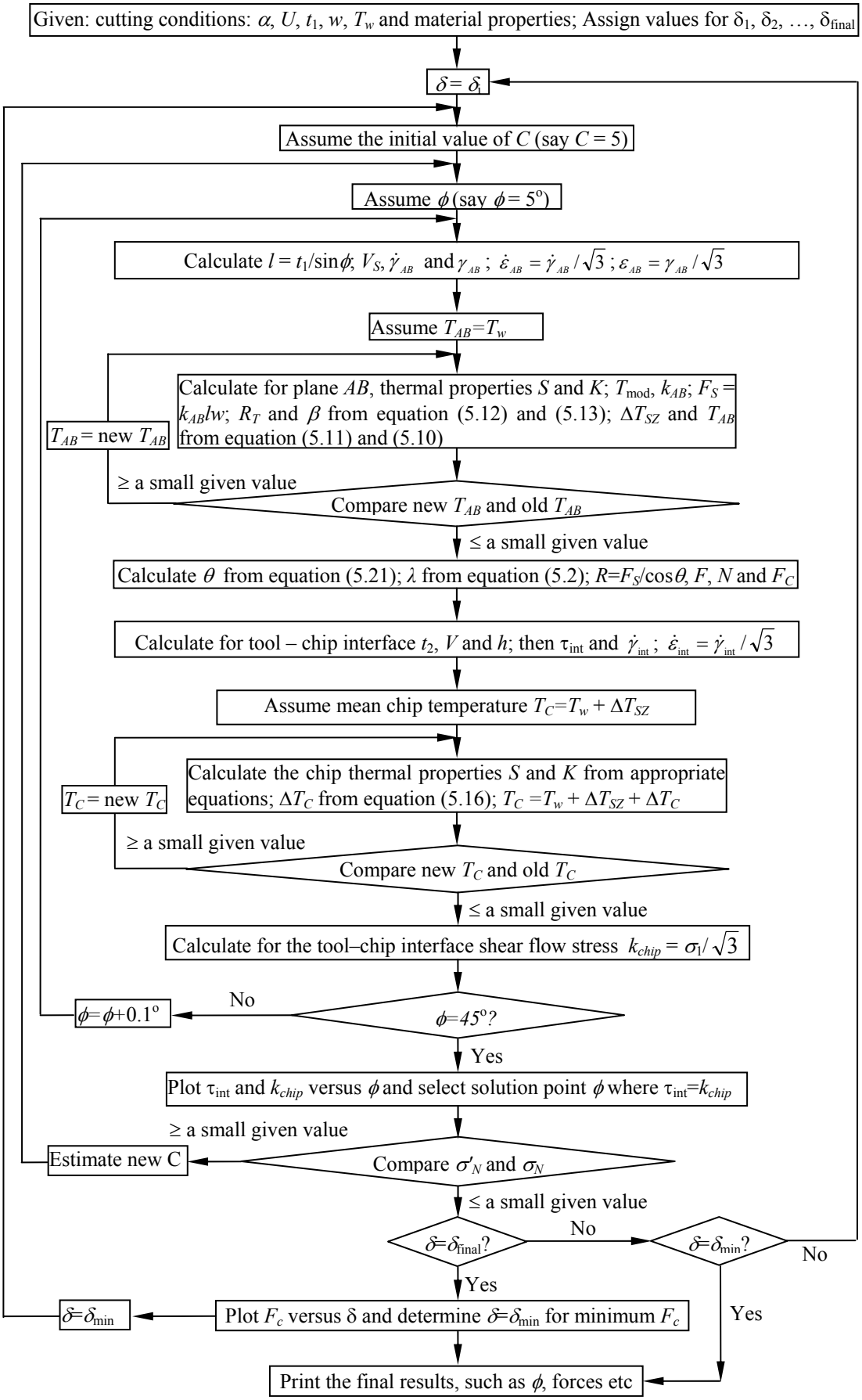


Figure 5.2 Diagram of the machining calculation with Oxley's theory

$$\sigma_N = \frac{F_N}{wl_c} \tag{5.23}$$

Then, C' can be found by fulfilling the condition $\sigma_N' = \sigma_N$.

The above procedure is iterated for a given range of δ and C' until all the equilibrium conditions are fulfilled. For Oxley's machining method, three parameters are very important for its accuracy: shear angle, strain-rate constant, the ratio of tool-chip interface plastic zone thickness to chip thickness. For practical computation, if the actual values C' and δ are not given in the possible range. It is very difficult to use this method to find the accurate shear angle ϕ. Furthermore, this method is based on the assumptions, such as the uniform distribution of normal stress at the chip-tool interface, resultant force intersects the tool face at a distance of $\overline{BG}$ /2 from B. These assumptions will reduce the accuracy of the final value of ϕ to some extent.

In this study, an improved method based on Oxley's theory is proposed. In this model, the shear angle and strain-rate constant are estimated based on the finite element simulation. With this method, it can ensure more accurate estimation of these two parameters.

5.2 Geometrical modeling of milling process

In the geometrical modeling of chip formation for the milling operation, the circular tooth path is widely used with the traditional method. However, for high-speed milling, the errors with this method are becoming more and more significant and cannot be ignored. In this section, a new method for modeling of milling process geometry is proposed, which is based on the analysis of the true tooth trajectory.

Unlike in turning processes, in slot milling, the relative movement between the tool and workpiece along the feed direction is achieved by the rotational motion and

translational feed motion of the spindle. So the motion of the cutter is like the trace of a point fixed on a circle that rolls along a line. If assumed that there is no tool wear, tool eccentricity and spindle tilt, the trajectory of cutting points relative to the workpiece is a trochoid curve, as shown in Fig. 5.3 (a), where r is the radius of the tool holder and R is the radius of the cutter. This trochoid curve results in that the instantaneous chip thickness varies periodically as a function of time-varying immersion.

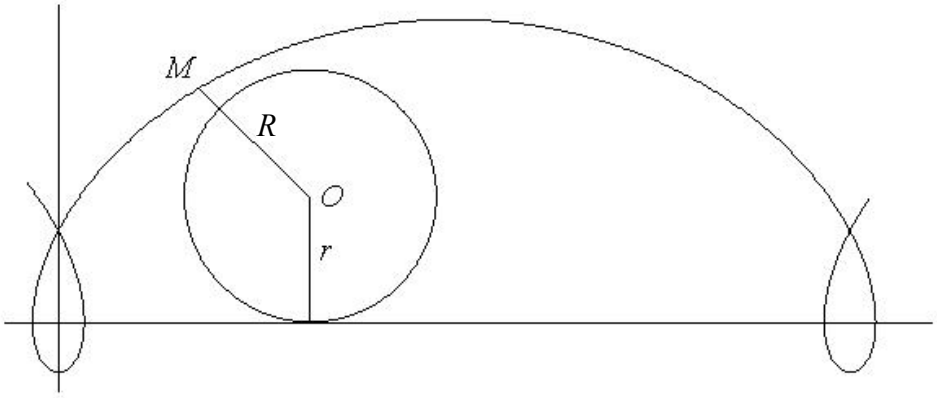
In milling, the undeformed chip thickness is determined by the distance between two consecutive cut surfaces. This distance varies during the chip formation of a cutting tooth and is measured in the direction perpendicular to the preceding cut surface (Kaczmarek, 1976). In this study, the inserted cutters were used in slot milling, so the helix angle is considered to be zero. For the cutter without helical angle, the undeformed chip thickness is uniform along each tooth and varies with its instantaneous angle of immersion.

In this study, only one flat ending milling insert is used. For the convenience of analysis, the point on the tip of the insert is selected to describe the true tooth trajectory. Fig. 5.3 (b) shows the simplified milling process model and coordinate system for the two consecutive cut, where f is the feed rate (mm/rev-tooth), ϕ is the instantaneous angle of immersion and $R = D/2$, D is the diameter of the cutter. The true tooth trajectories of two consecutive cut are two solid curves in Fig. 5.3 (b), and the undeformed chip thickness is decided by these two curves. The position of the center point of the cutter locates at point O and O', respectively when these two consecutive cut starts. When the tool tip is at point B for the second cut, the angular position of the

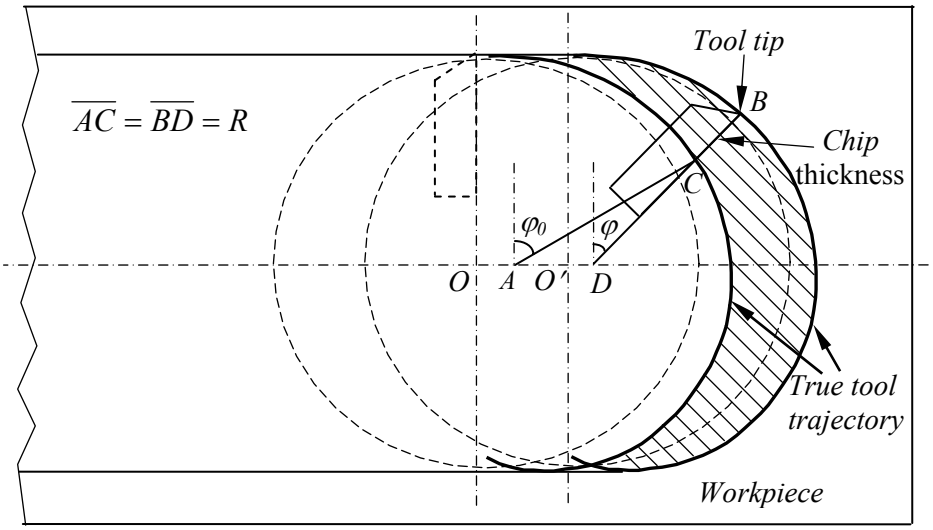
insert is φ. At this time, the undeformed chip thickness depends on coordinates of points B and C, and the position of point C is determined by the tool tip with the angular position φ_0 for the first cut. So when $0 < \varphi \leq 90^o$, the following equation can be derived:

$$\overline{OA} = \frac{\varphi_0}{2\pi} f \tag{5.24}$$

$$\overline{O'D} = \frac{\varphi}{2\pi} f \tag{5.25}$$



(a) Traditional trochoid curve



(b) The true tool trajectory during slot milling

Figure 5.3 Geometry of chip thickness of the milling process

Then, the length of *AD* can be obtained as following:

$$\overline{AD} = \overline{O'A} + \overline{O'D} = (f - \frac{\varphi_0}{2\pi} f) + \frac{\varphi}{2\pi} f = \frac{2\pi - \varphi_0 + \varphi}{2\pi} f \tag{5.26}$$

Using the law of sines for the triangle ΔACD, the following equation is obtained:

$$\frac{\sin(\pi/2 + \varphi)}{R} = \frac{\sin(\varphi_0 - \varphi)}{\overline{AD}} = \frac{\sin(\pi/2 - \varphi_0)}{\overline{CD}} \tag{5.27}$$

Let $\delta = \varphi_0 - \varphi$, so $\varphi_0 = \varphi + \delta$, then we get:

$$\frac{\sin(\pi/2 + \varphi)}{R} = \frac{\sin(\delta)}{\frac{2\pi - \delta}{2\pi} f} \tag{5.28}$$

After simplification, the following relationship is obtained:

$$(2\pi - \delta)a = 2\pi \sin(\delta),\ 2\pi \sin(\delta) + a\delta - 2\pi a = 0 \tag{5.29}$$

where $a = \frac{f}{R} \sin(\frac{\pi}{2} + \varphi)$.

After solving Eq. (5.29), the explicit expression of δ is given as:

$$\delta = a - \frac{1}{2\pi} a^2 + \frac{3 + 2\pi^2}{12\pi^2} a^3 - \frac{3 + 8\pi^2}{24\pi^3} a^4 + \frac{15 + 100\pi^2 + 18\pi^4}{240\pi^4} a^5 + O(a^6) \tag{5.30}$$

Then, the length of *CD* can be expressed in terms of δ and φ as:

$$\overline{CD} = \frac{\sin(\pi/2 - \varphi - \delta)}{\sin(\pi/2 + \varphi)} R \tag{5.31}$$

Finally, the instantaneous chip thickness at the position defined by angle φ is given by:

$$\overline{BC} = R - \overline{CD} = \frac{\sin(\pi/2 + \varphi) - \sin(\pi/2 - \varphi - \delta)}{\sin(\pi/2 + \varphi)} R \tag{5.32}$$

When $\varphi = 90^o$, the instantaneous chip thickness is equal to *f*; and when $\varphi > 90^o$, the following equations can be derived in the similar way, using the law of sines

$$\frac{\sin(\varphi - \pi/2)}{R} = \frac{\sin(\varphi - \varphi_0)}{\overline{AD}} = \frac{\sin(\varphi_0 - \pi/2)}{\overline{CD}} \tag{5.33}$$

Let $\delta = \varphi - \varphi_0$, so $\varphi_0 = \varphi - \delta$, and the following equation:

$$\frac{\sin(\varphi-\pi/2)}{R}=\frac{\sin(\delta)}{\frac{2\pi+\delta}{2\pi}f} \tag{5.34}$$

After simplification, the following relationship is obtained:

$$2\pi\sin(\delta)-a\delta-2\pi a=0 \tag{5.35}$$

where $a=\frac{f}{R}\sin(\varphi-\frac{\pi}{2})$

Then the instantaneous chip thickness $h(\varphi)$ at the position defined by angle φ is given by:

$$h(\varphi)=\overline{BC}=R-\overline{CD}=\frac{\sin(\varphi-\pi/2)-\sin(\varphi-\delta-\pi/2)}{\sin(\varphi-\pi/2)}R \tag{5.36}$$

where $\delta=a+\frac{1}{2\pi}a^2+\frac{3+2\pi^2}{12\pi^2}a^3+\frac{3+8\pi^2}{24\pi^3}a^4+\frac{15+100\pi^2+18\pi^4}{240\pi^4}a^5+O(a^6)$

After obtaining analytical solution to the undeformed chip thickness, the error percentage caused by the circular tool-path has been analyzed, as shown in Fig. 5.4. Obviously, this circular tool-path causes some error. In addition, the developed model for undeformed chip thickness can also be used for micro-milling process, where the circular tool-path causes more significant error.

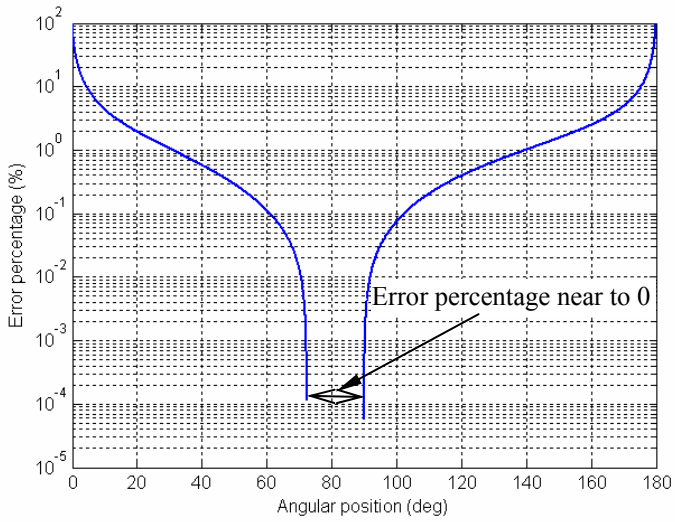


Figure 5.4 Error percentage caused by the traditional circular tool-path

5.3 Modeling for equivalent element representation

5.3.1 Effects of tool nose radius

Almost all the cutting tools used in industry have a nose radius. The tool nose radius can improve the surface finish obtained on the machined surface, and the strength and wear characteristics of the tool; and it will influence the chip flow direction and the cutting forces. It is of great practical importance to consider the effects of nose radius for accurate cutting force models. The process can be explained by taking a cutter with a nose radius r as shown in Fig. 5.5. The axial depth of cut is a, and the undeformed chip thickness $h(\varphi)$, which depends on the instantaneous position of the cutter and feed rate per tooth per revolution. If the axial depth of cut a, is larger than the tool nose radius r. The intersection surface of the chip can be divided into two regions, region I and region II, as shown in Fig. 5.5. The chip thickness in region I is equal to $h(\varphi)$. However, the chip thickness in region II is changing because of the nose radius of the tool.

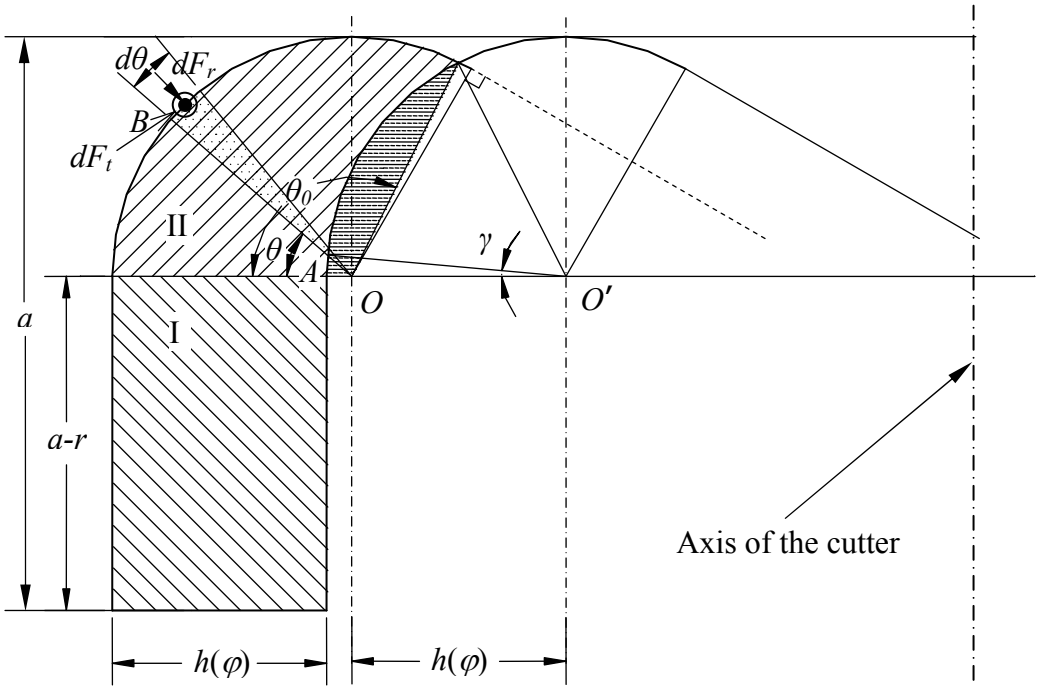


Figure 5.5 Mechanics of milling process with nose radius inserts ($a > r$)

In this study, high-speed slot milling of Ti-6Al-4V with a single indexable tool has been tried to investigate cutting mechanism. The indexable tool has a nose radius which is larger than the axial depth of cut. So only the nose of the tool is involved in cutting, and the process is illustrated in Fig. 5.6, which shows that the cutting action on the nose radius of the tool tip defines the surface finish of the machined surface.

For the convenience of calculation, the intersection surface of the chip is also divided into two regions as shown in Fig. 5.6.

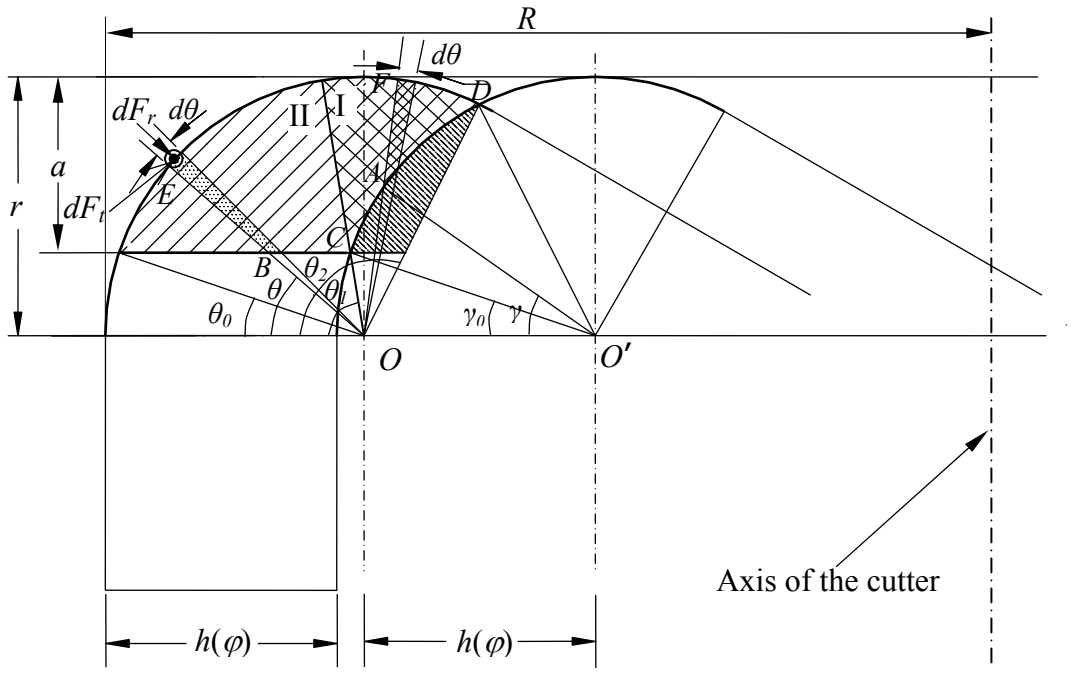


Figure 5.6 Mechanics of milling process with nose radius inserts ($a < r$)

In region I, using the law of sines for $\Delta AOO'$, where $\overline{OF}$ is defined by the angle θ.

$$\frac{\sin(\pi-\theta)}{r}=\frac{\sin[\pi-(\pi-\theta)-\gamma]}{h(\varphi)} \tag{5.37}$$

The following relationship is obtained:

$$\theta-\gamma=\arcsin[\frac{h(\varphi)}{r}\sin(\pi-\theta)] \tag{5.38}$$

so,

$$\gamma=\theta-\arcsin[\frac{h(\varphi)}{r}\sin(\pi-\theta)] \tag{5.39}$$

Then, the following can be obtained:

$$\begin{aligned}\cos\gamma &= \cos\{\theta - \arcsin[\frac{h(\varphi)}{r}\sin(\pi-\theta)]\} \\ &= \cos\theta\cos\{\arcsin[\frac{h(\varphi)}{r}\sin(\pi-\theta)]\} + \sin\theta\frac{h(\varphi)}{r}\sin(\pi-\theta) \\ &= \cos\theta\sqrt{1-\frac{h^2(\varphi)}{r^2}\sin^2(\pi-\theta)} + \sin^2\theta\cdot\frac{h(\varphi)}{r}\end{aligned} \tag{5.40}$$

Using the law of cosine for $\Delta AOO'$, the length of AO is calculated as:

$$\begin{aligned}\overline{AO} &= \sqrt{r^2 + h^2(\varphi) - 2rh(\varphi)\cos\gamma} \\ &= \sqrt{r^2 + h^2(\varphi) - 2rh(\varphi)[\sin^2\theta\cdot\frac{h(\varphi)}{r} + \cos\theta\cdot\sqrt{1-\frac{h^2(\varphi)}{r^2}\sin^2(\pi-\theta)}]} \\ &= \sqrt{r^2 + h^2(\varphi) - 2h^2(\varphi)\sin^2\theta - 2rh(\varphi)\cos\theta\cdot\sqrt{1-\frac{h^2(\varphi)}{r^2}\sin^2(\pi-\theta)}} \\ &= \sqrt{r^2 - 2rh(\varphi)\cos\theta\cdot\sqrt{1-\frac{h^2(\varphi)}{r^2}\sin^2(\pi-\theta)} + h^2(\varphi)\cos^2\theta - h^2(\varphi)\sin^2\theta} \\ &= \sqrt{(r^2 - h^2(\varphi)\sin^2\theta) - 2h(\varphi)\cos\theta\cdot\sqrt{r^2 - h^2(\varphi)\sin^2\theta} + h^2(\varphi)\cos^2\theta} \\ &= \sqrt{r^2 - h^2(\varphi)\sin^2\theta} - h(\varphi)\cos\theta\end{aligned} \tag{5.41}$$

If the ratio $h(\varphi)/r$ is very small, then the factor $h^2(\varphi)/r^2$ can be neglected, so the following equation is obtained:

$$\overline{AO} = r - h(\varphi)\cos\theta \tag{5.42}$$

So the corresponding area of differential (dA) chip can be approximated by:

$$dA = (r^2 - \overline{AO}^2)d\theta/2 \tag{5.43}$$

In region II, the differential chip area at the position defined by position θ is calculated with the following equation:

$$dA = (r^2 - \overline{OB}^2)d\theta/2 \tag{5.44}$$

Finally the whole intersection surface of the chip S_φ defined by chip thickness $h(\varphi)$ can be described as following:

$$S_\varphi = \frac{1}{2}\int_{\theta_0}^{\theta_1}(r^2 - \overline{OB}^2)d\theta + \frac{1}{2}\int_{\theta_1}^{\theta_2}(r^2 - \overline{AO}^2)d\theta = \frac{1}{2}\int_{\theta_0}^{\theta_1}(r^2 - \frac{(r-a)^2}{\sin^2\theta})d\theta +$$

$$\frac{1}{2}\int_{\theta_1}^{\theta_2}[h^2(\varphi)\sin^2\theta + 2h(\varphi)\cos\theta\sqrt{r^2 - h^2(\varphi)\sin^2\theta} - h^2(\varphi)\cos^2\theta]d\theta \quad (5.45)$$

After simplifying, S_φ can be expressed as following:

$$S_\varphi = \frac{1}{2}r^2(\theta_1 - \theta_0) + \frac{1}{2}(r-a)^2\frac{\cos\theta}{\sin\theta}\Big|_{\theta_0}^{\theta_1} - \frac{1}{2}h^2(\varphi)\sin\theta\cos\theta\Big|_{\theta_1}^{\theta_2} +$$

$$\frac{1}{2}h\sin\theta\sqrt{r^2 - (h(\varphi)\sin\theta)^2}\Big|_{\theta_1}^{\theta_2} + \frac{1}{2}r^2\arctan(\frac{h\sin\theta}{\sqrt{r^2 - (h(\varphi)\sin\theta)^2}})\Big|_{\theta_1}^{\theta_2} \quad (5.46)$$

where, $\sin\theta_0 = \frac{r-a}{r}$ so $\theta_0 = \arcsin\frac{r-a}{r}$

$$\sin\theta_1 = \frac{r-a}{\overline{CO}} = \frac{r-a}{\sqrt{r^2 + h^2(\varphi) - 2rh(\varphi)\cos\gamma_0}} \text{ and } \gamma_0 = \theta_0$$

$$\theta_1 = \arcsin\frac{r-a}{\sqrt{r^2 + h^2(\varphi) - 2rh(\varphi)\cos\theta_0}}$$

$$\theta_2 = \pi - \angle DOO' = \pi - \arccos\frac{h(\varphi)}{2r}$$

5.3.2 Equivalent elements of the real chips

As discussed in the above section, different parts of the nose involved in cutting are under different cutting load. Therefore, in order to investigate the effects of the nose and predict cutting process variables in slot milling, it is needed to simulate the cutting process around the tool tip. For this purpose, it needs to establish a model about 3-D milling process, which requires longer computation time even if for computers with faster computation speed. An alternative approach is to represent the 3-D milling with 2-D metal deformation process, which is also used by other researchers (Stephenson

and Bandyopadhyay, 1997; Ozel, 1998; Ozel and Altan, 2000). The main idea about the equivalent element representation is that the uneven intersection surface of the undeformed chip has been changed to an equivalent one which is suitable for 2-D deformation simulation. Normally, there are two types of 2D geometry representations: axisymmetric and plane strain. In (Stephenson and Bandyopadhyay, 1997), the uneven uncut chip area has been changed into the equivalent rectangular contact area. Ozel (1998) transformed the uneven intersection surface of the undeformed chip to the equivalent one which is suitable for axisymmetric deformation simulation. An equivalent axisymmetric intersection of the chip with uniform thickness is shown in Fig. 5.7. Then, the engagement angle between the nose radius and the chip θ_e can be derived as following:

$$\theta_e = \pi/2 - \theta_0 + \theta_3 \tag{5.47}$$

where $\theta_0 = \arcsin\dfrac{r-a}{r}$, and $\theta_3 = \arcsin(\dfrac{h(\varphi)}{2r})$

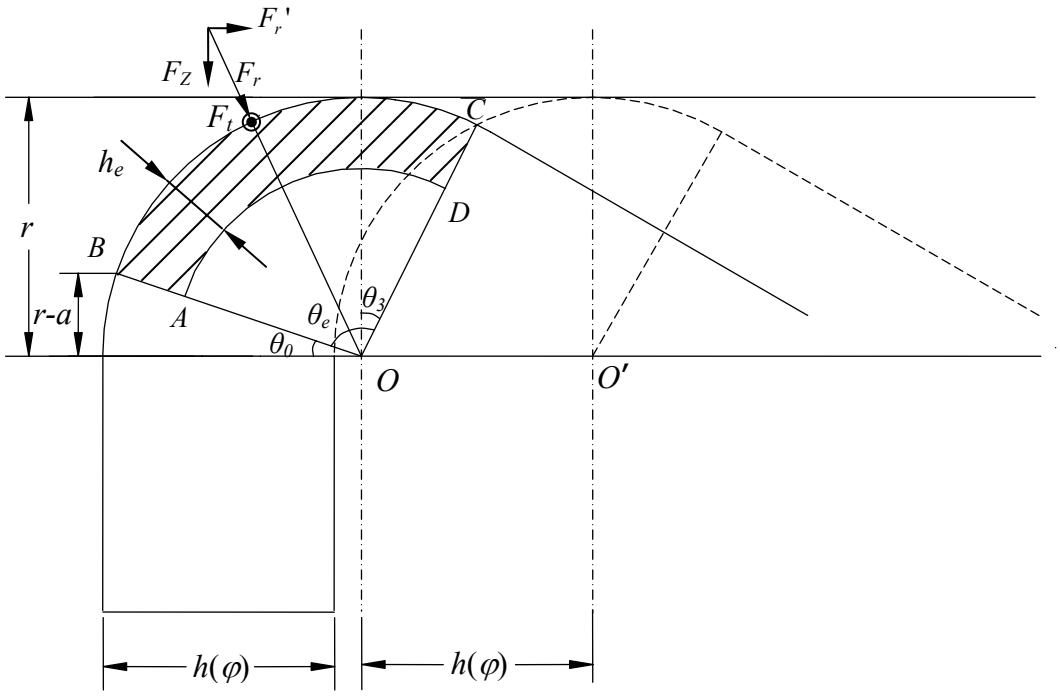


Figure 5.7 Equivalent chip element with uniform chip thickness by Ozel's method

After θ_e is obtained, the equivalent chip thickness for the corresponding equivalent element can be found as:

$$h_e = r - \sqrt{r^2 - 2\frac{S_\varphi}{\theta_e}} \tag{5.48}$$

In this study, because the axial depth of cut and feed rate are very small, the maximum equivalent chip thickness (around 0.025mm calculated with Ozel's method) is less than 0.05mm for all cases. As discussed in (Oxley, 1989), when the chip thickness is less than 0.05mm, the size effect is very obvious. More importantly, special care was taken to ensure that the cutting edge radius was much smaller than the smallest value of undeformed chip thickness. In this study, the radius of the cutting edge is 0.015 mm, so it is difficult to claim that it is much smaller than the undeformed chip thickness, whose maximum value is 0.025 mm. Thus, Ozel's transformation method is not suitable in this study.

Based on the observation of the geometry of uncut chip area, a new equivalent representation idea was used, as shown in Fig. 5.8. From Fig. 5.8 (a), it can be seen that the area of region I is equal to that of region II, so the non-uniform uncut area can be represented with its equivalent element of Fig. 5.8 (b). Then, the equivalent area is further transformed into a parallelogram ($A'BCD'$) as shown in Fig. 5.8 (c), where a_e is equal to the uncut chip area divided by undeformed chip thickness $h(\varphi)$. In Fig. 5.8 (c), the engagement angle between the nose radius and the chip θ_e can be derived in the same as that shown in Fig. 5.7: $\theta_e = \pi/2 - \theta_0 + \theta_3$, where $\theta_0 = \arcsin\frac{r-a}{r}$, and $\theta_3 = \arcsin(\frac{h(\varphi)}{2r})$. Thus, the undeformed chip thickness around the nose is represented by the equivalent chip of uniform chip thickness. The equivalent chip around the nose

radius can be simulated as a facing process in turning with depth of cut a_e and the changing feed-rate $h(\varphi)$, as shown in Fig. 5.9.

Therefore, with equivalent element representation, the flat slot milling process can be illustrated as turning process with changing feed rate $h(\varphi)$ as shown in Fig. 5.10. After this conversion, the plane strain deformation of the turning process can be simulated with FEM. The cutting forces components F_t (cutting speed direction) and F_r (thrust direction) can be predicted from the plane strain simulation.

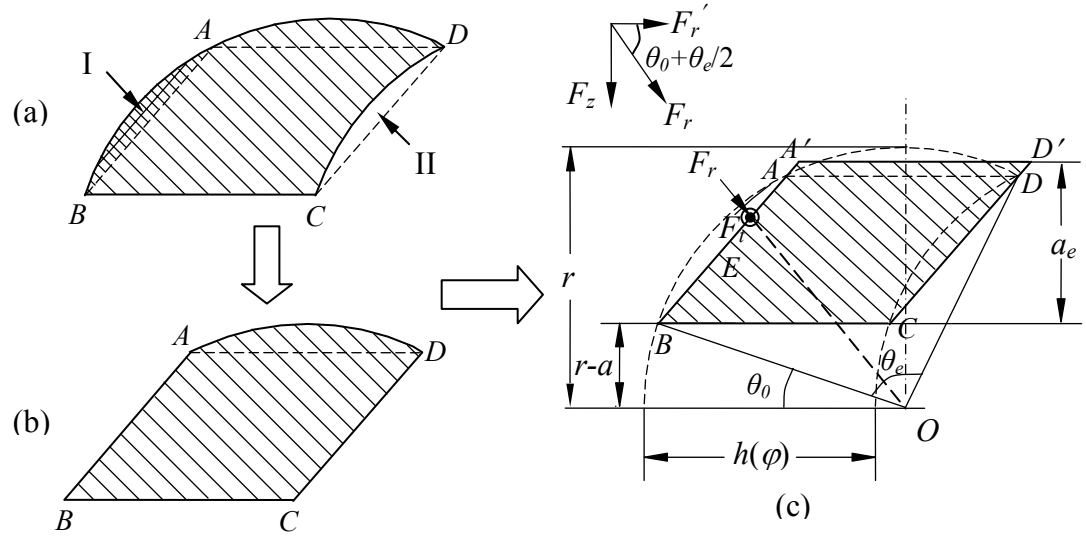


Figure 5.8 Equivalent chip element with uniform chip thickness

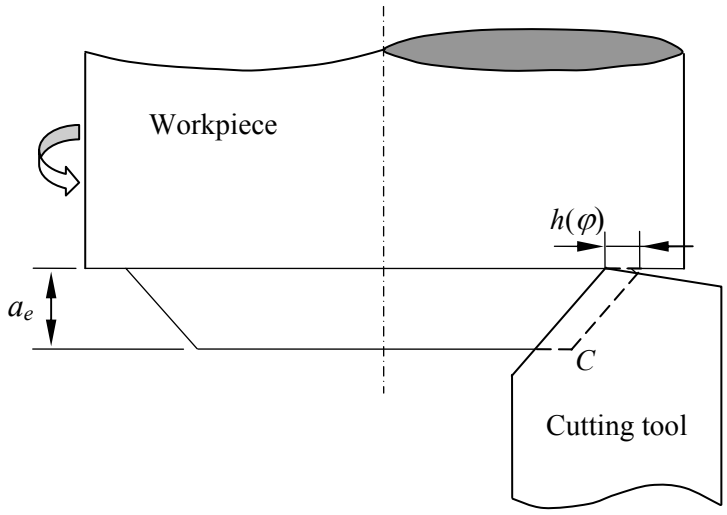


Figure 5.9 Equivalent facing process in turning with uniform chip thickness

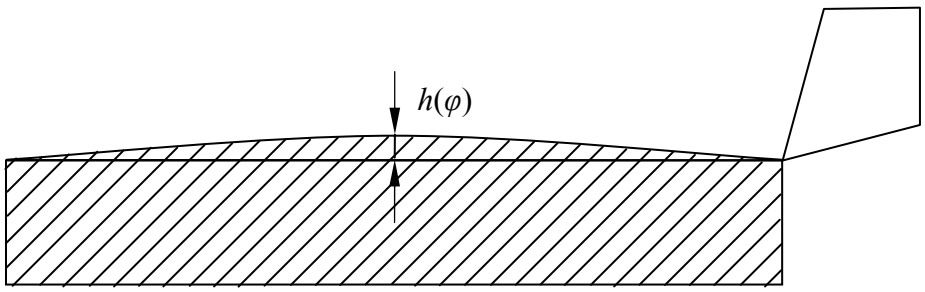


Figure 5.10 Illustration of FEM simulation of slot milling with a nose radius tool with a facing process in turning

5.3.3 Formulation of cutting forces

As described above, from the geometry relationship in Fig. 5.8 (c), the cutting forces components F_z and F_r' can be obtained as:

$$F_z = F_r \sin(\theta_0 + \frac{\theta_e}{2}) \tag{5.49}$$

$$F_r' = F_r \cos(\theta_0 + \frac{\theta_e}{2}) \tag{5.50}$$

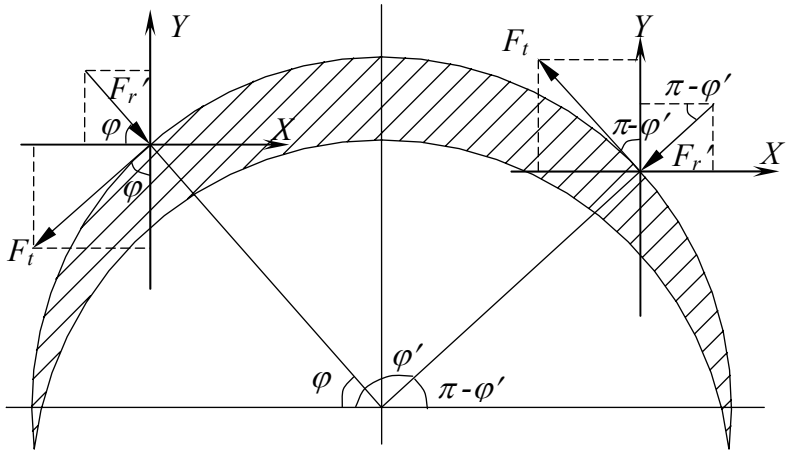


Figure 5.11 Mechanics of milling process

Then, the horizontal (x direction) and normal (y direction) components of the cutting forces acting on the cutter are derived from the equilibrium diagram shown in Fig. 5.11:

$$F_x = -F_t \sin\varphi + F_r' \cos\varphi = -F_t \sin\varphi + F_r \cos\varphi \sin(\theta_0 + \frac{\theta_e}{2}) \tag{5.51}$$

$$F_y = -F_t \cos\varphi - F_r' \sin\varphi = -F_t \cos\varphi - F_r \sin\varphi \sin(\theta_0 + \frac{\theta_e}{2}) \quad (5.52)$$

Normally, for tangential ($F_t(\varphi)$) and radial ($F_r(\varphi)$) cutting forces acting on the cutter in slot milling, the radial component F_r is assumed to be ξF_t, where ξ is a constant. Assumed that $\tan\varphi_0 = \xi \sin(\theta_0 + \frac{\theta_e}{2})$, then

$$F_x = -F_t \sin\varphi + F_t \tan\varphi_0 \cos\varphi = -\frac{F_t}{\cos\varphi_0} \sin(\varphi - \varphi_0) \quad (5.53)$$

$$F_y = -F_t \cos\varphi - F_t \tan\varphi_0 \sin\varphi = -\frac{F_t}{\cos\varphi_0} \cos(\varphi - \varphi_0) \quad (5.54)$$

Because of the effect of φ_0, F_x and F_y cannot reach their maximum values when F_r and F_t nearly reach their maximum ones at $\varphi \approx \pi/2$. The varying values of F_x and F_y are simulated in the next section.

In the practical experiments, the cutting forces in x, y and z direction are recorded, it needs to derive the cutting force components F_r' and F_t from measured forces F_x, F_y and F_Z. From Fig. 11, the derivation of F_r' and F_t is as follows:

$$F_r' = F_x \cos\phi - F_y \sin\phi \quad (5.55)$$

$$F_t = -F_x \sin\phi - F_y \cos\phi \quad (5.56)$$

For the mechanical cutting force model, the cutting forces in the tangential and radial direction are obtained by multiplying the specific cutting pressures with the uncut chip area.

$$F_t(\varphi) = K_t A_c(\varphi) \quad (5.57)$$

$$F_r(\varphi) = K_r A_c(\varphi) \quad (5.58)$$

After get the cutting forces from simulation, the specific cutting pressures in the tangential and radial direction can be estimated from above two equations. In Chapter

7, the estimated specific cutting pressures will be used to determine the cutting force and machining power constraints.

5.4 Prediction of cutting forces in slot milling

5.4.1 Modeling of flow stress properties of Ti-6Al-4V

The yield stress of metal under uniaxial conditions is defined as the flow stress or the effective stress, which depends on the strain and strain-rate, material properties and cutting temperature. The metals start deforming plastically when the applied stress reaches the values of the yield stress or flow stress (Altan, 1983).

The flow stress is mostly influenced by temperature, strain, strain-rate and material properties. Accurate and reliable flow stress models are very important for describing workpiece materials' deforming behavior during practical machining processes. Many researchers developed several techniques to determine the flow stress of metals.

The widely used constitutive model of flow stress is Johnson-Cook (JC) strength model, which was proposed by Johnson and Cook (1983). The JC model represents the flow stress $\bar{\sigma}$ of a material as the product of strain, strain-rate and temperature, shown as the following equation:

$$\bar{\sigma} = [A + B(\bar{\varepsilon})^n][1 + C\ln(\frac{\dot{\bar{\varepsilon}}}{\dot{\varepsilon}_0})][1 - (\frac{T - T_r}{T_m - T_r})^m] \tag{5.59}$$

The parameter A is the initial yield strength of the material at room temperature and a strain-rate of 1 s^{-1}; $\bar{\varepsilon}$ is the equivalent plastic strain; $\dot{\bar{\varepsilon}}$ is the strain-rate normalized by a reference strain-rate $\dot{\varepsilon}_0$. The temperature term is valid within the range from room temperature (T_r) to melting temperature of the workpiece material (T_m). The

parameters B, C, m and n are fitted to the experimental results obtained from the corresponding compression and tension tests. This model sacrifices the potential coupling of effects, such as temperature effects on strain-rate hardening, but it can be calibrated more easily. Therefore, some researchers chose the JC model as constitutive equation for deformation behavior of metals at higher strain rate and high temperature. Lee and Lin (1998a) investigated the deformation behavior of Ti-6Al-4V using the split Hopkinson bar (SHPB). They fitted SHPB test results into the JC model at the strain-rate of 2×10^3 s^{-1}. Meyer Jr. and Kleponis (2001) also studied high strain rate behavior of Ti-6Al-4V and low cost titanium.

Another constitutive model for metals is the Zerilli-Armstrong (ZA) strength model, which was proposed by Zerilli and Armstrong (1987). The ZA model is based on dislocation mechanics theory. There are two forms of the model, one to address face-centered cubic (fcc) materials, and the other to address body-centered cubic (bcc) materials. For the bcc case, the flow stress is computed from:

$$\overline{\sigma} = C_0 + C_1 e^{(-C_3 T + C_4 T \ln \dot{\overline{\varepsilon}})} + C_5 \overline{\varepsilon}^n \tag{5.60}$$

where T is absolute temperature, $\overline{\varepsilon}$ is the equivalent plastic strain, $\dot{\overline{\varepsilon}}$ is the strain-rate, and C_0, C_1, C_3, and C_4 are fitting constants. C_5 and n are the same as parameters B and n from the JC model. On the right hand side of Eq. (5.60), the effects of the strain rate and temperature on the flow stress are coupled together, which is different from those of the JC model. The coupling relationship between parameters makes it difficult to fit the parameters of the ZA model. Without additional constraints for the values of the parameters, many combinations of parameter values result in similar values for the total error (Meyer Jr. and Kleponis, 2001). In addition, Zerili and Armstrong (1987) did not recommend their model for temperature above about one-half of the absolute

melting temperature. When machining Ti-6Al-4V at high cutting speed, the maximum cutting temperature is somehow greater than 965K, one-half of the absolute melting temperature of Ti-6Al-4V. Therefore, in this study, the JC model is chosen to describe the deform behavior of Ti-6Al-4V.

Based on the published data listed in (Lee and Lin, 1998a; Lee and Lin, 1998b; Majorell et al., 2002), the Gauss-Newton algorithm with Levenberg-Marquardt modifications for global convergence was used to find the parameter estimates for the JC model. The estimated parameters are listed in Table 5.1, and those parameters for JC model found by Lee and Lin (1998a) and Meyer Jr. and Kleponis (2001) are also shown in Table 5.1 for the reference.

Table 5.1 Parameters of JC constitutive model for Ti-6Al-4V

Reference	*A*	*B*	*n*	*C*	*m*	Test
Lee and Lin (1998)	782.7	498.4	0.28	0.028	1.0	SHPB
Meyer and Kleponis (2001)	862.5	331.2	0.34	0.0120	0.8	SHPB
In this study	1165.5	236.6	0.29	0.0355	0.42	−

5.4.2 Modeling of cutting forces

In Oxley's model (1989), flow stress in the shear plane zone k_{AB} can be calculated according to Eq. (5.3) ($k_{AB} = \sigma_1 \varepsilon_{AB}{}^{n} / \sqrt{3}$). This is replaced by: $k_{AB} = \bar{\sigma} / \sqrt{3}$, where $\bar{\sigma}$ is the effective flow stress along *AB*, which can be calculated using the constitutive equation Eq. (5.59).

Based on the assumption of uniform normal stress distribution on the tool-chip interface, the value of C′ is calculated from equilibrium condition given in Eqs. (5.22) and (5.23) in Oxley's theory. In this study, according to the chip formation model from

(Oxley, 1989) and the constitutive equation of flow stress for Ti-6Al-4V, the value of C' is derived as follows.

The change rate of flow stress (dk/ds_2) normal to AB can be assumed to be only related to the actual strain-rate. Therefore, dk/ds_2 can be derived using the following equation (Oxley, 1989):

$$\frac{dk}{ds_2} = \frac{dk}{d\gamma}\frac{d\gamma}{ds_2} = \frac{dk}{d\gamma}\frac{d\gamma}{dt}\frac{dt}{ds_2} \tag{5.61}$$

where t is time. From Eq. (5.60), the following equation can be obtained at AB:

$$\frac{d\bar{\sigma}}{d\bar{\varepsilon}} = nB(\bar{\varepsilon}_{AB})^{n-1}[1 + C\ln(\frac{\dot{\bar{\varepsilon}}}{\dot{\bar{\varepsilon}}_0})][1 - (\frac{T - T_r}{T_m - T_r})^m] = \frac{nB(\bar{\varepsilon}_{AB})^{n-1}}{[A + B(\bar{\varepsilon}_{AB})^n]}\bar{\sigma}_{AB} \tag{5.62}$$

Then, the first term on the right-hand side of Eq. (5.61) can be obtained as:

$$\frac{dk}{d\gamma} = \frac{d\bar{\sigma}/\sqrt{3}}{\sqrt{3}d\bar{\varepsilon}} = \frac{nB(\bar{\varepsilon}_{AB})^{n-1}}{\sqrt{3}[A + B(\bar{\varepsilon}_{AB})^n]}k_{AB} \tag{5.63}$$

The second term on the right-hand side of Eq. (5.61) is the strain-rate, which is given in Eq. (5.9). The last term is the reciprocal of the cutting speed normal to AB, which can be presented as:

$$dt/ds_2 = 1/(V_C\sin(\phi)) \tag{5.64}$$

By substituting for Eq. (5.61) with Eqs. (5.9), (5.63) and (5.64), the following relation is obtained, and a similar derivation procedure is given in Huang and Liang (2003):

$$\frac{dk}{ds_2} = \frac{k_{AB}nB(\bar{\varepsilon}_{AB})^{n-1}}{\sqrt{3}[A + B(\bar{\varepsilon}_{AB})^n]}\frac{C'V_C\cos\alpha}{\cos(\phi - \alpha)l}\frac{1}{V_C\sin\phi} = \frac{2k_{AB}C'nB(\bar{\varepsilon}_{AB})^n}{l[A + B(\bar{\varepsilon}_{AB})^n]} \tag{5.65}$$

According to the stress equilibrium equation along AB from Oxley (1989), the following relation exists:

$$dp = \frac{dk}{ds_2}ds_1 \tag{5.66}$$

By applying the equation along AB, substituting for dk/ds_2 from Eq. (5.65), the next equation is given:

$$p_A - p_B = \frac{2k_{AB}C'nB(\bar{\varepsilon}_{AB})^n}{[A + B(\bar{\varepsilon}_{AB})^n]} \tag{5.67}$$

where p_A and p_B are the hydrostatic stresses at points A and B, respectively. Finally the unknown parameter C' is given by:

$$C' = \frac{(p_A - p_B)[A + B(\bar{\varepsilon}_{AB})^n]}{2k_{AB}nB(\bar{\varepsilon}_{AB})^n} \tag{5.68}$$

In Oxley's theory, the angle θ made by the resultant force R with AB is expressed as:

$$\tan\theta = 1 + 2(\frac{\pi}{4} - \phi) - \frac{\Delta k}{2k_{AB}}\frac{l}{\Delta s_2} \tag{5.69}$$

With the empirical stress-strain relation $\sigma = \sigma_1\varepsilon^n$, Eq. (5.69) is represented as Eq. (5.21), while in this study, by substituting for Eq. (5.69) with Eq. (5.65), the following equation is obtained:

$$\tan\theta = 1 + 2(\frac{\pi}{4} - \phi) - C'n\frac{B\bar{\varepsilon}_{AB}^{\ n}}{A + B\bar{\varepsilon}_{AB}^{\ n}} \tag{5.70}$$

Based on above description, for given values of tool rake angle α, the cutting speed V_C, the thickness t_1 and width of cut w of the undeformed chip, together with the thermal and flow stress properties of the workpiece material and the initial temperature of the work T_w (say, 20°C in all calculations), FEM can be employed to predict the cutting forces.

5.4.3 Determination of the values of ϕ, k_{AB} and C' by FEM

Recently, the application of FEM has achieved great progress with the dramatic progress in both computer hardware and software. Many researchers have developed user-oriented general-purpose program for analysis procedures of FEM. In this study,

FEM is used to simulate the cutting process, which is an extension of FEM application in the analysis of metal-forming. During practical machining, cutting heat is generated. At high cutting temperature, workpiece materials properties will wary considerably. Thus, the consideration of temperature effects in the analysis of plastic deformation during machining process is very important. So in cutting process, workpiece materials deform viscoplastically under the cutting load. Although there still exists elastic deformation except viscoplastic one, viscoplastic strains outweigh elastic strains. Therefore, it is reasonable to assume the cutting tool and workpiece combination with a rigid-viscoplastic material behavior.

Kobayashi et al. (1989) found that the governing equations for viscoplastic deformation are formally identical to those of plastic deformation, except that the effective stress is a function of strain, strain-rate and temperature. For the rigid-plastic materials, the deformation process is a boundary-value problem. Solutions to the boundary-value problem are the velocity distribution that satisfies the governing equations and the boundary conditions. The boundary conditions are that the velocity vector u is prescribed on a part of surface S_u together with traction F on the remainder of the surface S_F. The governing equations can be expressed mathematically as the following two equations (Kobayashi et al., 1989):

$$\pi = \int_V \overline{\sigma}\dot{\overline{\varepsilon}} dV - \int_{S_F} F_i u_i dS \tag{5.71}$$

$$\delta\pi = \int_V \overline{\sigma}\delta\dot{\overline{\varepsilon}} dV - \int_{S_F} F_i \delta u_i dS + K\int_V \dot{\overline{\varepsilon}}_V \delta\dot{\overline{\varepsilon}}_V dV \tag{5.72}$$

where $\overline{\sigma}$ is the effective stress, $\dot{\overline{\varepsilon}}$ is the effective strain-rate, F_i represents the surface tractions, the stress exists in the region V, traction F_i is applied over the surface S, u_i is admissible velocity, $\dot{\overline{\varepsilon}}_V$ is the volumetric strain-rate, and K is a penalty constant, which is a very large positive constant.

In heat transfer analysis of deforming material, the energy balance equation can be expressed by (Kobayashi et al., 1989):

$$\int_V k_1 T_{,ii} \delta T dV - \int_V \rho c \dot{T} \delta T dV + \int_V k \sigma_{ij} \dot{\varepsilon}_{ij} \delta T dV = 0 \qquad (5.73)$$

where k_1 is the thermal conductivity; $T_{,ii}$ is used for $T_{,i,i}$ with the comma denoting differentiation and repeated subscript meaning summation (Laplace differential operator applied to temperature T); ρ is the density; c is the specific heat; k is the heat generation efficiency.

Equations (5.71), (5.72) and (5.73) are the basic equations for the finite-element discretization. FEM is used to obtain the closed form solution to the velocity distribution. Once the solution for the velocity field that satisfied the basic equation is obtained, then the corresponding stress can be calculated using the flow rule and the known mean stress distribution.

The velocity distribution can be calculated with the minimum work rate principle. According to this principle, the material should always flow in the path of least resistance. Because of the effects of cutting heat, the equations for the flow analysis and the temperature calculation are strong coupled. The iteration process is used to determine the velocity distribution together with the cutting temperature:

Step 1: Assume the initial temperature field;

Step 2: Calculate the initial velocity field at the initial temperature;

Step 3: Calculate the initial temperature-rate field and the quantity;

Step 4: Update the nodal point positions and effective strain for the next step;

Step 5: Use the velocity field at the previous step to calculate the temperature;

Step 6: Calculate a new velocity field with the solution from Step 5;

Step 7: Use the new velocity to calculate the second temperature field;

Step 8: Repeat steps for (6) and (7) until both have converged;

Step 9: Calculated the new temperature rate field;

Step 10: Repeat steps (4) – (9) until the desired deformation state is reached.

One main problem for FEM simulation is to determine the boundary condition. The boundary conditions along the cutting tool/chip interface are very complicated. It is extremely difficult to determine the frictional stress at the tool/workpiece interface. The frictional stress has been influenced by factors such as cutting speed, feed rate and rake angle etc. Numerous works have been done to investigate the friction mechanism at the tool/chip interface. Zorev (1963) considered that at the tool/chip interface there are two regions on the tool rake face: sticking region and sliding region, as shown in Fig. 5.12. In the sticking region whose length is l_p, frictional stress is known to be equal to the local shear stress (k_{chip}). In the sliding region $\overline{BC}$, shear stress decreases on the rake face, Coulomb's friction law can be applied to calculate the frictional stress. Ozel (1998) found that variable friction coefficients as a function of normal pressure at the tool rake surface can ensure more accurate results in FEM simulations. In this study, this friction model has also been used. For this friction model, the frictional stress τ_f at the tool/chip interface is defined as:

$$\tau_f = \begin{cases} k_{chip}, & 0 \le x \le l_p \\ \mu_p \sigma_n(x), & l_p < x \le l_c \end{cases} \tag{5.74}$$

where μ_p is a constant friction coefficient.

From Eq. (5.74), it can be seen that there are two major unknown parameters: shear flow stress of the chip k_{chip} and the constant friction coefficient μ_p. According to

Oxley's theory (1989), the shear flow stress in the chip at the tool/chip interface can be calculated with Eq. (5.18). In the practical simulation, an initial value of μ_p is assumed to start the simulations. In this study, the initial value of μ_p is chosen at 1.5 according to the estimated value of the friction coefficient from experimental results.

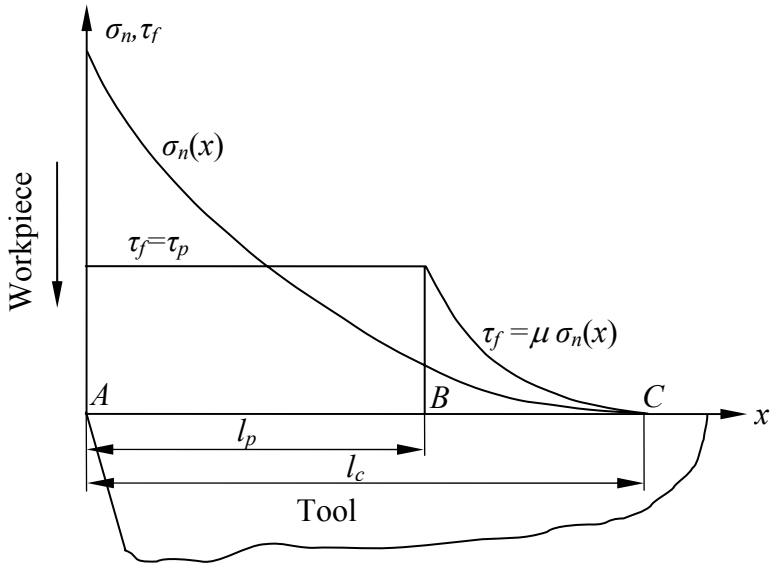


Figure 5.12 Curves representing normal stress (σ_n) and frictional stress (τ_f) distributions on the tool rake face proposed by Zorev (1963)

In FEM simulation at different instantaneous chip thickness, the values of shear angle ϕ are found by fitting the maximum strain-rate along the thin shear zone, and simultaneously the position of shear plane AB can be determined. Then, the shear stress k_{AB} and hydrostatic stresses at points A and B are found at different instantaneous chip thickness. Finally, the required value of C' is determined from Eq. (5.68), and the value of angle θ is obtained from Eq. (5.70). Finally, the cutting forces are found using Eq. (5.2). With this procedure, the cutting forces can be calculated accurately and realistically. So the assumption of uniform normal stress at the tool-chip interface used in Oxley's method is not needed any more.

5.5 Verification of the cutting force model

Using the approaches proposed in the above sections, undeformed equivalent chip geometries around the insert nose for three cutting conditions were calculated and process simulations were performed in order to predict process variables such as cutting forces, temperatures and tool stresses in 3-D flat end milling. The calculated chip geometry for the equivalent chip element representing the plane strain element was given in Table 5.2.

Before FEM simulation, many experiments have already been carried out. In order to avoid run-out effects, only one binderless CBN tool was used during slot milling. Fresh inserts were used in each experiment. The details of the tool material, cutting conditions and experimental setup are given in Chapter 3.

Table 5.2 Cutting parameters for the simulation of FEM

Simulation No.	Maximum predicted value of $h(\varphi)$ (mm)	Cutting speed (m/min)	Maximum predicted value of a_e (mm)	$\theta_0 + \frac{\theta_e}{2}$ (°)
1	0.075	350	0.075	78.391
2	0.100	350	0.100	76.702
3	0.100	400	0.100	76.702

The simulation of plastic deformation of chip formation when machining Ti-6Al-4V is shown in Fig. 5.13. These exists an obvious deformation zone between the boundary *CD* and *EF* is the chip formation zone.

The typical chip flow is given in Fig. 5.14, and at the same time the predicted cutting temperature at the tool is shown in Fig. 5.15. It can be seen from Fig. 5.15 that shear localization may exist, which would result in the generation segmented chips. Similar

experimental results have been achieved by other researchers (López de lacalle et al., 2000; Zoya and Krishnamurthy, 2000).

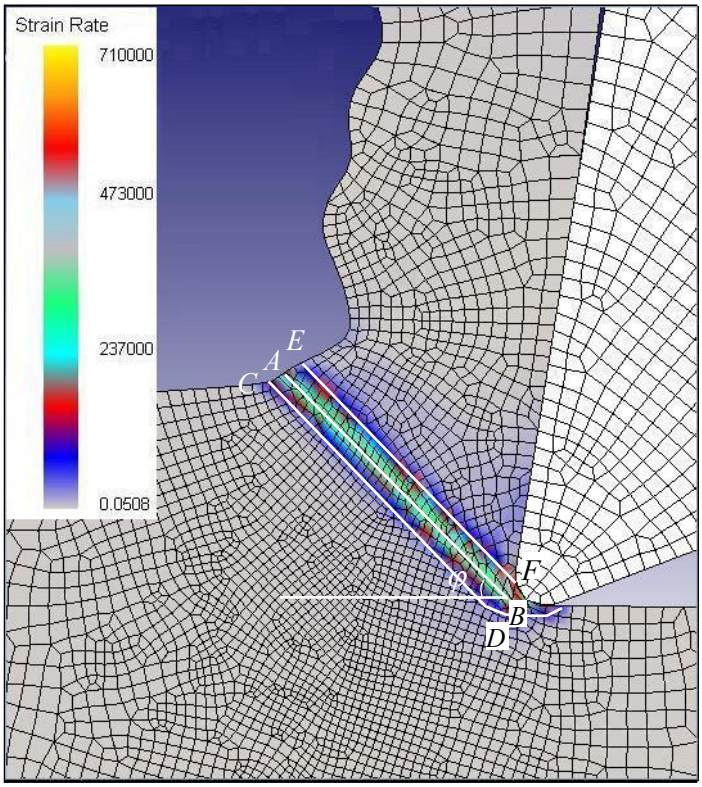


Figure 5.13 Deformation zones of FEM simulation in machining of Ti-6Al-4V

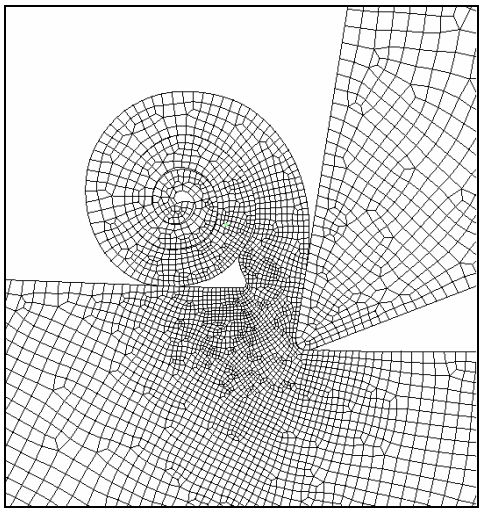


Figure 5.14 Process simulation with changing feed rate at rotation angle 30^{o} (a = 0.10mm, f = 0.10mm/r, v = 350m/min)

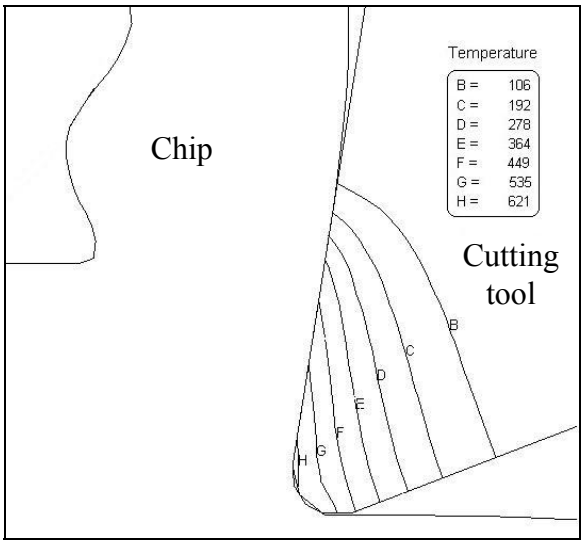


Figure 5.15 Temperature distribution at the position listed in Fig. 5.13

Figures 5.16-5.21 show the predicted cutting forces from FEM simulation and the corresponding experimental cutting forces. Unlike the measured cutting forces, the estimated cutting forces are fluctuating. The possible reason for this may be the frequent remeshing and the discretization of the cutting area. Similar conclusions have also been drawn by Ozel (1998). After fitting the trend of estimated fluctuating cutting forces, the mean values of estimated cutting forces are obtained (the solid line and dashed line in Figs. 5.16, 5.18 and 5.20). The estimated cutting force values of F_x and F_y are similar to those of the experimental results. This indicates that the cutting forces can be predicted with good accuracy for these two directions. But for all three cases, the estimated cutting forces of F_z are smaller than the experimental values. The difference may result from the plowing forces in milling, which is not taken into account in the cutting force model proposed in this study. However, the difference is not significant, so it means that the cutting forces can be predicted with reasonable accuracy for all three directions.

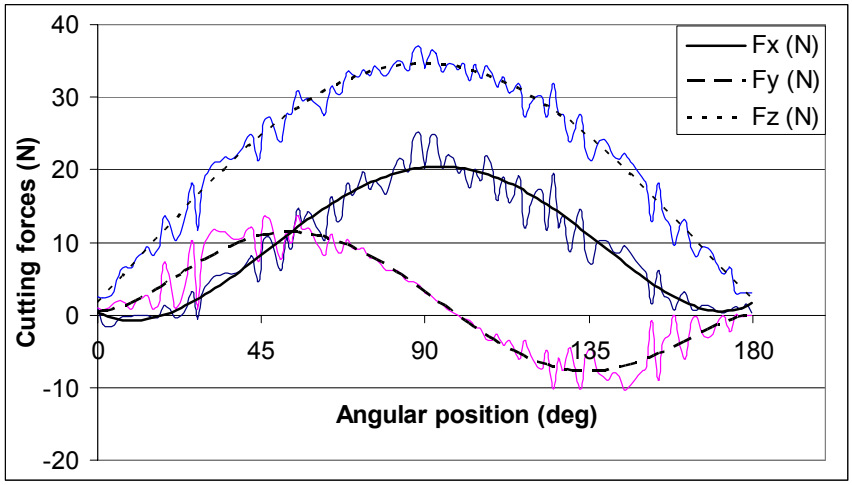


Figure 5.16 Estimated cutting forces at a = 0.075mm, f = 0.075mm/r and v = 350m/min

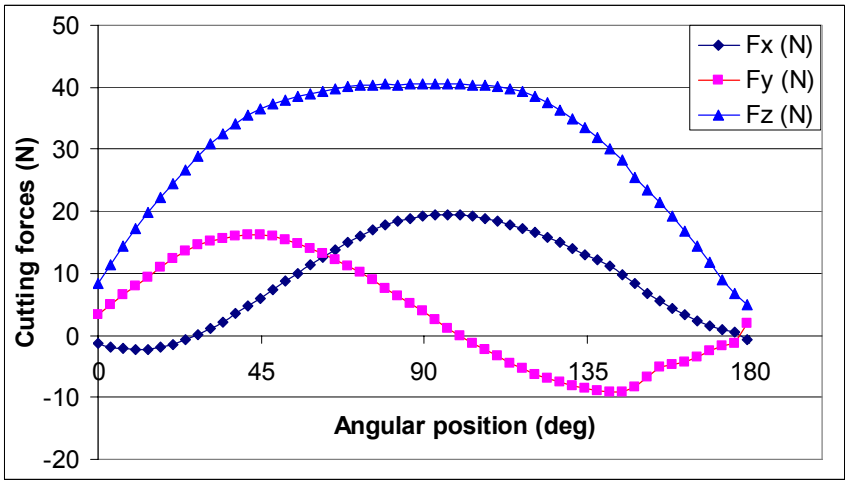


Figure 5.17 Experimental cutting forces at a = 0.075mm, f = 0.075mm/r and v = 350m/min

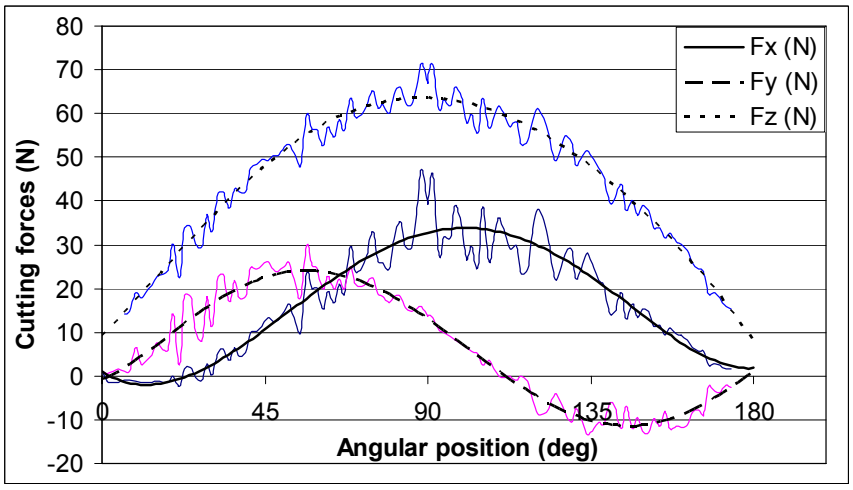


Figure 5.18 Estimated cutting forces at a = 0.10mm, f = 0.1mm/r and v = 350m/min

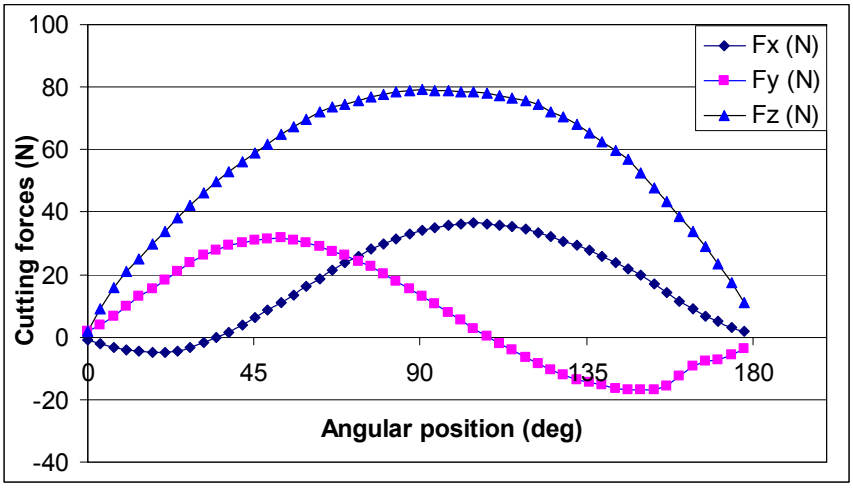


Figure 5.19 Experimental cutting forces at a = 0.10mm, f = 0.1mm/r and v = 350m/min

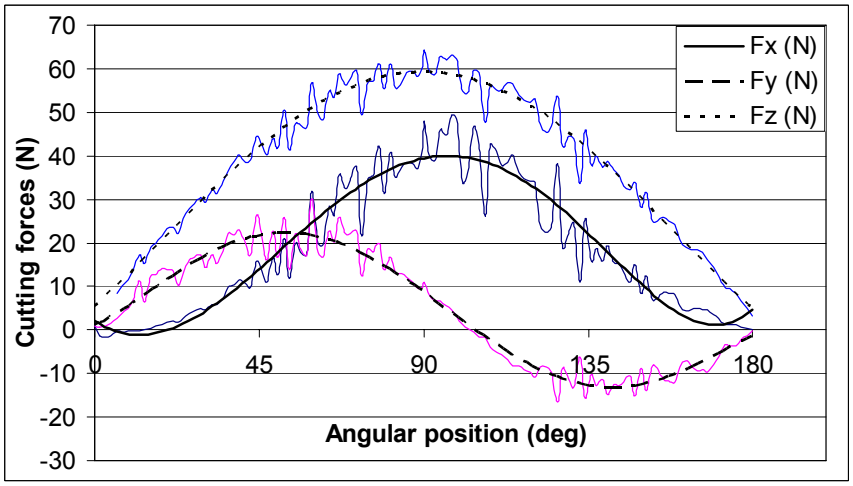


Figure 5.20 Estimated cutting forces at a = 0.10mm, f = 0.1mm/r and v = 400m/min

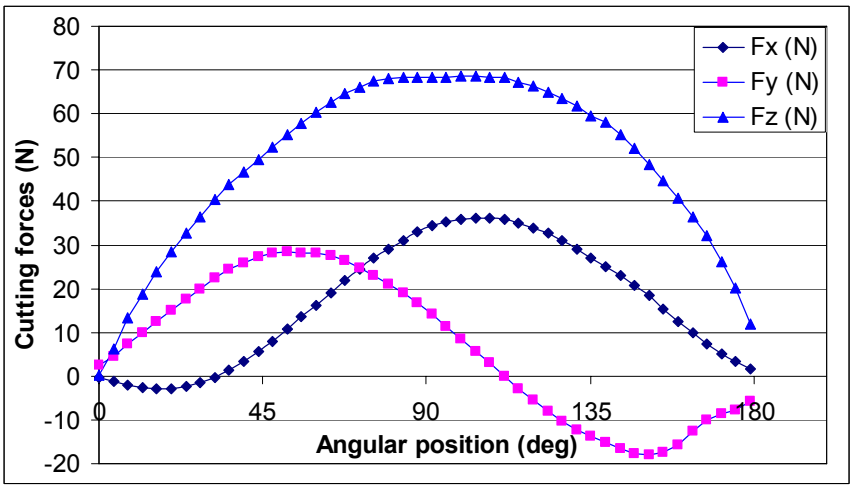


Figure 5.21 Experimental cutting forces at a = 0.10mm, f = 0.1mm/r and v = 400m/min

5.6 Concluding remarks

In this chapter, Oxley's predictive cutting force theory is overviewed briefly, which has been widely used to predict cutting forces, shear stress, etc. However, his model assumes that the tool is perfectly sharp and normal stress distributes uniformly at the tool-chip interface, which are not practical, particularly for HSM. Thus, a cutting force model for high-speed milling of Ti-6Al-4V is presented.

A hybrid cutting force model based on FEM simulation and Oxley's predictive machining theory is proposed to predict the cutting forces when high-speed milling Ti-6Al-4V. With this model, the impractical assumption of Oxley's theory is eliminated. It has been found that the force model can predict the cutting forces with reasonable accuracy.

Chapter 6

Development of a PGSA optimization algorithm

This chapter presents a parallel genetic simulated annealing (PGSA) algorithm that has been developed for optimization of high-speed milling. Firstly, the new hybrid method of GA and SA, referred to as GSA, is presented. Then, the model of hierarchical parallel GA is used to parallelize GSA in order to improve its efficiency. The implementation details of PGSA are discussed. Finally, the performance is evaluated against a standard set of test functions in comparison to other advanced search algorithm.

6.1 Introduction

Owing to its ability to achieve the global or near global optimum, GA has been applied to a large number of combinatorial optimization problems. The successful application of GA depends on the population size or the diversity of individual solutions in the search space. If GA cannot hold its diversity well before the global optimum is reached, it may prematurely converge to a local optimum. Although maintaining diversity is the predominant concern of GA, it also reduces the performance of GA. Thus, various techniques have been attempted to find a trade-off between the population diversity and the performance of GA (exploration and exploitation).

An alternative approach is to combine GA with other optimization techniques, such as simulated annealing (SA). This chapter presents a new hybrid of GA and SA, referred

to as genetic simulated annealing (GSA), which has been developed to avoid the premature convergence of GA by exploiting the local selection strategy of SA. GSA is also parallelized to improve its computation performance further. The chapter aims to demonstrate that parallel GSA is a powerful optimization strategy that overcomes the inherent weaknesses of conventional GA and PGA.

6.2 Genetic simulated annealing and its parallelization

6.2.1 Genetic simulated annealing

Recently, many researchers tried to combine GA and SA to provide a more powerful optimization method that has both good convergence control and efficient parallelization. In this section, a new tightly coupled hybrid of GA and SA is described as follows.

After crossover and mutation for a couple of individuals, there are four chromosomes: two parents and two offspring. In conventional GA, two parents are replaced by their offspring. But in PGSA, two chromosomes are chosen to form the next generation from these four individuals. The selection criterion is based on the fitness values of these four individuals. Individuals with higher fitness values have a greater probability of surviving into the next generation. Those with less fitness values are not necessarily discarded. Instead they are selected with a probability related to the current temperature (as in simulated annealing). In this selection process, a Markov chain is executed, which is composed of two offspring. Four parameters ($f_{best}, f_{worst}, T_t, f_i$) are involved to describe this selection process:

f_{best} — the best fitness value of two parents;

f_{worst} — the worst fitness value of two parents;

f_i — the fitness value of one offspring (i= 1,2);

T_t — controlling temperature;

During the course of the Markov chain at temperature T_t, the fitness value f_i (i =1, 2) of the trial chromosome is compared with f_{worst}. Chromosome i is accepted to replace the worst individual, if the following requirement is met:

$$\min\left\{1,\ e^{\frac{f_i - f_{worst}}{T_t}}\right\} \geq r$$

where r is a randomly generated number between 0 and 1. If chromosome i is accepted, the worst chromosome and the best one are updated then the course of the Markov continues again until it finishes. After the implementation of the Markov chain, the best and the worst individuals are survived into the next generation.

Initially, the mutation probability p_m of GSA is set to a higher value, and a simple annealing process is then used to adjust p_m. After every ten generations, the mutation probability p_m is updated with $p_m \times \alpha$ until it reaches to a certain value, where α is the cooling rate of SA. Thus, at the initial stage, when manipulating the cooling schedule of SA properly, the initial higher temperature can ensure that parents will be replaced by their offspring after crossover and mutation whether they are much fitter or not. More importantly, the initial higher mutation probability is capable of improving population diversity greatly, which can eliminate the premature convergence problem of conventional GA. On the contrary, at the later stage the mutation probability and the temperature become lower, and the chances for the fitter parents to be replaced decrease greatly. In this way, the current best individuals always remain in the next generation. Thus, the possibility of removing potentially useful individuals in the last

generation because of the mutation operation can be reduced. The pseudo-code of GSA is illustrated in Fig. 6.1, where $P(t)$ is the population of individuals at generation t.

```
begin
  t = 0
  initialize P(t) and temperature T_t
  evaluate P(t);
  while (not termination-condition) do
   begin
    t = t + 1
    select P(t) from P(t-1)
    select individuals for reproduction from P(t)
    do reproduction
        select two unused individuals P_1, P_2
        crossover & mutation; generate two children C_1,C_2
        evaluate C_1, C_2
        for i:= 1 to 2 do
            if min{1, exp((f_i-f_worst)/T_t)}> random[0,1)
            accept C_i & replace the corresponding parent
            update the new best and worst points
    until all selected parents finish reproduction
    T_t+1 = T_t × α; 0 <α < 1
    if( the modulus of t divided by 10 == 0 && p_m > 1/n)
       p_m = p_m×α
   end
end
```

Figure 6.1 Algorithm of genetic simulated annealing

In GSA, after conventional GA operator of crossover and mutation, local selection strategy of SA is applied to improve the convergence of GA. In addition, the higher initial mutation probability helps to maintain population diversity. Finally, good parallelizable property of GA is applied to parallelize GSA to improve its efficiency,

described in Section 6.2.2. Thus, GSA shows tighter coupling of these two paradigms as SA controls a number of distinct GAs running in parallel.

Based on above analysis, GSA can maintain good strengths of GA and SA. However, for problems with a large search space and costly function evaluation, a faster GSA is needed. Here, a parallel version of GSA (PGSA) is presented to improve its efficiency.

6.2.2 Parallel genetic simulated annealing

As described in Section 2.3.6 of Chapter 2, there are four categories about parallel GA. Among them, the hierarchical PGA is composed of coarse-grained PGA with master-slave or fine-grained PGA, so that it has the benefits of its components (Cantú-Paz, 2000). In this study, a master-slave/coarse-grained hierarchical PGA, which combines master-slave PGA and coarse-grained PGA together, has been borrowed to parallelize GSA.

In the master-slave/coarse-grained PGSA, the host program runs on the master processor, which decides on the global termination criterion. The whole population is equally divided into several subpopulations among the slave processors. Each slave processor runs a sequential GSA independently within its own subpopulation on one processor. The pseudo code of PGSA is shown in Fig. 6.2, where $P(t)$ is the population of individuals of generation t, *myrank* is the rank of the process, $slns_{migrate}$, $slns_{recv}$ and $slns_{delete}$ are the migrants, received individuals and the individuals to be replaced, respectively. If the migration conditions are satisfied, each processor, such as the source processor, implements the function *neighbor* to find the destination processors according to the migration topology. The migrant individuals $slns_{migrate}$ are selected

and sent to the destination processors. After the migrant individuals $slns_{migrate}$ are received on the destination processor as $slns_{recv}$, the individuals to be deleted $slns_{delete}$ are determined and replaced by $slns_{recv}$. The same program is executed on each processor, but on different data (their own population) until the global optimum is achieved.

```
begin
    t = 0
    initialize P(t)
    evaluate P(t)
    while (not termination-condition) do
        reproduction process of GSA
        if( migration-condition satisfies ) then
            dest = neighbor(myrank)
            slns_migrate = migrant_individual(P(t))
            send_string(dest, x_migrate)
            slns_recv = recv_string()
            slns_delete = delete_individual(P(t))
            replace_string (x_delete, x_recv, P(t))
        endif
    end
end
```

Figure 6.2 Pseudo code of parallel genetic simulated annealing

Although PGSA is related to the parallel hybrid method developed by Chen et al. (1998), there are some important differences between PGSA in this Chapter and Chen's PGSA (C-PGSA). In C-PGSA, each processing element (PE) maintains one solution, and each PE accepts a visiting solution from other PEs for crossover and mutation. In PGSA, each processor maintains its own subpopulation of solutions and different processors exchange their best solutions after a certain number of epochs.

Thus, the communication overhead between processors is much smaller. In C-PGSA, a normal SA-type probabilistic selection procedure is used to retain the proof of convergence of SA. In PGSA, a Markov chain is used to realize the local selection of SA, which can improve the selection performance of SA.

6.3 Full description of parallel genetic simulated annealing

6.3.1 Representation

PGSA uses a real-value coding scheme to represent the chromosome; and each chromosome vector is coded as a vector of real-value point numbers of the same length as the solution vector. Let $x = (x_1, x_2, \ldots, x_n)$ be the encoding of a solution, where x_i ($\in \boldsymbol{R}$) represents the value of the ith gene in chromosome x. Initially, x_i is selected within the desired domain, and reproduction operators of GA are carefully designed to preserve this constraint.

6.3.2 Selection

Based on their fitness function values, individuals are appropriately selected for recombination. The first step is to assign fitness values to all individuals according to their values of objective function(s). Sometimes the fitness value needs to be scaled for further use. Scaling is important to avoid early convergence caused by dominant effect of a few strong candidates in the beginning, and to differentiate relative fitness of candidates when they have very close fitness values near the end of run (Goldberg, 1989). In GA, there are mainly three selection procedures: proportional selection, tournament selection and rank-based selection (Blickle and Thiele, 1996). Proportional approach is usually called “roulette wheel” selection. Fitness values of individuals represent the width of slots of the wheel, and selection is based on the slot widths of

individuals (Pham and Karaboga, 2000). Individuals with larger slot widths will have a higher probability to be selected. In tournament approach, a sub-group is initially selected randomly from the population with or without replacement. Then, a 'tournament' competition takes place in this sub-group, and the winner is inserted into the next population (Blickle and Thiele, 1996). This process is repeated until the new population is formed. For rank-based approach, the individuals are sorted based on their fitness values. The rank N is assigned to the best individual, and the rank 1 is assigned to the worst one. Based on their ranks, the selection probability is then assigned to the individuals. Rank-based selection behaves in a more robust manner than the proportional one. Therefore in this study, the rank-based fitness selection approach is used.

6.3.3 Crossover and Mutation

The basic operators for producing new individuals in GA are crossover and mutation. *Crossover* may produce better individuals that have some genetic material of both parents. The conventional crossover operator combines substrings belonging to their parents. For real-value encoding, this type of crossover does not change the value of each variable; so it cannot perform the search with respect to each variable. Therefore, it is not suitable in this study and consequently, a modified crossover operator, called convex recombination (Dumitrescu et al., 2000), is used. It operates as follows.

Consider the crossover takes place at the position i, let $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ be the parents strings. Then, the offspring u and v are of the form:

$$u_i = a\, x_i + (1 - a)\, y_i,\ i = 1, \ldots, n;$$

$$v_i = a\, y_i + (1 - a)\, x_i,\ i = 1, \ldots, n.$$

where *a* is a random number in the interval [0, 1]. In this study, the two-point crossover is applied to each couple of individuals whose string length is greater than two. And for individuals whose string length is equal to two, one point crossover is used.

Mutation is the random alteration of gene values to each child after crossover. Usually crossover is regarded as the main operator of GA, and mutation is thought of as working in the background, which simply changes the value for a particular gene with a certain probability (mutation rate). Because mutation introduces variability or new individuals in the search space, it can avoid stagnation in the local optimal values. When used sparingly with reproduction and crossover, it is an insurance policy against premature loss of important notions (Goldberg, 1989).

The mutation rate is very important for the mutation operator to be effective. A very high mutation rate results in the mutation process to be a highly random, thus it helps to maintain the vast diversity of the population and also prevents the population from stagnating. However, the higher mutation rate also makes it difficult for GA to ensure its convergence and then makes GA ineffective. At later stages, it even increases the probability that good solutions will be destroyed. A very low mutation rate or no mutation at all may cause GA to converge a local optimum. The reason for this is that reproduction and crossover cannot introduce new genetic material into the search space for GA when individuals are similar to each other at later stages. So if the global optimum locates in the initialized search space, GA can find the optimal solution just with reproduction and crossover. But unfortunately, in real application there is no guarantee that the global optimum is in the initialized search space. In some cases, if

the global optimum is outside of the search space, it is impossible for GA to obtain the global optimal solution just with reproduction and crossover. Therefore, a trade-off should be made between the higher population diversity with a higher mutation rate and the risk of premature convergence posed by a lower one.

Some researchers have suggested that an optimal fixed mutation rate of $1/n$, where n is the length of the chromosome. This means that on average one bit per string has the chance to be chosen for mutation. Other researchers set the mutation rate to a low value (e.g., 0.01) so that accumulated good candidates will not be destroyed. In this study, the local selection of SA is applied after mutation, such that at the later stage, only better individuals were retained after mutation. The mutation rate can be larger than the recommended values by Hesser and Männer (1991) based on the dimension of the configuration space and population size.

For the real-value coding scheme, there are different mutation operators that can be used, such as uniform distributed mutation, Gaussian mutation, range mutation and non-uniform mutation. The first two mutation methods were used in this study. In the uniform distributed mutation, the mutation operator randomly chooses a number z in the interval defined by $[-A, A]$, where A is called the mutation range. The new point is given by: $x_m = x + z$ (Mühlenbein and Schlierkamp-Voosen, 1993). In Gaussian mutation, a random value z is chosen from a normal Gaussian distribution $N(0, \sigma)$, where σ is the standard deviation (Bäck, 1991).

Uniform distributed mutation is more commonly used for searches in a large region, and the Gaussian mutation performs better searches in a small local area (Dumitrescu

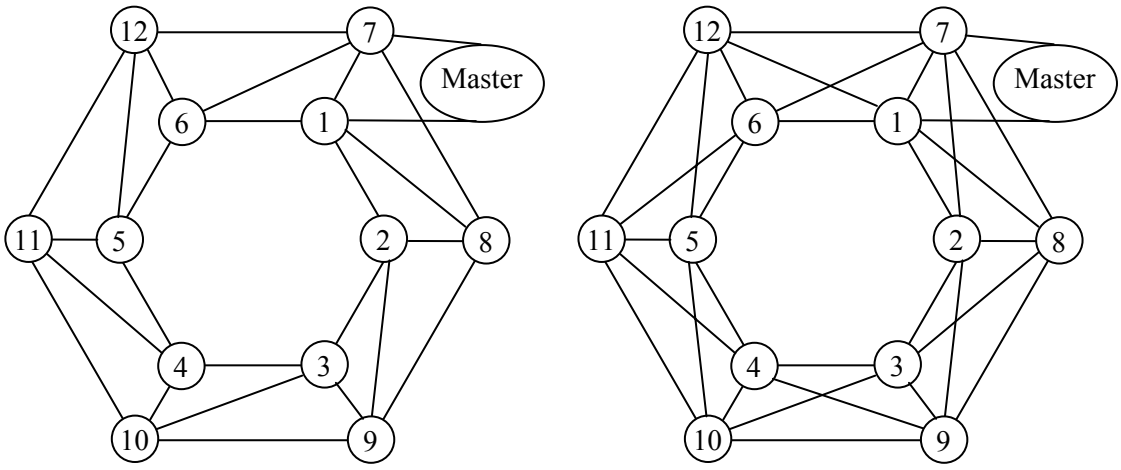
et al., 2000). In this study, at the initial stage, uniform distributed mutation is used. When the decreasing rate of the average fitness values is less than 0.01, Gaussian mutation is used.

6.3.4 Migration policy, rate, topology and frequency

In the implementation of the coarse-grained PGSA, some parameters of concern are: migration policy which determines how to select individuals to migrate, migration rate or the number of individuals to migrate, the frequency of migration, and the migration topology.

Two methods to choose the individual to migrate are often attempted: elitist strategy and probabilistic tournament selection. In the first method, the best individuals are sent to a neighboring subpopulation. In the second one, the fitter individual is selected via a probabilistic binary tournament with a certain probability. Gordon and Whitley (1993) compared these two methods and indicated that the elitist strategy performed better than tournament selection. Thus, the elitist strategy is used to choose the migration individuals. The top 1% of the best individuals are migrated to replace the worst individuals of other subpopulations. If 1% of population size is not an integer, it will be rounded off to the next integer that is greater than its fractional value. The topology is an important factor in the performance of PGA because it determines the speed with which good individuals disseminates to other sub-populations. The often used migration scheme is the ring topology, where individuals are exchanged between directionally adjacent sub-populations. The main advantage for the ring topology is that the communication time is smaller, but it cannot spread good solutions faster to all other sub-populations. Thus, the sub-populations are more isolated from each other

(Cantú-Paz, 1997). Cantú-Paz and Mejía-Olvera (1994) concluded that PGA with dense topologies finds the global solution using fewer function evaluations that PGA with sparsely connected ones. So more densely connected topology is needed to promote a better distribution of good individuals. Based on the ladder structure concept in (Mühlenbein et al., 1991), a similar ladder neighborhood relation is used to implement PGSA, and the 12-processor structure of this relation is shown in Fig. 6.3 (a), which shows that every slave processor is connected with other four slave processors. Furthermore, another topology is proposed as shown in Fig. 6.3 (b). Every processor except the master processor has five neighboring processors. The advantage of this topology is that it can spread the migrant individuals quickly among the slave processors. The frequency of migration is given in Section 6.4.1.



(a) 4-neighbor ladder structure (b) 5-neighbor ladder structure

Figure 6.3 Schematic diagram of the implementation of PGSA

6.3.5 Termination criterion

In PGSA, the host program, which runs on the master processor, decides on the global termination criterion. After certain iterations, each slave processor sends its top 1% of

the best individuals to the master processor. After receiving best individuals from all slave processors, the stop flag of PGSA is set on the master processor, if the following criterion is fulfilled:

$$\left| f_{best}^{k} - f_{best}^{k-\Delta k} \right| \le \varepsilon \left| f_{best}^{k} \right| \tag{6.1}$$

$$\text{or } \left| f_{best}^{k} - f_{best} \right| \le \varepsilon \tag{6.2}$$

where f_{best}^{k} is the best fitness value of an individual on the master processor at generation k, and the stop criterion is checked at every Δk generations; f_{best} is the global minimum and ε is a constant value 10^{-3}. For F1-F7 and F9, the Inequality (1) was used as the termination criterion. For Function F8, the Inequality (2) was chosen as the termination criterion, where $f_{best} = 0$. After setting the termination flag, the master processor sends it to every slave processor. According to the termination flag, each slave process decides whether GSA continues to run or not.

6.4 Numerical results and discussion

A set of test functions has been used to compare the performance of PGSA with other algorithms, which is taken from available literatures (Mühlenbein et al., 1991; De Jong, 1975; Griewank, 1981), as shown in Table 6.1. De Jong's suite of functions F1~F5 (De Jong, 1975) is extensively used in the GA community. Functions F6-F8 are also widely used to test the performance of global search algorithms. All tests were run on SUN Workstation network, which consists of 42 SUN Blade 2000 workstations with a fast Ethernet interconnect (100 M-Bytes/sec). The hardware and software details of Sun Blade 2000 are as follows: Ultra-SPARC III Cu 900 MHz processor, 2 G-Bytes memory, swap memory 4 G-Bytes, and Solaris 8 operating system.

Table 6.1 A set of standard test functions

Function	Function equation	Parameter intervals
F1	$f_1(x)=\sum_{k=1}^{3}x_k^2$	$-5.12\le x_i\le 5.12,\ \ i=1,2,3$
F2	$f_2(x)=100(x_2-x_1^2)^2+(1-x_1)^2$	$-2.048\le x_i\le 2.048,\ \ i=1,2$
F3	$f_3(x)=\sum_{i=1}^{5}\text{integer}(x_i)$	$-5.12\le x_i\le 5.12,\ \ 1\le i\le 5$
F4	$f_4(x)=\sum_{i=1}^{30}ix_i^4+Gauss(0,1)$	$-1.28\le x_i\le 1.28,\ \ 1\le i\le 30$
F5	$f_5(x)=0.002+\sum_{j=1}^{25}\frac{1}{j+\sum_{i=1}^{2}(x_i-a_{ij})^6}$	$-65.536\le x_i\le 65.536, i=1,2$
F6	$f_6(x)=nA+\sum_{i=1}^{20}[x_i^2-A\cos(2\pi x_i)]$	$-5.12\le x_i\le 5.12,\ \ 1\le i\le 20$
F7	$f_7(x)=\sum_{i=1}^{10}[-x_i\sin(\sqrt{\lvert x_i\rvert})]$	$-500\le x_i\le 500,\ \ 1\le i\le 10$
F8	$f_8(x)=\sum_{i=1}^{10}[-x_i^2/4000]-\prod_{i=1}^{10}\cos(x_i/\sqrt{i})+1$	$-600\le x_i\le 600,\ \ 1\le i\le 10$
F9	$f_9(x)=\sum_{i=1}^{n}[100(x_{i+1}-x_i^2)^2+(1-x_i)^2]$	$-5.12\le x_i\le 5.12,\ \ i=1,2,...,n$

An extensive performance evaluation for functions F1-F8 has been done for the PGA by Mühlenbein et al. (1991). In their investigation, the efficiency of the search method was demonstrated by varying the problem size *n*. Mühlenbein and Schlierkamp-Voosen (1993) increased the problem size *n* further to investigate the advantage of the scalability of their search method. The number of function evaluations was defined as efficiency criteria in (Mühlenbein and Schlierkamp-Voosen, 1993) and (Mühlenbein et al., 1991). For comparison, in this section, the efficiency of search methods is also defined by the number of function evaluations needed to obtain the optimal solutions. When the optimum of each function on the master processor is reached, it will send the termination signal to all slave processors. After these processors have received this termination signal, they stop running and send the number of function evaluations done so far to the master processor to sum them up.

6.4.1 Parameters selection for PGSA

Both GA and SA have many internal control parameters, and their performance is very sensitive to such parameters (Chen et al., 1998). Different parameters of GA or SA are used for different problems. PGSA, which is parallel version of the hybrid of GA and SA, also has many control parameters, and can be described as follows:

$$PGSA = (P_0, \lambda, \mu, \sigma, \delta, GSA, T_0, \alpha, t) \quad (3)$$

where P_0 is the initial population, λ is the number of subpopulations, μ is the population size of each subpopulation, σ is the migration interval in number of generations, δ is the number of neighbors, T_0 is the initial temperature, α is the cooling rate and t is the termination criterion.

Thus, in order to show the robustness of PGSA, there is no need to tune all these parameters to a specific function. Based on the recommended values from previous work (Cantú-Paz, 2000; Mühlenbein et al., 1991), the parameters for F1-F9 are shown in Table 6.2, where p_m and p_c are the respective mutation and crossover probabilities. Function F8 is much more difficult to get the global minimal value, so the larger population size and mutation rate were used. Furthermore in order to get more representative results, the average values for PGSA listed in the following tables were based on 50 runs across the test suite of functions.

6.4.2 Results and discussion for lower dimension problems

The computation results for F1-F9 using PGSA are shown in Table 6.3, where n is the problem size and FE indicates the number of function evaluations. For comparison, the average number of function evaluations with optimized PGA by Mühlenbein et al. (1991) is also included in Table 6.3. For Functions F1-F5, difference in performance

between PGA and PGSA is very small. These problems could be too small for PGSA or PGA to work ideally. The comparison of more complex functions F6-F9 shows more significant difference.

Table 6.2 Control parameters setting for Function F1-F9

Parameters	F1~F5	F6	F7	F8	F9
λ	8	8	8	16	20
μ	20	20	20	50	100
σ	10	10	10	20	20
δ	4/5	4/5	4/5	4/5	4/5
p_m	0.5	0.1	0.1	0.3	0.05
p_c	0.65	0.65	0.65	0.65	0.85
T_0	200	200	200	200	800
α	0.85	0.85	0.85	0.85	0.85

Table 6.3 Performance comparison between PGA (Mühlenbein et al., 1991) and PGSA

Function	n	PGA (Mühlenbein et al.)		PGSA	
		λ	*FE*	λ	*FE*
F1	3	8	1526	8	1287
F2	2	8	1671	8	1473
F3	5	8	3634	8	1769
F4	30	8	5243	8	4025
F5	2	8	2076	8	1476
F6	20	8	9900	8	6705
F7	10	8	8699	8	6006
F8	10	16	59520	16	25690
F9	50	—	—	20	537235

Functions F6-F7 with lower dimension are also easily solved by PGSA and PGA. But in (Mühlenbein et al., 1991), the global optimum of F7 was not found in 4 of the 50

runs. In this study, PGSA found the optimum of F7 in all 50 runs. Because PGSA can maintain a good diversity with a higher mutation probability at the initial stage, it can eliminate premature convergence to suboptimal minima. At the later stage, the local selection strategy of SA can ensure that best individuals are not discarded after crossover and mutation operator. Therefore, PGSA can approach or converge on the global minimum.

Griewank's function F8 is regarded as one of the most difficult test functions. Function F8 has its global minimum f_{best} = 0 at x_k = 0, and the local minima are located approximately at $x_k = m\pi\sqrt{k}$, where k = 1, …, n, and m is any integer value. Four suboptimal minima (≈0.0074) exist at $\vec{x} = (\pm\pi, \pm\pi\sqrt{2}, 0, ...0)$ in ten dimensions. The average number of function evaluations is 6600 by Griewank (1981), but only one of the four sub-optima was found. An average of 59520 evaluations is needed to solve this problem by Mühlenbein et al. (1991), but they did not comment on their results. PGSA found the minimum values (< 0.001) of F8 with only about half of the number of function evaluations using PGA. More importantly, in all 50 runs, these minimum values were found. Therefore, PGSA is able to obtain much better solutions with a higher convergence speed than PGA.

Function F9 with a dimension of 50 was extended from that of Rosenbrock. For this kind of functions, the mutation scheme is not efficient, and line recombination is able to leave the saddle point after a steep valley (Schlierkamp-Voosen and Mühlenbein, 1994). In this study, the line recombination was also tried for F9 instead of the crossover mentioned in section 6.3.3. Even with the use of line recombination, PGSA still converges to suboptimal points in some cases. A more powerful mutation operator

is needed to help individuals leave the suboptimal points. In this chapter, by considering the relationship between the objective-function values and the difference of two consecutive genes, a revised Gaussian mutation is proposed, which has non-zero expectation. Let $x = (x_1, \ldots, x_n)$ be a parent string, the new offspring u is of the form:

$$u_i = \begin{cases} x_i + N(0, \sigma), & i = 0 \\ x_i + N(\frac{x_{i-1} - x_i}{2}, \sigma), & 0 < i < n \end{cases} \quad (4)$$

where σ is the standard deviation. In our study, σ =0.005. We choose $(x_{i-1} - x_i)/2$ as the expectation because objective function values depend on the difference of values of two consecutive genes. A typical run using PGSA and PGA with the revised Gaussian mutation and PGSA with normal Gaussian mutation is shown in Fig. 6.4.

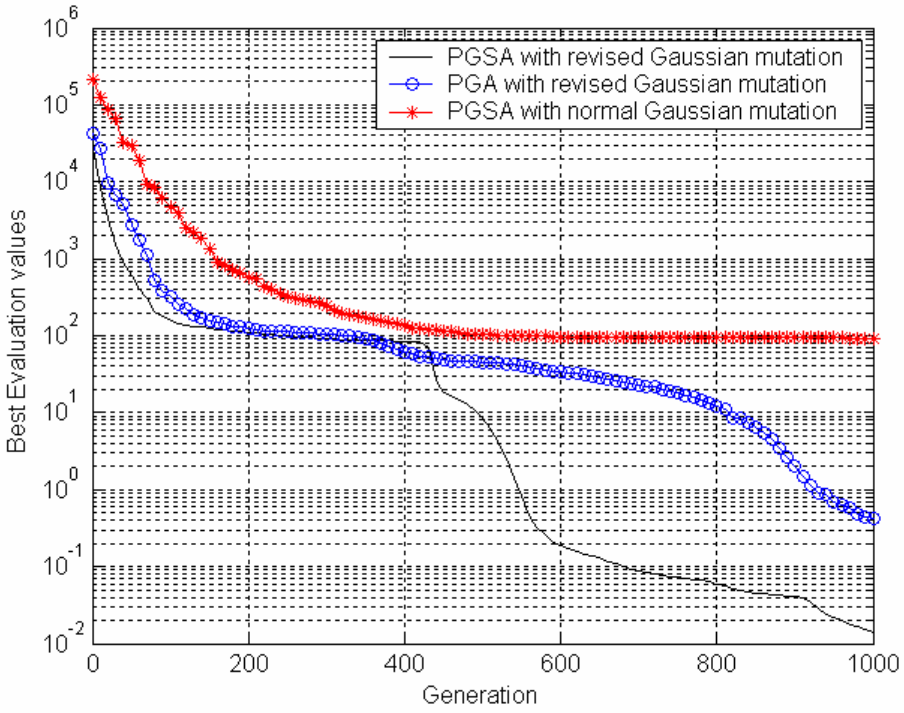


Figure 6.4 Best function evaluation values for Rosenbrock function (n = 50)

From Fig. 6.4, it can be seen that the best evaluation values decrease slightly following a steep decrease initially for these three cases. PGSA and PGA with the revised

Gaussian mutation can move out from such saddle points. However, it is very difficult for PGSA with normal Gaussian mutation to leave the saddle points. Even after 1000 generations, it still stagnates at the suboptimal points, while PGSA and PGA with the revised Gaussian mutation can exploit the more optimal points gradually. Hence, the revised Gaussian mutation is more efficient than the normal one, because it used information from its former consecutive genes to move the individuals towards more optimal points. On the contrary, normal Gaussian mutation just changes the value of a particular gene randomly, while the objective function value depends not only on this particular gene, but also on the difference between its value and its former consecutive one. Thus, it cannot guarantee that it will move out of the suboptimal points. However, when choosing $(x_{i-1} - x_i)/2$ as the expectation, the revised Gaussian mutation can solve this problem. In addition, with the help of the revised Gaussian mutation, the performance of PGSA outweighs that of PGA. The reason for this is discussed in following section.

6.4.3 Discussion of speed-up of PGSA

For Function F7, the sub-optima are far away from the optimization so that a parallel search with sub-populations seems to be promising. Function F7 with the dimension of 50 and 100 has been tried, and the testing parameters settings are shown in Table 6.4. The measured computation time and communication time are shown in Figs. 6.5 & 6.6, which indicates a super-linear speed-up when the number of processors increases from 6 to 10. However, when more processors are involved in computation, the decrease rate of computation time becomes smaller, and at the same time the communication overhead has increased. For F7 (n = 50), when 24 processors are used, the savings on

computation costs cannot compensate for the rapidly increasing costs of communication.

Table 6.4 PGSA's parameters setting for Function F7 (n = 50 and 100)

F7	σ	p_m	T_0	λ-μ								
n =50	30	0.05	400	6-134	8-100	10-80	12-68	14-58	16-50	18-44	20-40	22-36
n =100	50	0.05	600	6-334	10-200	14-142	16-126	18-112	20-100	22-92	24-84	26-78

When lesser processors are used for F7, there is only slight difference in computation time between 5-processor neighborhood ladder structure and 4-processor one. However, when more processors are involved in the computation, the difference in computation time between these two types of ladder structure becomes obvious, as shown in Figs. 6.5 & 6.6. Although, with 5-processor neighborhood ladder structure, the communication costs between processors are slightly higher than those with 4-processor one, the savings of computation time can compensate for the increased communication overhead. Therefore, 5-processor neighborhood ladder structure has been chosen as the migration topology.

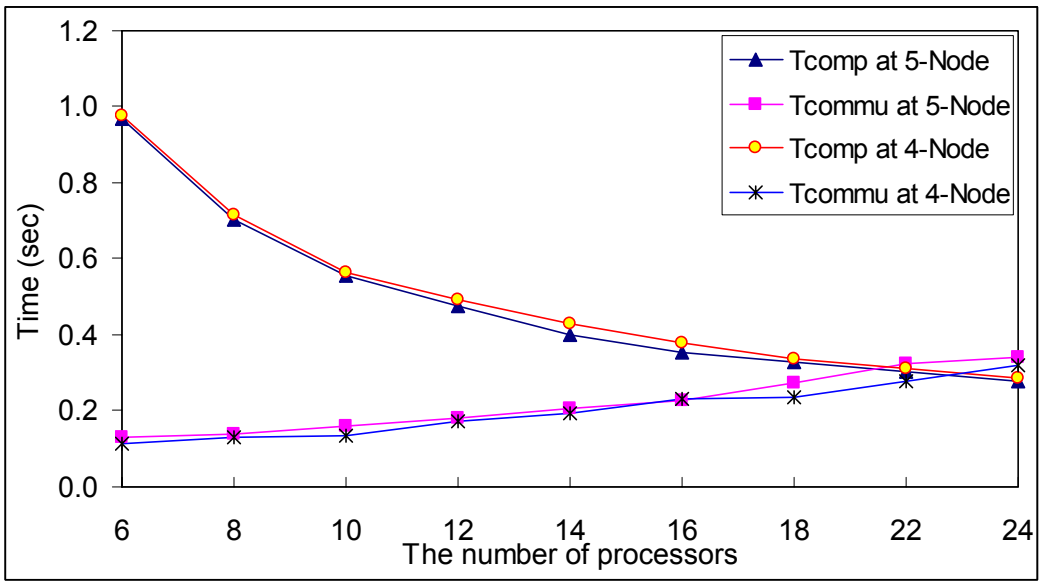


Figure 6.5 Computation time and communication overhead for F7 (n = 50)

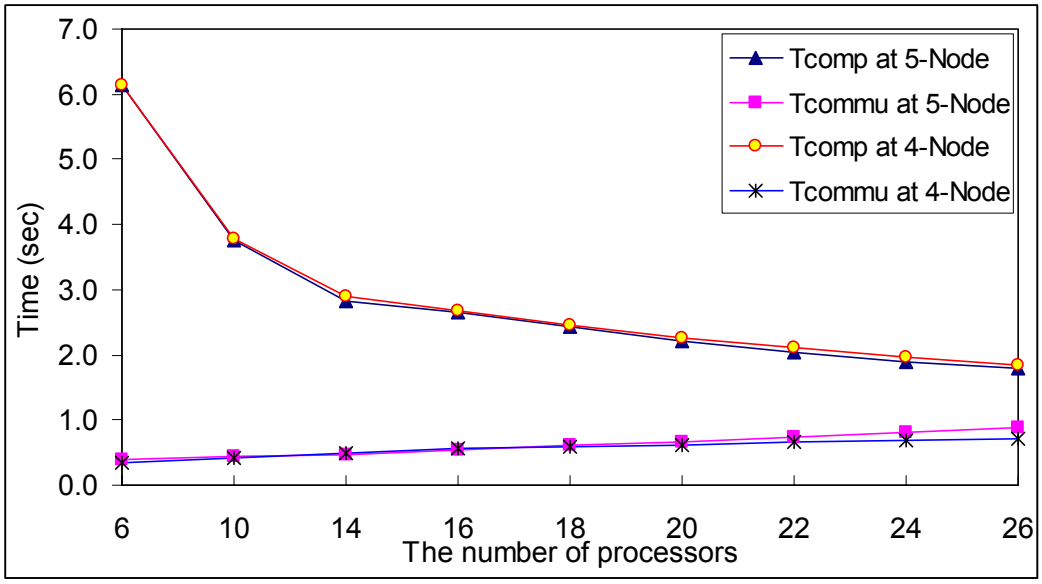


Figure 6.6 Computation time and communication overhead for F7 (n = 100)

Similarly, based on the trade-off between communication cost and computation time, the optimal number of processors used in this study has been determined.

6.4.4 Computation results for F6 and F7 with higher dimension

Mühlenbein et al. (1991) found the global minimum of F6 of dimension 400 and F7 of dimension 150 on a 64-processor computer using PGA. For comparison in this Chapter, F6 and F7 of much higher dimension have also been attempted using PGSA. The test parameters are listed in Table 6.5. The same termination criterion (as given in Section 6.3.5) was used in order to directly compare the efficiency of PGSA and PGA.

It is instructive to compare our results with the results from (Mühlenbein et al., 1991), as shown in Table 6.6. Function F6 with dimension 500 and 1000 and Function F7 with dimension 200 and 400 have also been investigated, which were not solved using PGA in (Mühlenbein et al., 1991). It can be seen that the number of function

evaluations using PGSA is much smaller than that using PGA. The performance of PGSA gets better with the higher problem size n.

Table 6.5 Control parameters setting for Function F6 & F7 with higher dimension

Dimension / Parameters	F6						F7				
	50	100	200	400	500	1000	50	100	150	200	400
λ	8	16	16	16	16	32	16	16	16	16	32
μ	20	20	40	80	100	100	50	100	150	200	200
σ	20	20	40	40	50	50	30	50	50	50	50
p_m	0.05	0.05	0.05	0.65	0.80	0.80	0.05	0.05	0.05	0.05	0.05
p_c	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65
T_0	400	500	1000	2000	2000	2000	400	600	800	1000	1200
α	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85

Table 6.6 Performance comparison between PGA (Mühlenbein et al.,1991) and PGSA

Function	n	PGA from Mühlenbein et al.			**PGSA**		
		λ	μ	*FE*	λ	μ	*FE*
F6	n=50	8	20	42753	8	20	*18378*
	n=100	16	20	109072	16	20	*77489*
	n=200	32	40	390768	16	40	*257894*
	n=400	64	40	7964400	16	80	*356333*
	n=500	—	—	—	16	100	*565348*
	n=1000	—	—	—	32	100	*1168424*
F7	n=50	32	20	119316	16	50	*70853*
	n=100	64	20	1262228	16	100	*273222*
	n=150	64	40	7041440	16	150	*689888*
	n=200	—	—	—	16	200	*1399284*
	n=400	—	—	—	32	200	*2846835*

A typical run is shown in Fig. 6.7. Initially, the difference in performance between PGSA and PGA is very small. After 100 generations, it becomes larger, and for PGA

the individuals of the sub-population become more similar than PGSA. In some cases, PGA even converges around suboptimal minima. Thus, the average evaluation values are smaller than those of PGSA. However, for PGSA, at this stage, the smaller suboptimal minima have been found. The difference between the average and the best evaluation values is larger than that of PGA, which means that more diverse individuals exist in sub-populations. As the optimization process proceeds for PGSA, the probability that the best individuals found are kept in the next generation becomes larger, when the controlling temperature gets lower. If the temperature is very small, the best individuals are always retained in the next generation. Normally, the mutation probability is set to a low value (e.g., 0.01) for GA or PGA, otherwise the accumulated good candidates are destroyed at the later stage. However, this negative effect of mutation has been eliminated for PGSA, because the local selection of SA takes place during this later stage and only better individuals are retained after mutation. Finally, because of the maintained diversity and the local selection of SA, PGSA can converge to the optimal solution within much smaller number of generations. It can be seen that PGSA found the global optimal solution after 700 generations, whilst PGA still cannot find its optimum even after 800 generations.

6.4.5 Computation results for F8 with higher dimension

Mühlenbein and Schlierkamp-Voosen (1993) used breeder genetic algorithm (BGA) to optimize Functions F6 and F7 with higher dimensions too. However, the number of function evaluations cannot be directly compared because a different termination criterion was used for BGA. In order to compare the efficiency of PGSA with that of BGA, the same termination criterion as that for BGA was used for Function F8. The computation results using BGA (Mühlenbein and Schlierkamp-Voosen, 1993) and

PGSA are listed in Table 6.7. The number of function evaluations scales almost exactly with $n \cdot ln(n)$ for Function F8 (Mühlenbein and Schlierkamp-Voosen, 1993). When the problem dimension size n is in the range of 100 to 400, the number of function evaluations scales exactly with $\sqrt{n}/\ln(n)$ in this study. With such scaling ability, the advantage of investigating the scaling of PGSA has been demonstrated. PGSA achieves much better performance with the higher problem size n. Thus the performance of PGSA is better in comparison to that of BGA.

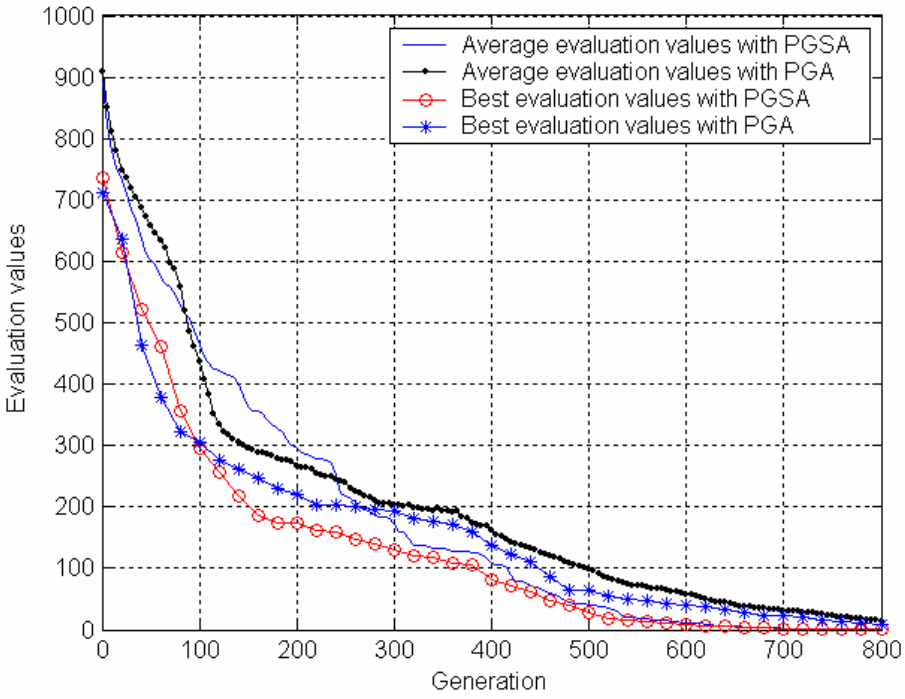


Figure 6.7 Average and best function evaluation values for Rastrigin function

Table 6.7 Performance comparison between BGA (Mühlenbein et al, 1993) and PGSA

Function	n	BGA (Mühlenbein et al., 1993)				**PGSA**			
		λ	μ	*FE*	$668n \cdot ln(n)$	λ	μ	*FE*	$58690\sqrt{n}/\ln(n)$
F8	n=20	1	500	66000	40023	14	50	*39707*	87614
	n=100	1	500	361722	307625	16	100	*123610*	127444
	n=200	1	500	748300	707855	16	100	*157568*	156654
	n=400	1	500	1630000	1600920	16	100	*196054*	195912

6.5 Concluding remarks

In this Chapter, a new GA and SA hybrid (GSA) is firstly presented, which inherits the strengths of GA and SA and overcomes their weaknesses. The extended ideas of simulated annealing (SA) were used in: (1) adjustment of the mutation rate; (2) the local selection of individuals retained in the next generation after crossover and mutation. Nine different functions have been tried with the proposed algorithm. The only difference between them lies in the objective function. If in future, any user wants to use our program for new problems, they only need to incorporate their objective function into the program and change some parameters if necessary. So the programmability of the algorithm is good.

The numerical results show that PGSA has faster convergence to global optimal solution than PGA, and thus, is superior to the conventional PGA. In comparison to the other advanced search method such as breeder genetic algorithm (BGA) (Mühlenbein and Schlierkamp-Voosen, 1993) using the same termination criterion, the performance of PGSA is better. More importantly, PGSA performs better with the larger problem size.

Chapter 7

Optimization of high-speed milling

In this chapter, the optimization of multi-pass milling has been investigated in terms of two objectives: production time and production cost. The optimization functions and the constraints are firstly given. Then, an advanced search algorithm parallel genetic simulated annealing (PGSA) was used to obtain the optimal cutting parameters under different axial depth of cut. In the implementation of PGSA, the fitness assignment is employed based on the concept of non-dominated sorting genetic algorithm (NSGA). Two application examples are given using PGSA. For the given examples, the algorithm PGSA can find a Pareto-optimal front which is composed of many Pareto-optimal solutions in a single simulation. The weighted average strategy is used to find the optimal cutting parameters along the Pareto-optimal front. Finally, based on the concept of dynamic programming, the optimal cutting strategy can be obtained.

7.1 Introduction

As described in Section 2.4 of Chapter 2, efforts are mainly focused on the single-objective optimization (production time or production cost). Little attention has been given to the multi-objective optimization of machining process although the production cost and production time are two main concerns for machining processes. In this chapter, two objectives of production time and production cost are considered in

the optimization of high-speed milling of Ti-6Al-4V, thus the optimization is posed as multi-objective optimization problem. It is clear that a minimum cost solution is usually not that one with minimum production time, and vice versa. Therefore, in such a scenario, no single solution is superior to the other one for these two objectives of machining, and such type of solution was called a non-dominated solution. The set of all non-dominated solutions is called the Pareto set or Pareto front (Steuer, 1986). Because there are no efficient multi-objective optimization algorithms available, the multi-objective optimization problems have traditionally been solved by converting them into single-objective optimization problems. The often used conversion methods include weighted sum approach, ε-perturbation method, Tchybeshev method, min-max method, goal programming method and others (Deb, 1999). When this type of classical methods is used, only one solution has been obtained, and it cannot guarantee that this is a Pareto-optimal solution. In addition, in order to find many Pareto-optimal solutions, the algorithm needs to be run for a number of times. Furthermore, the classical methods have been found to be sensitive to the convexity and continuities of the Pareto-optimal region (Deb, 1999).

Since GA can maintain a population of solutions, it is logical to use GA to find multiple Pareto-optimal solutions in a single simulation. However, for complicated engineering optimization problems, GA cannot guarantee to obtain the global optimal solution(s) for each run. In some cases, if GA cannot hold its diversity well before reaching the global optima, it may not find the global optimum, or even result in the premature convergence to local optimum. PGSA developed in Chapter 6 has avoided the premature convergence of GA by exploiting the local selection strategy of SA. In this chapter, this algorithm PGSA has been used to optimize high-speed milling of Ti-

6Al-4V. For comparison, the conventional PGA was also tried in this study with the same termination criterion as that used for PGSA.

For the optimization problems of multi-pass machining, two stages are needed for sequential method or algorithm. Firstly, the number of passes or depth of cut are given; then GA was used to find the optimal cutting parameters for a given depth of cut in the second stage. This methodology has been used by Dereli et al. (2001) and Shunmugam et al. (2000). On one hand, this method takes much more time to implement two stages work sequentially. On the other hand, it cannot guarantee the global optimal solution to be obtained with the conventional GA. PGSA not only improves the efficiency of GSA which has been verified in Chapter 6, but also ensures that global optimal solution(s) can be achieved for the optimization of multi-pass milling.

7.2 Objective function

The minimum production time and the minimum production cost have been chosen as the objective function. The production time for a component is the total time required to produce a component and is composed of the following items: (i) Set up time T_s which is the time taken to load and unload each component and to return the tool to its starting point, (ii) Process adjustment and quick return time T_a, (iii) Machining time T_m, and (iv) Tool change time per component T_{tc}. Therefore, the total production time per component is: $T_{pr} = T_s + T_a + T_m + T_{tc}$.

For a multi-pass operation in which N_p passes are required to remove the total depth of cut, the production time is given as:

$$T_{pr} = T_s + \sum_{i=1}^{N_p} (T_{a_i} + T_{m_i} + T_{tc} \frac{T_{m_i}}{T}) \qquad (7.1)$$

Recently with the new development of cutting tools, their lives are long enough to machine the components without significant wearing out the tools during the process of machining. Thus, in this study, the replacement of the worn tool can be neglected. The tool-changing time is determined by the indexing time required by the automatic tool changer (ATC) of the machine tool to swap a used cutter from the spindle with a tool in the tool magazine. This time is independent of the cutting parameters. In addition, the set up time, the loading/ unloading time and the process adjustment time are not influenced by the cutting parameters. Thus, the final optimization model can be simplified as:

$$T_{pr}' = \sum_{i=1}^{N_p} T_{m_i} \tag{7.2}$$

The machining time (T_m) for the slot milling operations can be calculated in the following form:

$$T_{m_i} = \frac{L}{f_{Z_i} z N_i} = \frac{L}{f_{Z_i} z (1000 v_i)/(\pi D)} = \frac{\pi D L}{1000 f_{Z_i} z v_i} \tag{7.3}$$

where z is the number of teeth on the cutter.

By substituting Eq. (7.3) into Eq. (7.2), the final simplified model for minimum production time per component can be expressed as:

$$T_{pr}' = \sum_{i=1}^{N_p} \left(\frac{\pi D L}{f_{Z_i} z 1000 v_i}\right) \tag{7.4}$$

The optimization model for minimum production cost can similarly be derived. The total production cost per component comprises the fixed cost and variable cost. The fixed cost includes cost of raw material, set-up cost and adjustment cost which are all independent of machining time. The variable cost is a function of the machining time, which includes machining cost and cost of changing the cutting tool. The set up cost, adjustment cost and machining cost include the operator's salary and the

corresponding overhead associated with his employment and the depreciation of the machining tool and the overhead associated with its use. For the tool-changing cost, the cost of the cutter is added besides the cost of the operator and the machine tool depreciation. The tool-changing cost depends on the frequency of tool replacement. Therefore, the average production cost for each component can be written as:

$$C_{pr} = C_{mat} + (C_l + C_o)T_s + (C_l + C_o)T_a + (C_l + C_o)T_m + [(C_l + C_o)T_{tc} + C_t]\frac{T_m}{T} \quad (7.5)$$

The first three items are non-productive cost, which is constant for a particular operation. Since they are not influenced by the cutting parameters, the final optimization model for minimum product cost can be described as:

$$C_{pr}' = (C_l + C_o)T_m + (C_l + C_o)T_{tc}\frac{T_m}{T} + C_t\frac{T_m}{T} \quad (7.6)$$

Based on the above analysis, tool-changing time is independent of the ratio of the machining time to the tool life for a multi-tool optimization approach. Thus the model can be further simplified as:

$$C_{pr}' = (C_l + C_o)T_m + C_t\frac{T_m}{T} \quad (7.7)$$

The empirical tool-life equation $vT^{\alpha}f^{\beta}a^{\gamma} = C$ is used to estimate the practical tool life of the cutting tool. A multi-element regression method based on the experimental results is used to determine the values of constant C and exponents α, β, γ.

In the optimization model, three parameters are considered: depth of cut, cutting speed and feed rate. Among them, depth of cut is the dominant parameter. Although removing the total depth of cut is preferable with only a single pass, this situation is

impossible or not applicable for most machining cases. Because with larger depth of cut, the cutting tool cannot stand larger cutting forces and its tool life is very short. In addition, this type of operation is also constrained by the machine tool, which cannot supply large enough power to remove larger volume of workpiece materials in a single pass. In this chapter, the optimization of multi-pass milling has been investigated with regard to optimal cutting passes, and three corresponding cutting parameters: depth of cut, cutting speed and feed rate.

In all, four parameters: the number of passes N_p, corresponding depth of cut, cutting speed and feed rate, are involved in the optimization of milling process. However, the number of passes depends on the depth of cut for every pass and total depth of cut. For the given total depth of cut and the number of passes, the depth of cut per pass to be removed per pass is determined. Then, the cutting strategy can be determined by the number of cutting passes and the corresponding depth of cut, cutting speed and feed rate for every pass. Finally, the optimization problem is to find the optimal cutting strategy so as to achieve the selected objective function.

The determination of the number of cutting passes can be considered as a one dimensional walk problem, in which the whole distance (total depth of cut) can be divided into possible sections (depth of cuts). The length of each section should always be smaller than the maximum depth of cut and greater than the minimum depth of cut allowed for the combination of cutting tools and the workpiece.

Based on the experimental results and the general recommended optimal working ranges, the value of depth of cut changes from 0.075mm to 0.150mm in this study with

the minimum increment of depth of cut is 0.025mm. So the possible values of depth of cut (length of possible sections) are 0.075mm, 0.100mm, 0.125mm and 0.150mm. Assuming that the total depth of cut is 0.30mm, the determination of the number of cutting passes can be illustrated as the one dimensional walk procedure as shown in Fig. 7.1. The total depth of cut means the walking distance from point *N*(0, 0) to any point at the bottom line *AB* along the possible ways, which are represented as the solid line. The problem is to find all alternative possible passes, that are composed of the certain number of possible sections, to any point of *N*(1, 0), *N*(2, 0), *N*(3, 0) and *N*(4, 0). The procedure starts from the point *N*(0, 0) (free surface of the workpiece). The stock can be machined from the outer end of the workpiece with four possible depth of cut: 0.075mm, 0.100mm, 0.125mm and 0.150mm. Then, this procedure is continued until any one of destination points is reached. At the same time, all possible passes are stored for future use. In all, to remove the total depth of cut, five alternative cutting strategies are available, as shown in Table 7.1.

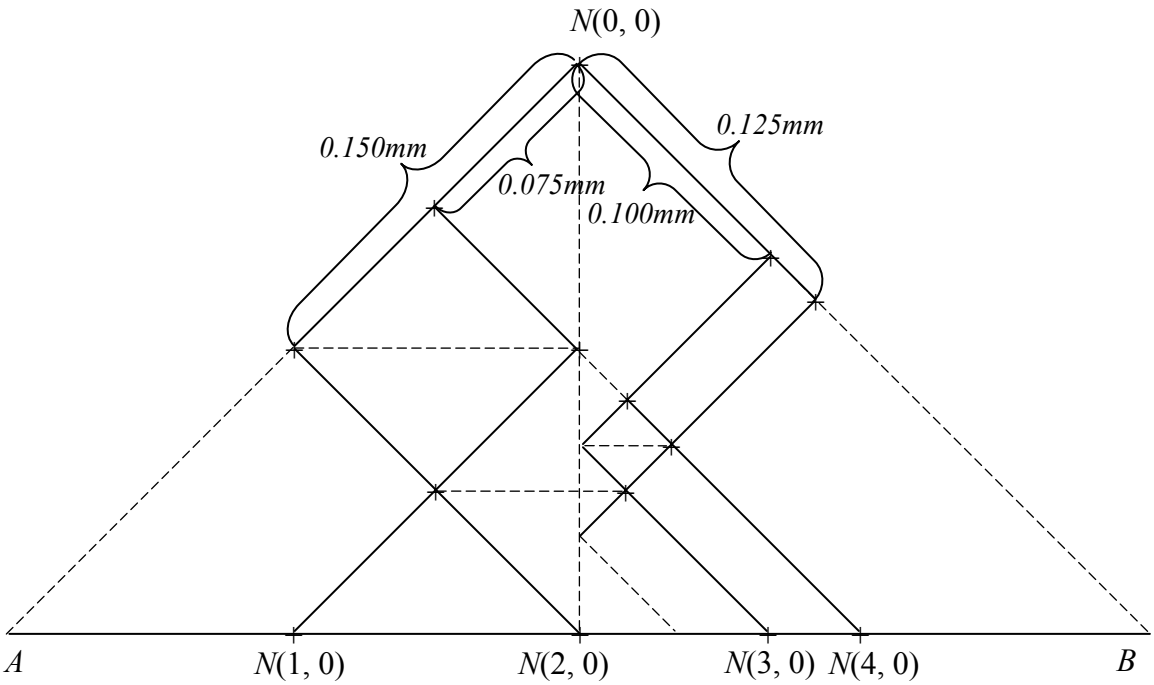


Figure 7.1 Schematic representation of sectioning strategy

After all of the alternative cutting strategies are determined by the procedure analyzed above, the optimization algorithm is used to find the optimal cutting conditions for each pass. After finding the optimal cutting conditions for all possible cutting strategies shown in Table 7.1, the best strategy is selected as the optimal one according to the corresponding optimization objective(s). Finally the number of passes and the cutting parameters (depth of cut, feed rate and cutting speed) associated with each pass can be used in practical machining.

Table 7.1 Alternative cutting strategies for a total depth of cut of 0.3 mm

Cutting strategy No.	Number of sections to be cut for each strategy			
	Pass 1 (mm)	Pass 2 (mm)	Pass 3 (mm)	Pass 4 (mm)
1	0.075	0.075	0.075	0.075
2	0.075	0.075	0.150	–
3	0.075	0.100	0.125	–
4	0.100	0.100	0.100	–
5	0.150	0.150	–	–

7.3 Constraints

For a meaningful optimization of the machining process, a certain number of constraints should be fulfilled. The following constraints are considered in this study.

7.3.1 Available feed rates and cutting speeds

Optimal feed rate and cutting speed should be in the range of the available feed rates and spindle speeds on the machine tool.

$$f_{Z\min} \le f_Z \le f_{Z\max} = \frac{f_{\max}}{zn_{\min}}$$

$$v_{\min} \le v \le v_{\max} = \frac{\pi D n_{\max}}{1000}$$

7.3.2 Available power

The required power for machining should be less than the effective power transmitted by the machine tool:

$$P_c \le P_m \eta$$

The cutting power P_c is the product of the cutting speed v and the cutting force F_C. So the maximum cutting force allowed at cutting speed v can be calculated by the following equation:

$$F_C = 60000 P_m \eta / v \qquad (7.8)$$

The mechanistic models are used to obtain the cutting forces:

$$F_C = K_t S_i \qquad - \qquad (7.9)$$

Gu et al. (1992) and Stephenson and Bandyopadhyay (1997) indicated that the specific cutting pressure increases for low values of the average uncut chip thickness and the cutting forces normally decrease slightly with increasing cutting speed. So in this study, the slight effect of cutting speed on the specific cutting pressure is neglected. But the uncut chip thickness is small, whose effects on the specific cutting pressure cannot be ignored. In order to get more accurate values of K_t, the maximum uncut chip thickness has been divided into several sections. Before optimization, the mean uncut chip thickness of each section together with average cutting speed of available speed range and axial depth of cut are used to simulate the milling process based on the cutting force model described in Chapter 5. After getting the cutting forces from simulation, the specific cutting pressure can then be calculated from Eq. (7.9). During

the optimization process, when the uncut chip thickness is located in any section, the corresponding value of K_t is used to predict the cutting forces. For the cutting forces constraint, only the maximum value of cutting forces needs to be concerned.

7.3.3 Available cutting forces

Due to the cutting force in milling, the tool holder is subject not only to deflection but also to torsion. Thus, the cutting parameters should be checked with regard to the tool holder strength and its deflection.

$$\frac{(F_t d_a / 2) d_a / 2}{J} \leq \tau_s$$

The deflection of the cutter holder should be kept within a certain value. Assuming the cutter holder as a cantilever beam, the deflection of the cutter at the end can be calculated by:

$$\delta = \frac{F_r L_s^{\,3}}{3EI} \tag{7.10}$$

For the finishing operations, the maximum deflection value δ_{max} is 0.01mm.

7.3.4 Surface finish

The maximum surface finish should be smaller than or equal to the required surface finish under the existing machining conditions. According to the geometry of chip round nose radius in Fig. 5.5, it can be seen that $R_{max} = r(1\text{-}\arcsin(f_z/2r))$. The surface finish values for slot milling can be calculated with the following equation:

$$R_a = \frac{R_{\max}}{4} = \frac{r(1 - \arcsin(f_z / 2r))}{4} \tag{7.11}$$

All the constraints are summarized in Table 7.2.

Table 7.2 Constraints and their expressions in terms of common variables

No.	Constraint	Expression in variables
1	Feed rate constraint	$f_{Z\min} \le f_z \le f_{Z\max} = \frac{f_{\max}}{zn_{\max}}$
2	Cutting speed constraint	$v_{\min} \le v \le v_{\max} = \frac{\pi D n_{\max}}{1000}$
3	Horsepower constraint	$F_c v \le 60000 P_m \eta$
4	Arbor strength constraint	$\frac{F_t d_a^{\ 2}}{4J} \le \tau_s$
5	Arbor deflection	$\frac{F_r L_s^{\ 3}}{3EI} \le \delta_{\max}$
6	Surface finish	$R_a = \frac{r(1-\arcsin(f_z/2r))}{4} \le R_{a_{\max}}$

7.4 Implementation details of PGSA

In this chapter, the hierarchical parallel GA model was used to realize the parallelization of GSA. The general architecture of PGSA is shown in Fig. 7.2. At first, a master task starts up on the master processor to control the operations of the algorithm by broadcasting problem parameters, such as constraints and depth of cut. According to the depth of cut, slave processors are classified into several sub-groups. There is no migration between different sub-groups.

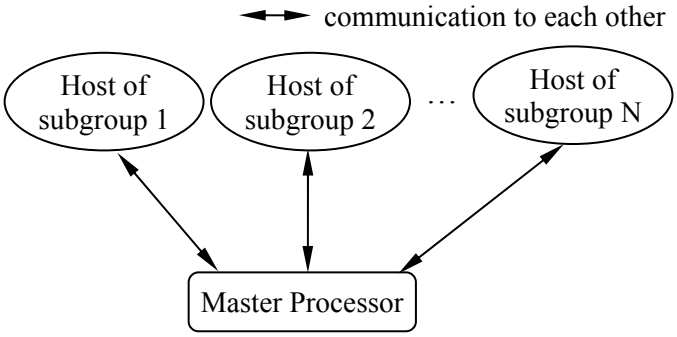


Figure 7.2 Schematic diagram of PGSA's architecture

In each sub-group, the combination model of coarse-grained PGA and master-slave PGA is used to parallelize GSA. A ladder neighborhood relation is adopted in each sub-group, and the 8-processor structure of this relation is shown in Fig. 7.3. The host program, which runs on the host processor, decides the stop criterion for its sub-group. After certain iterations, the host processor sends the stop flag to the eight slave processors, which are marked from 1 to 8 in Fig. 7.3. According to the stop flag, each slave processor decides whether to continue to run GSA or not. The stop flag of PGSA is set on the host processor, if the following criterion is fulfilled:

$$\left| f_{best}^{k} - f_{best}^{k-\Delta k} \right| \le \varepsilon \left| f_{best}^{k} \right|$$

where f_{best}^{k} is best fitness value of a solution on the host process at generation k, and after every Δk generations the stop criterion is checked; f_{best} is the global minimum and ε is a constant value 0.001.

On each slave processor, a sequential GSA runs within its own subpopulation. If the migration conditions are satisfied, each slave processor, say the source processor, finds out the destination processors according to the migration topology. Then the migrant solutions are selected on the source processor and sent to the destination processors. After the migrant solutions are received by the destination processor, the solutions to be deleted are determined and replaced by the received migrant solutions. The same program is executed on each slave processor, but on different data (their own population) until the global optimum is achieved.

After all subgroups converge to their optimal cutting parameters, these parameters are sent to the master processor. Finally the master processor decides the optimal cutting strategy after receiving the optimal solutions from sub-groups.

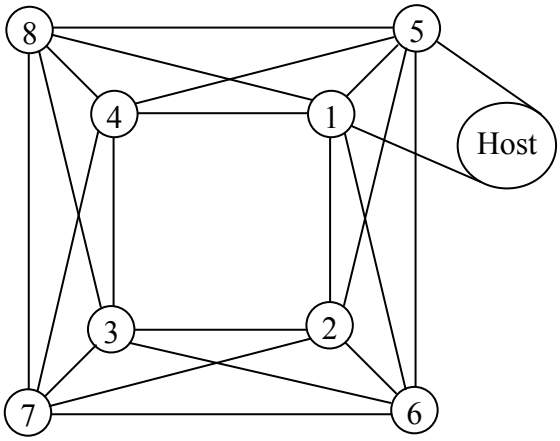


Figure 7.3 Schematic diagram of the implementation of each subgroup

7.4.1 Assignment of fitness values

In this chapter, the fitness assignment is based on the concept of Srinivas and Deb's non-dominated sorting genetic algorithm (NSGA) (Deb, 1999). Solutions are firstly classified into different non-domination levels according to their individual objective function values. Then a sharing function method is used to assign fitness to each solution on the basis of its non-domination level. The mechanisms to classify the non-domination levels are shown as follows:

Step 1: Assuming that a population consists of n solutions, each solution has ($m > 1$) objective function values. All solutions are marked as the 'active' and start with $k = 1$.

Step 2:For each active solution x_i in the population, compare it with the other solution x_j ($j = 1, \ldots, n$ and $j \neq i$), if x_j is no worse than solution x_i for all m objectives, and solution x_j is strictly better than solution x_i in at least one objective, x_i is dominated by x_j.

Step 3: After all solutions ($i = 1, \ldots, n$) in the population have been considered, those that are not dominated by any other solutions are identified and non-dominated

solutions are assumed to be the *k*-th non-dominated front of the population. These non-dominated solutions are set as inactive members.

Step 4: If there are still some active or un-classified solutions left, increment *k* by one and go to Step 2 until all of the solutions have been classified into a level of non-domination.

After the classification to different non-domination levels, the fitness values of solutions are assigned according to their non-dominated levels. Since solutions at lower levels of non-domination are better, those at a lower level get higher fitness values. Thus, a selection mechanism that selects solutions with higher fitness provides a search direction towards the Pareto-optimal region.

Two stages are used to assign the fitness values to solutions. Firstly, all solutions at a particular non-domination level are assigned an identical dummy fitness. Thereafter, based on the crowding of solutions in the non-domination front, lonely solutions are emphasized by using a sharing function strategy. On the contrary, a solution's dummy fitness value is reduced by a factor according to the sharing strategy, if it has many neighbors in the same front.

The shared fitness assignment starts from the first non-dominated front. Initially, all solutions in the first non-dominated front are assigned a fitness value equal to a certain positive value. Based on the fitness sharing strategy, the dummy fitness of each solution is updated with a factor, which depends on the number and proximity of neighboring solutions. Then, the shared fitness of solutions in the first front is calculated based on the dummy fitness. Subsequently the smallest shared fitness value

is also determined for later use. The smallest shared fitness value from the previous level is assigned to all solutions in following non-domination level as their initial dummy fitness values. With the sharing method, the shared fitness of each solution in the following front can be calculated, and the smallest shared fitness value is also determined for later use. If there still exists a non-domination level without assigned shared-fitness values, this procedure is continued till all solutions are assigned a shared fitness.

During the shared fitness assignment, any solution in the previous front has a higher shared fitness value than solutions in the current front. This can help to maintain the search direction towards the Pareto-optimal region. Apart from the search direction, it is also important to maintain the population diversity among same non-domination level, so that a good distribution of solutions in the Pareto-optimal front can be obtained. The fitness sharing method was originally introduced by Holland (1975); then Goldberg and Richardson (1987) have improved it further. The main principles of it are presented as follows.

The shared fitness value f_i' of a solution i in the k^{th} non-domination front is given as follows:

$$f_i' = f_k / m_i \tag{7.12}$$

where f_k is the dummy fitness for all solutions in the k^{th} non-domination front, and m_i is known as the niche count with which the fitness f_k is shared. The niche count m_i is calculated as:

$$m_i = \sum_{j=1}^{N_k} sh(d_{ij}) \tag{7.13}$$

where N_k denotes the number of solutions in the k^{th} non-domination front, and d_{ij} represents the distance between two solutions i and j in the k^{th} front. The sharing function (sh) measures the similarity level between two solutions in the same front (Sareni and Krahenbuhl, 1998). It returns one if the elements are identical, zero if their distance is higher than a threshold of dissimilarity, and an intermediate value at intermediate level of dissimilarity. The most widely used shared function is as follows:

$$sh(d_{ij}) = \begin{cases} 1-(d_{ij}/\sigma_s)^{\alpha}, & if \quad d_{ij} \le \sigma_s \\ 0 & otherwise \end{cases} \tag{7.14}$$

where σ_s is the threshold of dissimilarity, and α is a constant parameter which regulates the shape of the sharing function. α is commonly set to one with the resulting sharing function referred to as the triangle sharing function (Sareni and Krahenbuhl, 1998).

In this study, the real-value vector is used to encode the solutions, so the distance d_{ij} between two solutions i and j can be measured as the Euclidian distance:

$$d_{ij} = \sqrt{\sum_{t=1}^{n}(x_t^{(i)} - x_t^{(j)})^2} \tag{7.15}$$

where n is the length of the chromosome. In practical application, the values of the range for each variable in the chromosome are used to normalize the distance d_{ij} as follows:

$$d_{ij} = \sqrt{\sum_{t=1}^{n}(\frac{x_t^{(i)} - x_t^{(j)}}{x_t^u - x_t^l})^2} \tag{7.16}$$

where x_t^u and x_t^l are the upper and lower bounds of variable x_t.

This fitness assignment procedure starts from the first non-domination front. After all solutions in each front update their shared fitness values, the minimum shared fitness value of them is set as the initial dummy fitness values for solutions in the next non-

domination front. This procedure is continued until all of the solutions have their shared fitness values.

7.4.2 Selection

Based on the shared fitness assignment, solutions are appropriately selected for recombination. As described in Section 6.3.2 of Chapter 6, there are mainly three selection procedures: proportional selection, tournament selection and rank-based selection. For rank-based approach, the solutions are sorted based on their fitness values, and the rank N is assigned to the best solution and the rank 1 to the worst one. Then based on their ranks, the selection probability is assigned to the solutions. Rank-based selection behaves in a more robust manner than the proportional one. Therefore in this chapter, rank-based fitness selection approach is chosen.

7.4.3 Crossover and Mutation

The values of cutting speed and feed rate are taken in a continuous domain, so it is suitable to choose real-valued encoding for solving the optimization problem. For real-value encoding, the conventional crossover operator combines substrings belonging to their parents. Obviously, it will not change the value of each variable, so it cannot perform the search with respect to each variable. Therefore the conventional crossover operator is not suitable in this study and consequently, a modified crossover operator, called convex recombination, has been tried. The details of this operator are shown in Section 6.3.3 of Chapter 6.

For the real-value encoding scheme, there are different mutation operators that can be used, such as uniform distributed mutation, Gaussian mutation, range mutation and

non-uniform mutation. As described in Section 6.3.3 of Chapter 6, uniform distributed mutation seems to be used for searches in a large region, and Gaussian mutation performs better searches in a small local area. The first two mutation methods have been used in this chapter. At the initial stage, uniform distributed mutation is used, when the decreasing rate of the average fitness values is less than 0.01. The mutation type is changed to Gaussian mutation.

7.4.4 Migration policy, rate, frequency and topology

In the implementation of PGSA, the master-slave/coarse-grained PGSA was running in each sub-group. The host program runs on the host processor, which decides on the termination criterion for its sub-group, and GSA is executed on the slave processors with their own sub-populations. After certain intervals, migration will take place among slave processors in their neighborhood according to the migration topology. Only the processors in the same sub-group can exchange their solutions. Therefore, there are some parameters to be concerned: migration rate, migration policy, frequency of migration, and logical arrangement of the subpopulations. Two methods to choose the solution to migrate are often attempted: elitist strategy and probabilistic tournament selection. Gordon and Whitley (1993) compared these two methods and indicated that elitist strategy performed better. Thus we used elitist strategy to choose migration solutions and the top 1% best solutions are migrated to replace the worst solutions of other subpopulations. If 1% of population size is not an integer, it will be rounded off to the smallest integer that is greater than its fractional value. In this study, the 5-neighbor ladder migration scheme is used, where solutions are exchanged between two connected slave processors as shown in Fig. 7.3, which also shows that every slave processor is connected to five other slave processors. The advantage of this topology is

that it spreads the migrant solutions quickly among the slave processors. After every ten intervals, slave processors within a same sub-group exchange their migration solutions according to the migration topology.

For PGA and PGSA, there are many control parameters. Many researchers have done a lot experiments to discuss the effects of these parameters on the final solutions. Although for different depths of cut, the search space is different, in order to show the robustness of PGSA, there is no tuning of all of these parameters to a specific depth of cut. Based on the research results in (Goldberg 1989) and (Dumitrescu et al. 2000), and the parameters for GSA used in (Wang et al. 2004), the parameters of PGSA used in this study are shown in Table 7.3.

Table 7.3 Control parameters setting of PGA and PGSA used in each sub-group

Parameters	PGA	PGSA
Number of sub-group	4	4
Number of subpopulations of each sub-group	8	8
Number of solutions of each subpopulation	100	100
Number of neighbors of each slave processor	5	5
Migration interval (generations)	10	10
Number of migration solutions	top 1% best	top 1% best
Mutation probability	0.2	0.5
Crossover probability	0.70	0.70
Cooling rate	0.85	
Initial temperature	600	

7.5 Application examples

All tests of PGA and PGSA were run on SUN workstation network, which consists of 42 SUN Blade 2000 workstations with a fast Ethernet interconnect. In this study, 37 workstations have been used.

7.5.1 Example 1

An example is optimized with the proposed methodology based on the PGA and PGSA. The final dimension of the workpiece is shown in Fig. 7.4. Specifications of the required parameters and values of the constants are given as follows: Type of machining is pocket milling the bottom surface of the pocket, and total depth of cut is 0.60mm. P_m= 18.5kW, η= 0.7, d_a = 18mm, L_s= 100mm, τ_s=120MPa, E =211GPa, Cutting speed ranges: 200~500m/min, feed rate ranges: 0.050~0.300mm/r, Tool material is binderless CBN, z =1, D =20mm, workpiece material is titanium alloys Ti-6Al-4V, its tensile strength is 930MPa, L=160 mm, B =37mm, a = 5mm, T_a = 0.1min/part, T_s =10min, C_o = \$1.45, C_l = \$0.45, C_t = \$100.

According to the description about the cutting force constraint in Section 7.3.2, the specific cutting can be calculated as shown in the following table.

Table 7.4 Cutting pressure estimation K_t (N/mm^2) under different cutting conditions

Max. chip thickness value(mm)	K_t (N/mm^2)			
	a = 0.075mm	a = 0.100mm	a = 0.125mm	a = 0.150mm
0.050-0.075	4361.2	4159.8	3344.9	3179.3
0.076-0.100	3823.5	3565.3	3430.9	3325.0
0.101-0.150	3557.6	3397.1	3345.0	3263.0

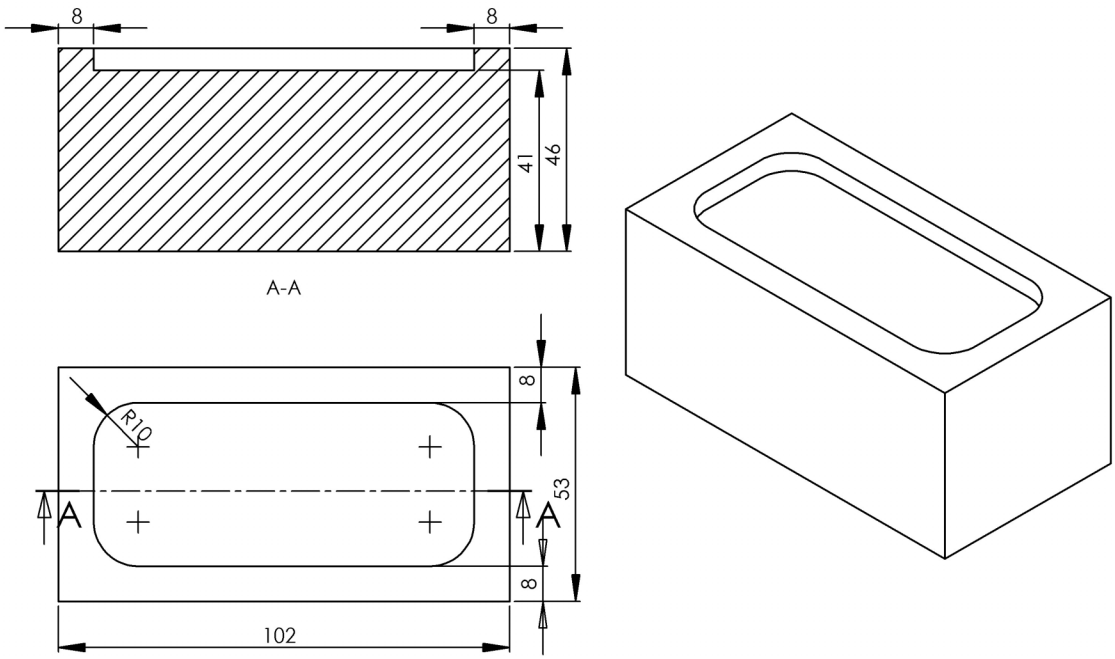


Figure 7.4 Geometrical dimension of test part 1

After the termination criterion is satisfied, multiple Pareto-optimal solutions are obtained as shown in Figs. 7.5-7.8 at four different axial depths of cut. All the non-dominated points are fitted with Polynomial curve fitting methodology, the fitted lines are indicated as the Pareto-optimal front as shown in Figs. 7.5-7.8. Although there are still some dominated solutions in the population, Fig. 7.9 shows that more than 80% of solutions are converted close to the Pareto-optimal front.

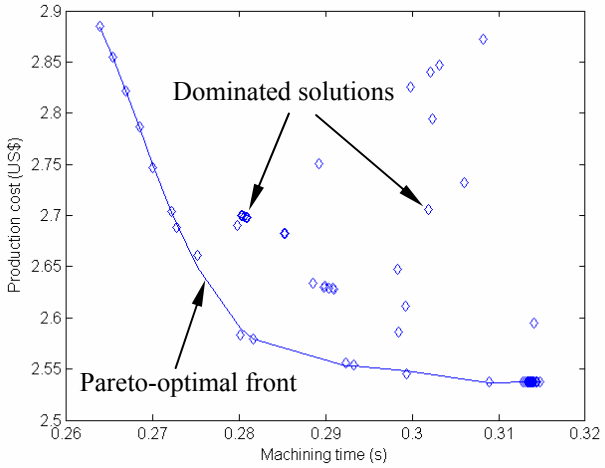


Figure 7.5 Population after termination of simulation at a = 0.075mm for test part 1

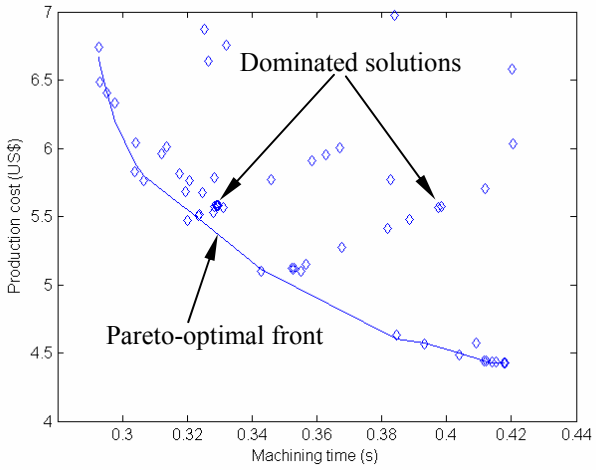


Figure 7.6 Population after termination of simulation at a = 0.100mm for test part 1

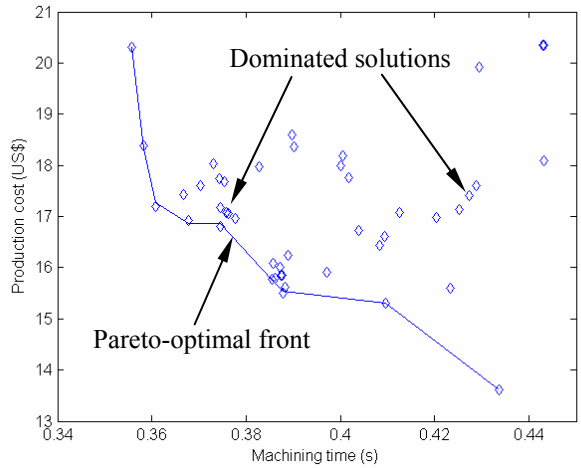


Figure 7.7 Population after termination of simulation at a = 0.125mm for test part 1

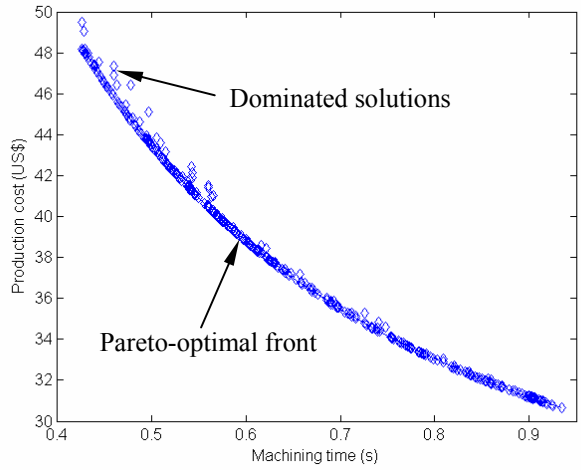


Figure 7.8 Population after termination of simulation at a = 0.150mm for test part 1

With non-dominated sorting PGSA, multiple Pareto-optimal solutions are obtained in a single simulation. Normally, a single solution or several particular solutions are chosen from the obtained Pareto-optimal solutions by some techniques, such as post-optimal techniques, compromise programming, marginal rate of return and weighted average, etc. (Deb, 1999). If the optimal solutions at certain user-specified weighted average of the objective functions locate in a convex Pareto-optimal region, weighted average strategy can be used to choose the optimal solution(s). In practical machining, the cutting strategy with less production time is more preferred; and Figs. 7.5-7.8 show that in the smaller production time region, the corresponding Pareto-optimal region is convex. In addition, the weighted average method can be implemented easily. Therefore, in this chapter this method is used to find the optimal solution from the Pareto-optimal front. Unlike the classical weighted average scheme for multi-objective optimization, this procedure described here is used to choose the optimal solution from the Pareto-optimal solutions found. For the classical weighted average scheme, based on the values of weighted average of the objectives, only one solution would be found to be the optimal one. Furthermore, with a different weight vector, a different optimal solution can be found. For the given example, the values of these two objectives are in different scale. Thus, data normalization is needed to rescale the values in the same range. Here Min-Max normalization is used to transform the data into the range [0, 1]. After rescaling the values of these two objectives, the weighted average strategy is then employed to find the optimal solutions from the Pareto-optimal front. Suppose for a given weight vector (w_1, w_2), the optimal solution is found by searching the minimum value of $w_1\overline{T}_{pri} + w_2\overline{C}_{pri}$ for all the solutions in the Pareto-optimal front, where $\overline{T}_{pri}$ and $\overline{C}_{pri}$ are normalized values of the production time and production cost,

respectively. The optimal solutions chosen with a weight vector (w_1, w_2) = (0.80, 0.20) for all cases are shown in Table 7.5.

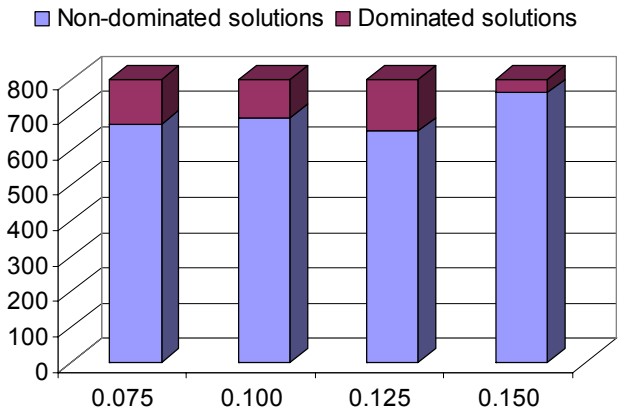


Figure 7.9 Number of non-dominated and dominated solutions under different depths of cut

Table 7.5 Optimal solutions to test part 1 with weighted average strategy

Depth of cut (mm)	Cutting speed (m/min)	Feed rate per tooth (mm/r)	Production time (min)	Production cost (US$)
0.075	417.62	0.085	0.264	2.88
0.100	436.12	0.073	0.293	6.48
0.125	437.12	0.060	0.361	17.17
0.150	438.30	0.050	0.427	48.12

For comparison, PGA has also been used to optimize cutting parameters with the same termination criterion. Based on 30 runs, the average computation times with PGA and PGSA were also recorded, as listed in Table 7.6, which shows that less computation time is need with PGSA than that with PGA. The reason for this may be that PGSA maintains better population diversity at the initial stage, and the local selection of SA after crossover and mutation can narrow down the search space at the later stage. Full details for this explanation are given in Chapter 6. Thus, PGSA is more effective for optimizing the milling process.

Table 7.6 Average computation time using PGA and PGSA

Optimization method	Calculation time (min)			
	a = 0.075mm	a = 0.10mm	a = 0.125mm	a = 0.150mm
PGA	7.875071	9.983201	10.99675	13.70998
PGSA	6.123077	6.621389	6.342372	7.308660

where a is axial depth of cut

After getting the optimal cutting parameters at different depths of cut, it needs to evaluate the production time and production cost under different depths of cut. Table 7.5 indicates that with the increase of the depth of cut, the production time per pass increases slightly. But at higher depth of cut, less cutting passes are needed, so it takes slightly less time to remove the same amount of depth of cut with higher depth of cut. However, the production cost increases significantly or even super-linearly with the increase in the depth of cut. Therefore, in considering the trade-off between the production time and production cost, it is better to select the cutting conditions at the lower depth of cut of 0.075mm than those for the other three cases. So the most optimal cutting strategy is at depth of cut 0.075mm, cutting speed 417.62m/min, feed rate per tooth 0.085mm/r, and eight cutting passes are needed to remove the stock.

Because the BCBN tool is a new type of cutting tools, in the machining data book, there are no recommended cutting conditions available for high-speed milling of Ti-6Al-4V with BCBN tools. Fortunately, Zareena (2002) used the same type of tool for high-speed milling of Ti-6Al-4V and recommended optimal cutting conditions based on the experiment results. In this study, almost the same experimental setup as Zareena's has been used. Thus, it is trustable to compare our results with the recommended cutting conditions by Zareena (2002). Zareena (2002) drew a conclusion that better cutting performance of BCBN tools is found at higher cutting speed

(400m/min), low feed rate per tooth (0.05mm/rev) and low axial depth of cut (0.05mm). Under such cutting condition, the production time and production cost are 0.469min and US$3.06, respectively. Obviously, the cutting performance in terms of production time and production cost under this condition is worse than that at the depth of cut 0.075mm, cutting speed 417.62m/min and feed rate per tooth 0.085mm/r. Therefore, PGSA is effective to find more optimal cutting conditions for high-speed milling of Ti-6Al-4V.

The test component was machined to verify the optimal cutting parameters obtained from optimization simulation. The new inserts were employed and the oil based coolant was used. Results proved to be satisfactory.

7.5.2 Example 2

Now a more complicated example is considered, its final dimension is shown in Fig. 7.10. The BCBN tool is used for finishing machining the bottom surface of the pocket and the slot in the inner island, and the total depth of cut is 0.30mm.

The width of the pocket and slot is greater than the diameter of the cutter, so two passes are needed to remove one layer of the bottom surface. Contour-parallel tool pass pattern has been chosen as the pocketing strategy. For the pocketing, the first pass is to reduce the inward for the outer contour, and the second one is to reduce the outward for the inner island. In the first cutting pass, the radial depth of cut is equal to the diameter of the cutter, the same procedure as that used in Section 7.5.1 can be used to find the optimal cutting strategy. While for the second pass, the difference between the width of pocket and the diameter of the cutter is the actual radial depth of cut. Different radial depth of cut results in different maximum instantaneous undeformed

chip thickness. The same situation exists for the two cutting passes of the slotting milling in the inner island. In the following parts, we focus on the optimization of the second pass, where the maximum value of $h(\varphi)$ is less than that of feed rate per tooth.

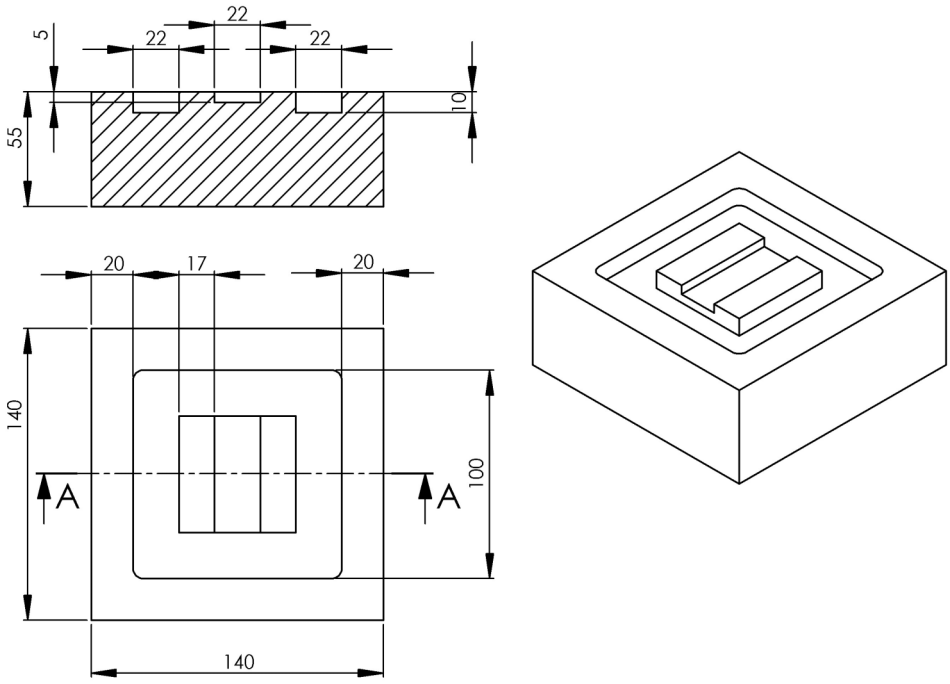


Figure 7.10 Geometrical dimension of test part 2

When the radial depth of cut is less than the diameter of the cutter, assumed that there is no insert throw and tool eccentricity, the ideal cutting zone geometry is shown in Fig. 7.11, where a_r is the radial depth of cut.

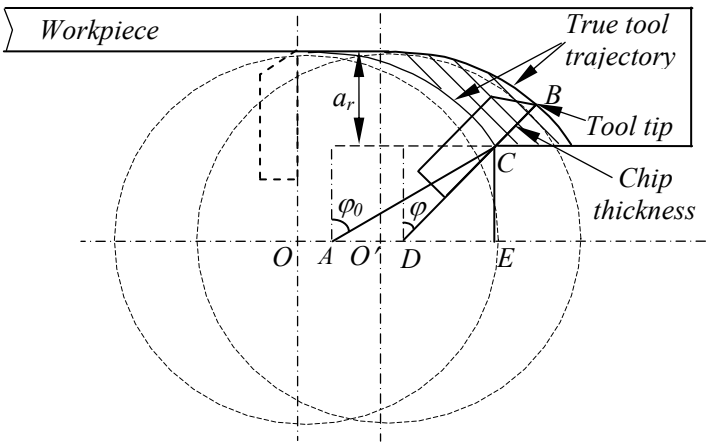


Figure 7.11 Cutting zone geometry for the second pass of pocketing

The cutting forces and machining power constraints of the optimization model are more relative to the maximum uncut chip thickness. Therefore, the maximum uncut chip thickness needs to be predicted based on the cutting zone geometry as shown in Fig.7.11. According the Eq. (5.32) of instantaneous chip thickness in Chapter 5, the maximum chip thickness $\overline{BC}$ in Fig. 7.11 can be derived as follows:

$$\overline{BC} = R - \overline{CD} = \frac{\sin(\pi/2+\varphi) - \sin(\pi/2-\varphi-\delta)}{\sin(\pi/2+\varphi)} R \tag{7.17}$$

where $\varphi = \varphi_0 - \delta$, $\varphi_0 = \frac{\pi}{2} - \arcsin(\frac{R-a_r}{R})$, Let $a = \frac{f}{R}\sin(\frac{\pi}{2}+\varphi)$, then

$$\delta = a - \frac{1}{2\pi}a^2 + \frac{3+2\pi^2}{12\pi^2}a^3 - \frac{3+8\pi^2}{24\pi^3}a^4 + \frac{15+100\pi^2+18\pi^4}{240\pi^4}a^5 + O(a^6)$$

In order to get the value of $\overline{BC}$, the only unknown parameter is φ. In this study, since f/R is very small, it means a is very small, then δ can be approximated as a.

Let $a_0 = f/R$, then the following equation can be got:

$$\varphi = \varphi_0 - \delta = \varphi_0 - a = \varphi_0 - a_0\cos\varphi \tag{7.18}$$

After solving the above equation, φ can be represented as follows:

$$\varphi = \varphi_0 - a_0\cos\varphi_0 - \sin\varphi_0\cos\varphi_0 a_0^{\,2} + O(a_0^{\,3}) \tag{7.19}$$

After obtaining the value of φ, the value of $\overline{BC}$ can be calculated further. Then, the constraints of the cutting forces and machining power can be justified. With the algorithm described in Section 7.3, multiple Pareto-optimal solutions are obtained in a single simulation as shown in Figs. 7.12-7.15, which show that Pareto-optimal regions have a convex shape, so weighted average strategy can also be used to choose the optimal solution from the Pareto-optimal regions as described in Section 7.5.1. With a weight vector (w_1, w_2) = (0.80, 0.20), the optimal cutting parameters at four axial depths of cut are obtained, as shown in Table 7.7.

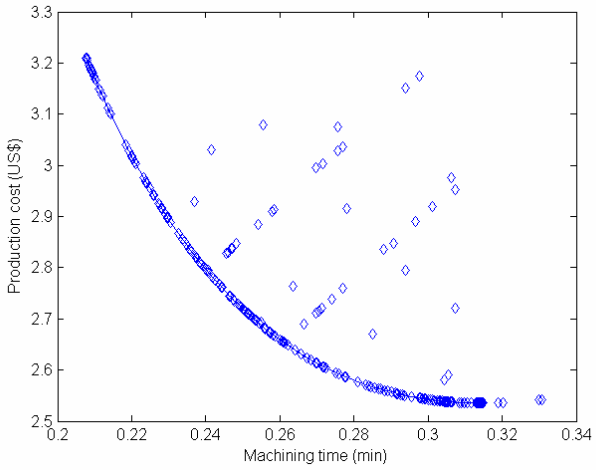


Figure 7.12 Population after termination of simulation at a = 0.075mm for test part 2

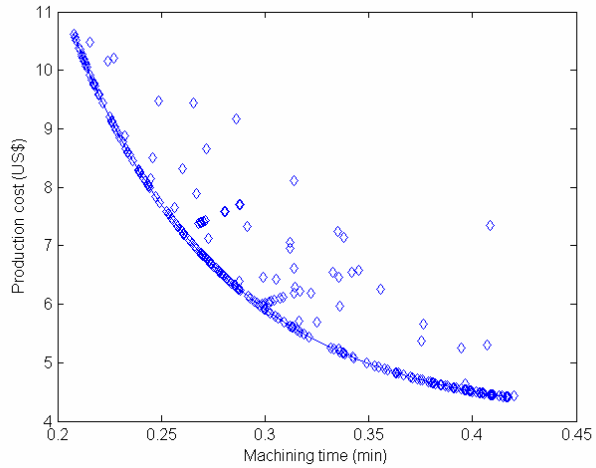


Figure 7.13 Population after termination of simulation at a = 0.100mm test part 2

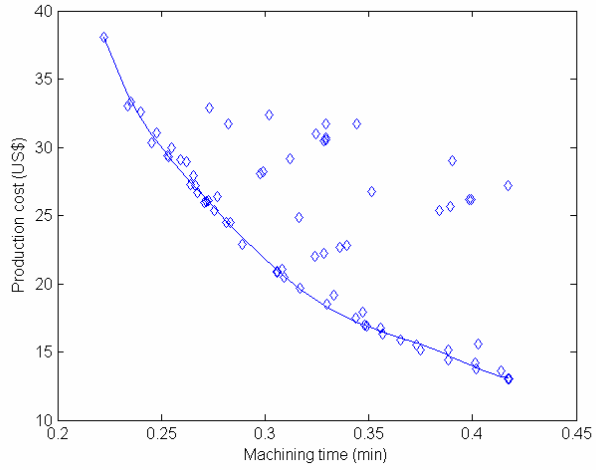


Figure 7.14 Population after termination of simulation at a = 0.125mm test part 2

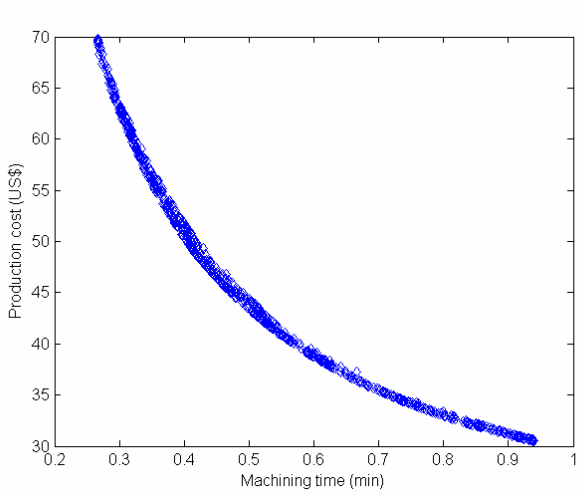


Figure 7.15 Population after termination of simulation at a = 0.150mm test part 2

Table 7.7 Optimal solutions to test part 2 with weighted average strategy

Depth of cut (mm)	Cutting speed (m/min)	Feed rate per tooth (mm/r)	Production time (min)	Production cost (US$)
0.075	450.00	0.100	0.2078	3.2100
0.100	450.00	0.100	0.2078	10.6193
0.125	449.62	0.089	0.2337	32.9934
0.150	450.00	0.078	0.2682	68.3204

Table 7.7 shows that the production cost increases significantly or even super-linearly with the increase in the depth of cut. Although with higher depth of cut, it can save the production time to a certain extent, it results in a significant increase in the production cost. Therefore, cutting conditions at the lower depth of cut are preferred. Finally, cutting parameters at the depth of cut of 0.075mm are much optimal than those for the other three cases.

When considering the production time and production cost of two passes, the optimal cutting strategy is obtained when the depth of cut is equal to 0.075mm. For first pass, the optimal cutting speed and feed rate per tooth are 417.62mm/min and 0.085mm/r,

respectively, which is same as that obtained in first application example. For the second pass, the optimal cutting parameters are at depth of cut 0.075mm, cutting speed 450.0 m/min and the feed rate 0.100mm/r. In all, four cutting passes are needed to remove the stock. Test component was also tried, and the experimental results were proved to be satisfactory.

7.6 Conclusions

In this chapter, PGSA is presented to optimize the cutting conditions in pocket milling (feed rate, speed and depth of cut) subject to a set of constraints; and two objectives are concerned. For this multi-objective optimization problem, the concept of NSGA is used to handle multi-objectives and assign the shared fitness values to each solution. With this fitness assignment, the selection operator of the proposed PGSA algorithm results in the search direction towards the Pareto-optimal region. From the Pareto-optimal region, the weighted average strategy was then used to choose the optimal solutions based on the user's preference. For all the possible depths of cut of the given examples, after obtaining the Pareto-optimal solutions with PGSA, the weighted average strategy was used to choose the optimal cutting parameters. When the optimal cutting parameters were obtained, the values of production time and production cost under all possible optimal cutting strategies were determined. Finally, based on the comparison of the values of two objective functions among possible cutting strategies, the optimal cutting strategy was obtained. From comparison with the recommended cutting parameters from the other researcher and practical machining tests, it can be seen that PGSA together with non-dominated sorting methodology is much more suitable for multi-objective optimization of the cutting parameters for milling operation.

Chapter 8

Conclusions

8.1 Main contributions

In this study, the main work is composed of four parts: investigation of machining mechanism of high-speed milling of Ti-6Al-4V; modeling of milling geometry and cutting forces; development of advanced search algorithm PGSA; and multi-objective optimization of high-speed milling process, which is based on the preceding three parts. The algorithm proposed in this thesis has been benchmarked with a standard test set, and models of milling process developed have been verified through practical experiments. The main contributions of this study are summarized as follows:

1. **Analysis of the cutting performance of BCBN for high-speed milling of Ti-6Al-4V and investigation of tool wear mechanism of BCBN tools**. In this study, the BCBN tools have longer tool life than that of CBN and PCD tools obtained by Zareena (2002). With the excellent mechanical properties (such as hardness and strength at higher temperature), the binderless CBN has much potential to be a new cutting tool material for high-speed machining of titanium alloys with superior cutting performance. The higher cutting speed and lower feed rate are more optimal cutting conditions. From the observed performance of BCBN tools in high-speed milling of Ti-6Al-4V, it has been found that non-uniform flank wear is the dominant wear pattern of BCBN tools. There is a strong bond at the tool/workpiece interface, and the adhered workpiece material

to the flank face is helpful in reducing wear rate. However, when the adhered workpiece material is subsequently removed, it not only causes accelerated attrition wear on the flank face, but also takes away some removal of aggregate of tool materials. Flank wear of the BCBN tool results not only from attrition, but also from the removal of aggregate of tool materials. Diffusion-dissolution wear for BCBN tools also exists, but it is not the dominant wear mechanism.

2. **Modeling of the milling process geometry and cutting force of high-speed milling of Ti-6Al-4V.** After solving transcendent function about the uncut chip thickness, the analytical solution to it has been obtained accurately. The nose radius effect has been considered in the computation of the uncut chip area. Then, a new equivalent element representation for the non-uniform uncut chip area is presented. In addition, the JC model of flow stress for Ti-6Al-4V is proposed to describe its deformation behavior. Based on above mentioned, a hybrid cutting force model has been developed, based on Oxley's cutting force predictive theory and FEM simulation of high-speed milling of Ti-6Al-4V. The value of strain-rate constant C' is determined from FEM simulation, thus the assumption about the uniform distribution of normal stress at tool-chip interface in Oxley's theory is no longer needed. Experimental milling tests have been done to verify the developed force model. It is found that the force model can predict the cutting forces with reasonable accuracy.

3. **Hybridization of GA and SA and its parallelization**. In this study, a new tightly coupled hybrid model of GA and SA (GSA) is presented, which enhances the strengths of GA and SA and overcomes their weaknesses. In

GSA, SA is incorporated in GA to adjust the mutation rate and to select solutions retained in the temporary population after crossover and mutation. Then, GSA is parallelized to improve its efficiency, and in the parallelization of GSA, a new migration topology has been tried. A set of standard test functions is used to benchmark PGSA in terms of efficiency criteria of the number of function evaluations. The numerical results show that PGSA has faster convergence to global optimal solution than PGA. From the comparison with the other advanced search method BGA, the performance of PGSA is also better using the same termination criterion. More importantly, PGSA has good scaling ability and programmability.

4. **Optimization of high-speed milling of Ti-6Al-4V.** Two objectives are considered: minimum production time and minimum production cost. The idea of non-domination Pareto-optimal is employed to handle the fitness assignment of these two objective functions in the implementation of PGSA. With this fitness assignment, the selection operator of PGSA results in the search direction towards the Pareto-optimal region. For the given optimization examples of milling process, firstly PGSA found the Pareto-optimal front at different depths of cut. Then, the weighted average strategy was used to choose the optimal cutting parameters along the Pareto-optimal region. Finally, the optimal cutting strategy was determined by evaluating the values of production time and production cost under all possible cutting strategies. From the comparison with the recommended cutting parameters from previous work, PGSA together with non-dominated sorting methodology is suitable for multi-objective optimization of the cutting parameters for milling operation.

8.2 Recommendation for future work

In this thesis, the constitutive model of flow stress for Ti-6Al-4V has been used in FEM simulation. If the flow stress data of other workpiece materials, such as Inconel 718, or stainless steel is known, the methodology developed in this study can be applied to simulate the high-speed milling process of this type of materials. Another important process parameter, cutting temperature can also be investigated with the help of this method.

The cutting force model developed in this study is based on high-speed milling with indexable inserts. It can be extended into high-speed milling with ball-end cutters or the milling cutters with various helix angles.

When the adhered workpiece material is taken away from the flank face of BCBN tools, it will cause accelerated flank wear. Investigations can be carried out to find how to keep the adhered material on the flank face as long as possible. For this investigation, cutting temperature at the tool and workpiece interface is needed to be measured, because the temperature is the main reason for the adhesion.

The developed multi-objective optimization algorithm could be used for cutting parameters optimization for other machining processes such as multi-pass turning, drilling or grinding. The developed methodology has been developed in this study, which may also be integrated with existing computer-aided design and manufacturing (CAD/CAM) systems for use in the process planning and optimization in manufacturing industry.

REFERENCES

Agapiou, J.S. Optimization of machining operations based on a combined criterion, part 2: Multipass operations, Journal of Engineering for Industry, Transactions of the ASME, 114(4), pp.508-513. 1992.

Alba, E. and J.M. Troya. A survey of parallel distributed genetic algorithms, Complexity, 4, pp.31-52. 1999.

Altan, T., S.I. Oh and H.L. Gegel. Metal forming: fundamentals and applications. pp. 45-46, Metals Park, OH: American Society for Metals. 1983.

Altintas, Y. Modeling Approaches and Software for Predicting the Performance of Milling Operations at MAL-UBC, presented in CIRP 2000 machining workshop, University of New South Wales, Sydney. (www.malinc.com/pdfs/JMTSRevie.pdf)

Altintas, Y. and E. Budak. Analytical prediction of stability lobes in milling, CIRP Annals - Manufacturing Technology, 44(1), pp.357-362. 1995.

Armarego, E.J.A. and R.H. Brown. The machining of metals. pp. 206-253, Englewood Cliffs, N.J.: Prentice-Hall. 1969.

Armarego, E.J.A., A.J.R. Smith and J. Wang. Constrained optimization strategies and CAM software for single-pass peripheral milling, International Journal of Production Research, 31(9), pp.2139-2160. 1993.

Armarego, E.J.A., J.P. Wang and N.P. Deshpande. Computer-aided predictive cutting model for forces in face milling allowing for tooth run-out, CIRP Annals - Manufacturing Technology, 44(1), pp.43-48. 1995.

Bäck, Th., F. Hoffmeister and H.P. Schwefel. A Survey of evolution strategies, In Proc. Fourth Int'l Conf. Genetic Algorithms, ed by R.K. Belew and L.B. Booker, pp.2-9. M. Kaufmann, San Mateo, Calif. 1991.

Baydar, C. A hybrid parallel simulated annealing algorithm to optimize store performance. In Workshop on GECCO 2002, July 2002, New York, USA.

Bellman, R.E. Dynamic programming. pp. 3-40, Princeton: Princeton University Press. 1957.

Bhaumik, S.K., C. Divakar and A. K. Singh. Machining Ti-6Al-4V alloy with a wBN-CBN composite tool, Materials & Design, 16, pp.221-226. 1995.

Blickle, T. and L. Thiele, A comparison of selection schemes used in evolutionary algorithms, Evolutionary Computation, 4, pp.361-394. 1996.

Brookes, C.A., R.D. James and F. Nabhani. Turning aerospace titanium alloys, Industry diamond review, 36, pp.89-93. 1991.

Brown, D.E., C.L. Huntley and A.R. Spillane. A parallel genetic heuristic for the quadratic assignment problem, In Proc. Third Int. Conf. on Genetic Algorithms, 1989, Morgan Kauffmann, San Mateo, USA, pp. 406-415.

Cantú-Paz, E. A Survey of Parallel Genetic Algorithms. Technical Report IlliGAL 97003, University of Illinois at Urbana-Champaign. 1997.

Cantú-Paz, E. and M. Mejía-Olvera. Experimental results in distributed genetic algorithms. In International Symposium on Applied Corporate Computing, 1994, Monterrey, Mexico, pp. 99-108.

Cantú-Paz, E. Efficient and accurate parallel genetic algorithms. pp. 1-119, Boston: Kluwer Academic Publishers. 2000.

Carrino, L., G. Giuliano and G. Napolitano. Finite element simulation of orthogonal metal cutting, In Computational Methods in Contact Mechanics VI, ed by C.A. Brebbia, pp. 105-114. Southampton: WIT press. 2003.

Chen, M.C. and D.M. Tsai. A simulated annealing approach for optimization of multi-pass turning operations, International Journal of Production Research, 34(10), pp.2803-2825. 1996.

Chen, H. and N. Flann. Parallel Simulated Annealing and Genetic Algorithms: A Space of Hybrid Methods, In Proc. Int'l Conf. Evolutionary computation − PPSN III, Lecture Notes in Computer Science, Vol. 866, ed by Y. Davidor, H.P. Schwefel and R. Manner, pp. 428-438. Berlin: Springer-Verlag. 1994.

Chen, H., N.S. Flann and D.W. Watson. Parallel genetic simulated annealing: a massively parallel SIMD algorithm, IEEE Transactions on Parallel and Distributed Systems, 9, pp.126-136. 1998.

Chipperfield, A. and P. Fleming. Parallel Genetic Algorithms. In Parallel and distributed computing handbook, ed by A.Y. Zomaya, pp. 1118-1143. New York: McGraw-Hill. 1996.

Dearnley, P.A. and A.N. Grearson. Evaluation of principal wear mechanisms of cemented carbides and ceramics used for machining titanium alloy IMI 318, Materials Science and Technology, 2, pp.47-58. 1986.

Deb, K. Multi-objective evolutionary algorithms: Introducing bias among Pareto-optimal solutions, KanGAL report 99002, Indian Institute of Technology, Kanpur, India, 1999.

De Jong, K.A. An Analysis of the Behavior of a Class of Genetic Adaptive Systems. Ph.D. Thesis, University of Michigan. 1975.

Dereli, T., I.H. Filiz and A. Baykasoglu. Optimizing cutting parameters in process planning of prismatic parts by using genetic algorithms, International Journal of Production Research, 39(15), pp.3303-3328. 2001.

Draper, N.R. and H. Smith. Applied regression analysis. pp. 1-69, New York: Wiley. 1981.

Duffuaa, S.O., A.N. Shuaib and M. Alam. Evaluation of optimization methods for machining economics models, Computers & Operations Research, 20(2), pp.227-237. 1993.

Dumitrescu, D., B. Lazzerini, L.C. Jain and A. Dumitrescu. Evolutionary computation. pp.187-211, Boca Raton: CRC Press. 2000.

Ermer, D.S. Optimization of the constrained machining economics problem by geometric programming, Transactions of the ASME, Journal of Engineering for Industry, 93(4) pp.1067-1072. 1971.

Ezugwu, E.O., J. Bonney and Y. Yamane. An overview of the machinability of aeroengine alloys, Journal of Materials Processing Technology, 134, pp.233-253. 2003.

Ezugwu, E.O. and Z.M. Wang. Titanium alloys and their machinability – a review, Journal of Materials Processing Technology, 68, pp.262-274. 1997.

Goldberg, D.E. Genetic algorithms in search, optimization and machine learning, pp. 1-145, Reading, Massachusetts: Addison –Wesley. 1989.

Goldberg, D.E. and J. Richardson. Genetic Algorithms with Sharing for Multimodal Function Optimization, In Proc. the Second international conference on Genetic Algorithms, July 1987, Cambridge, MA, USA, pp. 41-49.

Gordon, V.S. and D. Whitley. Serial and parallel genetic algorithms as function optimizers, In Proc. Fifth Int'l Conf. Genetic Algorithms (ed by S. Forrest), 1993, M. Kaufmann, San Mateo, Calif., pp. 177-183.

Gray, P., W. Hart, L. Painton, C. Phillips, M. Trahan and J. Wagner. A survey of Global Optimization Methods, http://www.cs.sandia.gov/opt/survey/. 1997.

Griewank, A.O. Generalized descent for global optimization, Journal of Optimization Theory and Applications, 34, pp.11-39. 1981.

Gu, F., S.G. Kapoor, R.E. DeVor, and P. Bandyopadhay. A Cutting Force Prediction Model for Face Milling with a Step Cutter, Trans. of the North American Manufacturing Research Institution of SME, 20, pp.361-368. 1992.

Hesser, J. and R. Männer. Towards an optimal mutation probability for genetic algorithms, In Proc. Parallel problem solving from nature: 1st workshop, PPSN I, Lecture Notes in Computer Science, Vol. 496, ed by H.P. Schwefel and R. Männer, pp. 23-32. Berlin: Springer-Verlag. 1991.

Hiroyasu, T., M. Miki and M. Ogura. Parallel Simulated Annealing using Genetic Crossover, In Proc. the IASTED International Conference on Parallel and Distributed Computing Systems, 2000, Las Vegas, USA, pp.145-150.

Holland, J. Adaptation in natural and artificial systems. pp. 1-19, Michigan: University of Michigan Press. 1975.

Huang, Y. and S.Y. Liang. Cutting forces modeling considering the effect of tool thermal property - Application to CBN hard turning, International Journal of Machine Tools and Manufacture, 43(3), pp.307-315. 2003.

Ingber, L. and B. Rosen. Genetic Algorithms and Very Fast Simulated Reannealing: A comparison, Mathematical and Computer Modeling, 16(11), pp.87-100. 1992.

ISO 8688-2, Tool life testing in milling – Part 2: End milling, pp. 1-20. 1989.

Jacobus, K., R.E. DeVor, S.G. Kapoor and R.A. Peascoe. Predictive model for the full biaxial surface and subsurface residual stress profiles from turning, Transactions of the ASME, Journal of Manufacturing Science and Engineering, 123(4), pp. 537-546. 2001.

Jawaid, A., C.H. Che-Haron and A. Abdullah. Tool wear characteristics in turning of titanium alloy Ti-6246, Journal of Materials Processing Technology, 92-93, pp.329-334. 1999.

Jang, D.Y. A unified optimization model of a machining process for specified conditions of machined surface and process performance, International Journal of Production Research, 30(3), pp.647-663. 1992.

Jha, N.K. A discrete data base multiple objective optimization of milling operation through geometric programming, Transactions of the ASME, Journal of Engineering for Industry, 112(4), pp.368-374. 1990.

Johnson, G.R. and Cook, W.H. A constitutive model and data for metals subjected to large strains, high strain rates and high temperatures, In Proc. the 7th International Symposium on Ballistics, 1983, The Hague, Netherlands, pp. 541-547.

Juan, H., S.F. Yu and B.Y. Lee. The optimal cutting-parameter selection of production cost in HSM for SKD61 tool steels, International Journal of Machine Tools and Manufacture, 43 (7), pp. 679-686. 2003.

Kaczmarek, J. Principles of machining by cutting, abrasion and erosion. pp. 262-293, Warsaw, Poland: Peter peregrinus limited. 1976.

Kayacan, M.C., I.H. Filiz, A.I. Sonmez, A. Baykasoglu and T. Dereli. OPPS-ROT: An optimized process planning system for rotational parts, Computers in Industry, 32(2), pp.181-195. 1996.

Kee, P.K. Development of constrained optimization analyses and strategies for multi-pass rough turning operations, International Journal of Machine Tools and Manufacture, 36(1), pp.115-127. 1996.

Kilic, S.E., C. Cogun and D.T. Sen. A computer-aided graphical technique for the optimization of machining conditions. Computers in Industry, 22(3), pp.319-326.1993.

King, R.I. and R.L. Vaughn. A synoptic review of high-speed machining from Salomon to the present, In High speed machining: presented at the winter annual

meeting of ASME, ed by R. Komanduri, K. Subramanian, B.F. von Turkovich, New Orleans, Louisiana, pp. 1-13, New York: ASME, 1984.

König, W. and N. Neises. Turning TiAl6V4 with PCD, Industrial Diamond Review, 53, pp.85-88. 1993.

Kobayashi, S., S.I. Oh and T. Altan. Metal forming and the finite-element method. pp. 1-110, New York: Oxford University Press. 1989.

Kuljanic, E., M. Fioretti, L. Beltrame and F. Miani. Milling titanium compressor blades with PCD cutter, CIRP Annals-Manufacturing Technology, 47, pp.61-64. 1998.

Lee, C.Y. and J.Y. Choi. A genetic algorithm for job sequencing problems with distinct due dates and general early-tardy penalty weights, Computers & Operations Research, 22(8), pp.857-869. 1995.

Lee, W.S. and C.F. Lin. High-temperature deformation behaviour of Ti6Al4V alloy evaluated by high strain-rate compression tests, Journal of Materials Processing Technology, 75(1-3), pp.127-136. 1998a.

Lee, W.S. and C.F. Lin. Plastic deformation and fracture behaviour of Ti-6Al-4V alloy loaded with high strain rate under various temperatures, Materials Science and Engineering A, 241(1-2), pp.48-59. 1998b.

Li, H.Z. Theoretical modeling and simulation of the dynamic processes in milling. Ph.D Thesis, National University of Singapore. 2001.

Liu, Y.M. and C.J. Wang. Modified genetic algorithm based optimization of milling parameters, International Journal of Advanced Manufacturing Technology, 15(11), pp.796-799. 1999.

López de lacalle, L.N., J. Pérez, J.I. Llorente and J.A. Sánchez. Advanced cutting conditions for the milling of aeronautical alloys, Journal of Materials Processing Technology, 100(3), pp.1-11. 2000.

Machado, A.R. and J. Wallbank. Machining of titanium and its alloys. A review, Proceedings of the Institution of Mechanical Engineers, Part B: Management and Engineering Manufacture, 204, pp.53-60. 1990.

Majorell, A., S. Srivatsa and R.C. Picu. Mechanical behavior of Ti-6Al-4V at high and moderate temperatures—Part I: Experimental results, Materials Science and Engineering A, 326 (2), pp.297-305. 2002.

Mahfoud, S.W. and D.E. Goldberg. Parallel recombinative simulated annealing: A genetic algorithm, Parallel Computing, 21, pp.1-28. 1995.

Matthew, J. and Jr. Donachie. Titanium: a technical guide. pp. 79-84, Ohio: Materials Park, ASM International. 2000.

Martellotti, M.E. An analysis of the milling process, Transactions of the ASME, 63, pp.677-700. 1941.

Martellotti, M.E. An analysis of the milling process, part II: Down milling, Transactions of the ASME, 67, pp.233-251. 1945.

McQuillan, A.D. and M.K. McQuillan. Titanium. pp. 361-387, New York: Academic Press. 1956.

Merchant, M.E. Basic mechanics of the cutting process, ASME Journal of the applied mechanics, 67, pp.168-175. 1944.

Merchant, M.E. Mechanics of the metal cutting process, Journal of applied physics, 16(6), pp.318-324. 1945.

Metropolis, N., A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller and E. Teller. Equation of state calculation by fast computing machines, Journal of Chemical Physics, 21, pp.1087-1092. 1953.

Meyer Jr., H.W. and D.S. Kleponis. Modeling the high strain rate behavior of titanium undergoing ballistic impact and penetration, International Journal of Impact Engineering, 26(1-10), pp.509-521. 2001.

Michalewicz, Z. Genetic algorithms + data structures = evolution programs. pp. 33-93, New York: Springer-Verlag. 1996.

Montgomery, D. and Y. Altintas. Mechanism of cutting force and surface generation in dynamic milling, Transactions of the ASME, Journal of Engineering for Industry, 113(2), pp.160-168. 1991.

Montgomery, D.C. Design and analysis of experiments. pp. 228-289, New York: Wiley. 1997.

Montgomery, D.C. Introduction to statistical quality control. pp. 82-149, New York: Wiley. 2001.

Moufki, A., A. Molinari and D. Dudzinski. Modeling of orthogonal cutting with a temperature dependent friction law, Journal of the Mechanics and Physics of Solids, 46(10), pp. 2103-2138. 1998.

Mühlenbein, H. and D. Schlierkamp-Voosen. Predictive models for the breeder genetic algorithm I. Continuous parameter optimization, Evolutionary Computation, 1, pp.25-49. 1993.

Mühlenbein, H., M. Schomisch and J. Born. The parallel genetic algorithm as function optimizer, Parallel Computing, 17, pp.619-632. 1991.

Myers, R.H. and D.C. Montgomery. Response surface methodology: process and product optimization using designed experiments. pp.279-350, New York: Wiley. 1995.

Nabhani, F. Machining of aerospace titanium alloys, Robotics and Computer-Integrated Manufacturing, 17, pp.99-106. 2001.

Ng, E.G., D.K. Aspinwall, D. Brazil and J. Monaghan. Modeling of temperature and forces when orthogonally machining hardened steel, International Journal of Machine Tools and Manufacture, 39 (6), pp. 885-903. 1999.

Oxley, P.L.B. The mechanics of machining: an analytical approach to assessing machinability. pp.23-135, Chichester (England): E. Horwood. 1989.

Ozel, T. Investigation of high speed flat end milling process: prediction of chip formation, cutting forces, tool stresses and temperatures. Ph.D Thesis, The Ohio State University. 1998.

Ozel, T. and T. Altan. Process simulation using finite element method - prediction of cutting forces, tool stresses and temperatures in high-speed flat end milling, International Journal of Machine Tools and Manufacture, 40(5), pp.713-738. 2000.

Pham, D.T. and D. Karaboga. Intelligent optimisation techniques: genetic algorithms, tabu search, simulated annealing and neural networks. pp. 51-91, New York: Springer. 2000.

Petropoulos, P.G. Optimal selection of machining rate variable by geometric programming, International Journal of Production Research, 11(4), pp.305-314. 1973.

Root, R.R. and K.M. Ragsdell. A survey of optimization methods applied to the design of mechanisms, Transactions of the ASME, Journal of Engineering for Industry, 98(3), pp.1036-1041. 1976.

Sandvik hard materials, Cemented carbide rod blanks for Metal Cutting, pp.4, Sandvikens Tryckeri, 2001.

Sareni, B. and L. Krahenbuhl. Fitness sharing and niching methods revisited, IEEE Transactions on Evolutionary Computation, 2(3), pp.97-106. 1998.

SAS Institute. SAS user's guide: statistics, pp.3-13, Cary, N.C., 1985.

Schlierkamp-Voosen, D. and H. Mühlenbein. Strategy Adaptation by Competing Subpopulations, In Proc. Int'l Conf. Evolutionary computation − PPSN III, Lecture Notes in Computer Science, Vol. 866, ed by Y. Davidor, H.P. Schwefel and R. Manner, pp.199-208. Berlin: Springer-Verlag. 1994.

SEI news, http://www.sei.co.jp/sn/2000/09/p1.html, Sep. 2000.

Shaw, M.C. The size effect in metal cutting, Sadhana-Academy Proceedings in Engineering Sciences, 28, pp.875-896. 2003.

Shih, A.J. Finite element simulation of orthogonal metal cutting, Journal of Engineering for Industry, Transactions of the ASME, 117(1), pp. 84-93. 1995.

Shin, Y.C. and Y.S. Joo. Optimization of machining conditions with practical constraints, International Journal of Production Research, 30(12), pp.2907-2919. 1992.

Shunmugam, M.S., S.V. Bhaskara Reddy and T.T. Narendran. Selection of optimal conditions in multi-pass face-milling using a genetic algorithm, International Journal of Machine Tools and Manufacture, 40(3), pp.401-414. 2000.

Siekmann, H.J. How to machine titanium, Tool Engineer, pp.78-82. 1955.

Sirag, D.J. and P.T. Weisser. Toward a unified thermodynamic genetic operator. In Proc. Second Int'l Conf. Genetic Algorithms, ed by J.J. Grefenstette, Hillsdale, N.J.: Lawrence Erlbaum Associates, pp.116-122. 1987.

Somlo, J. and J. Nagy. A new approach to cutting data optimization. In Advances in Computer-aided Manufacture, ed by D. McPherson, pp. 293-303. Amsterdam: North-Holland Pub. Co. 1977.

Sönmez, A.I., A. Baykasoglu, T. Dereli, and I.H. Filiz, Dynamic optimization of multipass milling operations via geometric programming, International Journal of Machine Tools and Manufacture, 39(2), pp.297-320. 1999.

Spiewak, S. Improved model of the chip thickness in milling, CIRP Annals - Manufacturing Technology, 44(1), pp.39-42. 1995.

Stephenson, D.A. and P. Bandyopadhyay. Process-independent force characterization for metal-cutting simulation, Transactions of the ASME, Journal of Engineering Materials and Technology, 119 (1), pp. 86-94. 1997.

Steuer, R.E. Multiple criteria optimization: theory, computation, and application, pp. 138-164, New York: Wiley. 1986.

Sumitomo Electric Industries, Performance cutting tools, pp. 122, Sumitomo Electric Industries: Japan. 2000.

Sumiya, H., S. Uesaka and S. Satoh. Mechanical properties of high purity polycrystalline CBN synthesized by direct conversion sintering method, Journal of Materials Science, 35, pp.1181-1186. 2000.

Tanaka, Y., H. Tsuwa and M. Kitano. Cutting mechanism in ultra-high speed machining, ASME Paper No. 67-Prod-14, 1967.

Tolouei-Rad, M. and I.M. Bidhendi. On the optimization of machining parameters for milling operations, International Journal of Machine Tools and Manufacture, 37(1), pp.1-16. 1997.

Trent, E.M. and P.K. Wright. Metal cutting, pp. 21-96, Boston: Butterworth-Heinemann. 2000.

Uesaka, S. and H. Sumiya. Mechanical properties and cutting performances of high purity polycrystalline CBN compact, American Society of Mechanical Engineers, Manufacturing Engineering Division, MED 10, pp.759-766. 1999.

Uesaka, S., H. Sumiya, H. Itozaki, J. Shiraishi, K. Tomita and T. Naka. Cutting Tools Using High-Purity Polycrystalline Cubic Boron Nitride Sintered Bodies, SEI technical review, 50, pp.34-40. 2000.

Usui, E. and T. Shirakashi. Mechanics of machining - from descriptive to predictive theory, American Society of Mechanical Engineers, Manufacturing Engineering Division, MED 7, pp.13-35. 1982.

van Luttervelt, C.A., T.H.C. Childs, I.S. Jawahir, F. Klocke and P.K. Venuvinod. Present situation and future trends in modelling of machining operations. Progress report of the CIRP working group 'Modelling of Machining Operations', CIRP Annals - Manufacturing Technology, 47(2), pp.587-626. 1998.

Varanelli, J.V. and J.C. Cohoon. Population-Oriented Simulated Annealing: A Genetic/Thermodynamic Hybrid Approach to Optimization. In Proc. Sixth Int'l Conf. Genetic Algorithms, 1995, M. Kaufman, San Francisco, Calif., USA, pp.174-181.

Wang, J. Multiple-objective optimization of machining operations based on neural networks, International Journal of Advanced Manufacturing Technology, 8(4), pp.235-243.1993.

Wang, J. Computer-aided economic optimization of end-milling operations, International Journal of Production Economics, 54(3), pp.307-320. 1998.

Wang, J. and E.J.A. Armarego. Computer-aided optimization of multiple constraint single pass face milling operations, Machining Science and Technology, 5(1), pp.77-99. 2001.

Wang, J., T. Kuriyagawa, X.P. Wei and D.M. Guo. Optimization of cutting conditions for single pass turning operations using a deterministic approach, International Journal of Machine Tools and Manufacture, 42(9), pp.1023-1033. 2002.

Wang, Z.G., Y.S. Wong and M. Rahman. High-speed milling of titanium alloys using binderless CBN tools, International Journal of Machine Tools and Manufacture, 45(1), pp.105-114. 2005.

Yang, X.P. and C. R. Liu. Machining titanium and its alloys, Machining Science and Technology, 3, pp.107-139. 1999.

Zareena, A.R. High-speed machining of titanium alloys. Master thesis, National University of Singapore. 2002.

Zener, C.M. A Mathematical Aid in Optimizing Engineering Designs, Proceeding of the National Academy of Science, 47(4), pp.537-539. 1961.

Zerilli, F.J. and R.W. Armstrong. Dislocation-mechanics-based constitutive relations for material dynamics calculations, Journal of Applied Physics, 61(5), pp.1816-1825. 1987.

Zorev, N.N. Interrelationship between shear processes occurring along tool face and on shear plane in metal cutting. In Proceedings of the Conference on International Research in Production Engineering, ASME, 1963, New York, USA, pp. 42–49.

Zoya, Z.A. and R. Krishnamurthy. The performance of CBN tools in the machining of titanium alloys, Journal of Materials Processing Technology, 100(1-3), pp.80-86. 2000.

PUBLICATION LIST

Journal papers

[1] Z.G. Wang, Y.S. Wong and M. Rahman, Optimisation of multi-pass milling using genetic algorithm and genetic simulated annealing, International Journal of Advanced Manufacturing Technology, Vol.24 (9-10), pp. 727-732, 2004.

[2] Z.G. Wang, Y.S. Wong and M. Rahman, High-speed milling of titanium alloys using binderless CBN tools, International Journal of Machine Tools and Manufacture, Vol.45 (1) pp. 105-114, 2005.

[3] Z.G. Wang, M. Rahman and Y.S. Wong, Tool wear characteristics of binderless CBN tools used in high-speed milling of titanium alloys, Wear, 258, 2005, pp. 752-758.

[4] Z.G. Wang, M. Rahman, Y.S. Wong and J. Sun, Optimization of multi-pass milling using parallel genetic algorithm and parallel genetic simulated annealing, International Journal of Machine Tools and Manufacture, (Article in press).

[5] Z.G. Wang, Y.S. Wong and M. Rahman, Development of a parallel optimization method based on genetic simulated annealing algorithm, Parallel Computing, (Accepted for publication).

[6] Z.G. Wang, M. Rahman and Y.S. Wong, A hybrid cutting force model for machining of Ti6Al4V, *CIRP annals*, 2005, (Accepted for publication).

[7] N. He, Z.G. Wang, C.Y. Jiang and B. Zhang, Finite element method analysis and control stratagem for machining deformation of thin-walled components, Journal of Materials Processing Technology, 139(1-3), 2003, pp. 332-336.

[8] J. Sun, G.S. Hong, Y.S. Wong, M. Rahman, and Z.G. Wang, Effective Training data selection in Tool Condition Monitoring System, submit to International Journal of Machine Tools and Manufacture, (Accepted for publication).

[9] L. Li, N. He, M. Wang and Z.G. Wang, High speed cutting of Inconel 718 with coated carbide and ceramic inserts, Journal of Materials Processing Technology, 129(1-3), 2002, pp. 127-130.

[10] Z.G. Wang, Y.S. Wong, M. Rahman and J. Sun, Multi-objective optimization of high-speed milling with parallel genetic simulated annealing, International Journal of Advanced Manufacturing Technology (submitted).

[11] J. Sun, Y.S. Wong, M. Rahman, G.S. Hong, and Z.G. Wang, Tool Condition Identification Framework in Titanium Machining, submit to Journal of Engineering Manufacture, Proceedings of the Institution of Mechanical Engineers, Part B.

Conference papers

[1] Z.G. Wang and Y.S. Wong and M. Rahman, Development of the parallel optimization method based on genetic simulated annealing, In: Maarten Keijzer (ed.), *Late Breaking Papers at the 2004 Genetic and Evolutionary Computation Conference*, June 26-30, 2004, Seattle, Washington, USA, CD-ROM.

[2] Z.G. Wang, M. Rahman and Y.S. Wong, Modeling of cutting forces during machining of Ti6Al4V with different coolant strategies, 8th *CIRP International Workshop on Modeling in Machining Operations*, , Chemnitz, Germany, 2005, pp. 275-282.

[3] Z.G. Wang, M. Rahman and Y.S. Wong, Multi-niche crowding in the development of parallel genetic simulated annealing, *Genetic & Evolutionary Computation Conference, 2005*, Washington DC, USA.