2009

Physical and statistical analysis of functional process variables for process control in semiconductor manufacturing

Xi Zhang
University of South Florida

Physical and Statistical Analysis of Functional Process Variables for Process Control in Semiconductor Manufacturing

by

Xi Zhang

A dissertation submitted in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy
Department of Industrial and Management Systems Engineering
College of Engineering
University of South Florida

Major Professor: Qiang Huang, Ph.D.
José L. Zayas-Castro, Ph.D.
Michael Weng, Ph.D.
Yuncheng You, Ph.D.
Shekhar Bhansali, Ph.D.

Date of Approval:
July 16, 2009

Keywords: interaction patterns, nonlinear dynamic model, interaction structure, statistical quality control, interaction mechanism

Dedication

To my beloved parents

Acknowledgements

I would like to sincerely thank my advisor Prof. Qiang Huang for offering me an opportunity to conduct research and pursue my doctorate degree under his guidance. Prof. Huang's serious attitude on research and in-depth knowledge in applied statistics and engineering greatly inspire me, and his countless encouragement and critical mentoring in my research work present me with confidence and maturity both in research and my life. I would also like thank Prof. Qiang Huang for allowing me involving in NSF research project (NSF CMMI grant # 0600066) and providing me an opportunity of parallel computation training program. Without him, I could not accomplish my dissertation.

I would like to thank my dissertation members Prof. Jose Zayas, Prof. Michael Weng, Prof. Yuncheng You and Prof. Shekhar Bhansali for their valuable suggestions in my research. I would like to thank Prof. Ashok Kumar and his research group for experiment guidance. Additionally, I would express my sincerely appreciation to Dr. Hui Wang for his assistance and enlightening me in my research. I would also like to thank Ms. Gloria Hanshaw, Ms. Jackie Stephens, Ms. Catherine Burton and Mr. Rafael. Moreover, I would like to thank my friends, Ms. Yixin Wang, Mr. Shaoqiang Chen, Mr. Gang Liu, Mr. Yang Tan, Mr. Yu An, Mr. Ozan Ozcan and other workmates.

Finally, I wish to express my pure-hearted love to my beloved parents Jing Luo and Chunsheng Zhang. Without their constantly supports, continual encouragements and endless loves, I would have never finished my dissertation.

Table of Contents

List of Tables

List of Figures

Physical and Statistical Analysis of Functional Process Variables for Process Control in Semiconductor Manufacturing

Xi Zhang

ABSTRACT

The research aims at modeling and analyzing the interactions among functional process variables (FPVs) for process control in semiconductor manufacturing. Interaction is a universal phenomenon and different interaction patterns among system components might characterize the system conditions. To monitor and control the system, process variables are normally collected for observation which could vary with time and present in a functional form. These FPVs interact with each other and contain rich information regarding the process conditions. As an example in one of the semiconductor manufacturing processes, changes of interactions among FPVs like temperature and coefficient of friction (COF) might characterize different process conditions.

This dissertation systematically developed a methodology to study interaction among FPVs through statistical and physical modeling. Three main topics are discussed in this dissertation: (1) Interaction patterns of FPVs under varying process conditions are studied both through experiments and statistical approaches. A method based on functional canonical correlation analysis (FCCA) is employed to extract the interaction patterns between FPVs and experiments of wafer polishing processes are conducted to verify the patterns of FPVs under varying process conditions. (2) Interaction among FPVs is further studied based on physics for process condition diagnosis. A mathematical

model based on nonlinear dynamics is developed to study the strength of interaction and their directionalities, and advanced statistical control charts followed by this nonlinear dynamics model are established for process monitoring. (3) Complex interaction structures among multiple FPVs are analyzed based on nonlinear dynamics for a better understanding of process mechanism. An approach with extended nonlinear dynamics model is proposed to characterize process conditions, and combined engineering knowledge, complex interaction structure patterns are concluded accordingly for interpretation of process mechanism.

The main contribution of this dissertation is to propose a novel methodology based on nonlinear dynamics, which could investigate interactions between components of systems and provide physical understanding of process mechanism for process monitoring and diagnosis. Through studies on interaction among FPVs in semiconductor manufacturing, this research provides guidance for improvement of manufacturing processes. Not limited to manufacturing, the developed methodology can be applied to other areas such as healthcare delivery.

Chapter 1

Introduction

Interaction among components of a complex system is a universal phenomenon and usually can be defined as a kind of action that occurs as two or more components having an effect on one another. Studying interaction mechanism is important because interaction may characterize system. Therefore, to monitor and control the system, interaction might be analyzed through collecting process variables. These process variables usually vary with time and interact with each other, containing rich information regarding the process conditions. As an example in one of the semiconductor manufacturing processes, changes of interactions among process variables like temperature and coefficient of friction (COF) might characterize different process conditions. Hence, to understand the interaction mechanism may bring insights for process improvement.

However, understanding the interaction mechanism is a quite challenging issue. There is still a lack of a science base to develop the interaction model upon which methods of detecting and diagnosing process conditions can be built. The difficulty of establishing such a science base lies in the complexity of the interaction phenomenon. The term of interaction, often vaguely defined, has been used interchangeably with “correlation”, “dependence” and “synchrony” [1], which reflects the different aspects or understanding of the interaction phenomenon. For example in statistics, term interaction

means effects of various changes operate simultaneously, which is different from "interaction" in manufacturing process that dominant factors might trig interaction. It is therefore very challenging to mathematically define the interaction. In addition, FPV signals especially from semiconductor manufacturing have complex spatio-temporal patterns, irregularity and large noise. Moreover, the interaction structure among three or more process variables can be extremely complex. Thus, modeling, detection and diagnosis of the interaction are very challenging.

Therefore, an essential analysis of interaction of FPVs for process improvement is required. The goal of this dissertation is to study the interaction of FPVs and generate knowledge of modeling, detection and diagnosis of the interaction to achieve an insightful understanding of interaction mechanism through combing nonlinear dynamics theory, the engineering knowledge and advanced statistical tools.

1.1 Studies in Semiconductor Manufacturing Processes

With the rapid development of technology in semiconductor manufacturing, more complex manufacturing processes are developed. For example according to "Moore's Law", the demand is the semiconductor industry with respect to the number of transistors per chip will be doubled every 2 year [2], and shrinking of device dimensions will bring complex manufacturing processes which could significantly impact the quality of products. Hence, advanced process control techniques are essentially needed for quality assurance in semiconductor manufacturing. To monitor and control the processes, FPVs are collected for observation. These FPVs usually interaction with each other and contain rich information and could characterize the varying process conditions. Hence, it is of great interest to understand the interaction mechanism for process improvement because

further reduction of surface finish variations can affect 20% of wafer yield and impact a revenue stream of $2.8 billion in a single wafer fab [3].

In semiconductor manufacturing, usually many process factors are involved to affect product quality. One of the important procedures in semiconductor manufacturing processes is chemical-mechanical planarization (CMP) processes. CMP has been widely employed during semiconductor fabrication for planarizing the top surface of semiconductor wafers. In addition to abrasion due to applied mechanical pressure, slurry containing chemicals and particles is continuously fed onto the polishing pad to react with wafer materials on the interface for accelerated and improved planarization performance.

CMP studies roughly follow three categories (1) physics-driven modeling of polishing mechanism and predicting material removal rate (MRR) [4], (2) experimental investigation of process parameters' effects on CMP performance, and (3) data-driven analysis of process variables for condition monitoring and diagnosis. Since MRR is a commonly used criterion to evaluate CMP performance, most research has been focusing on prediction of MRR based on a given set of process variables. Luo and Dornfeld [5] extended Preston's Equation and considered additional process parameters such as wafer hardness, pad roughness, abrasive size and geometry for better prediction. Considering the pad conditioning, Yi [6] first investigated the kinematical relationship between wafer and polishing pad. He also employed the distributed LuGre dynamic friction model to study the wafer/pad friction characteristics. Sorooshian *et al.* [7] studied the effect of pad temperatures, and introduced a new energy parameter into new Preston's Equation by employing an Arrhenius argument. Osseo-Asare [8] and Kaufman *et al.* [9] considered

the chemical reaction between the slurry chemicals and wafer materials. They proposed a Tungsten CMP model by introducing oxidation reduction reaction which occurs in the passivation layer.

In experimental investigation of CMP, process parameters such as slurry characteristics, pad temperature, polishing velocity, COF, and their effects on MRR have been studied. Mudhivarthi *et al.* [10] found that COF decreases with a rise of pressure and platen velocity, and MRR and COF increase when slurry temperature increases. Sorooshina *et al.* [11] investigated the effect of slurry temperature on COF, and they found that COF shows an increasing trend as polishing temperature rises. Seal *et al.* [12] and Du *et al.* [13] concluded that the COF increases with increased peroxide concentration in the slurry and they interpreted this phenomenon as the cause of surface chemical decomposition of polyurethane material. Li *et al.* [14][15] studied the effect of slurry characteristics on friction mechanism and they found slurry with different surfactants and abrasive sizes can significantly alter COF profile. They also found MRR decreases when slurry flow rate increases at a fixed relative rotating velocity.

The least studied area is the data-driven analysis of sensing process variables for online process change detection and diagnosis. Hocheng *et al.* [16] investigated the distribution of the pad temperature and established a regression model to detect the end point. Ganesan *et al.* [17][18] provided their wavelet-based approaches based on sequential probability ration test to identify the delamination and end point online. Wang *et al.* [19] first studied the timing correlation between CMP process variables based on a new phase nonlinear dynamics model and used the model for process change detection.

It is known that factors such as applied force, pad property, and slurry flow rate would jointly (not independently) impact the quality of polished wafers. The observed chemical and mechanical process variables are expected to be strongly correlated during wafer polishing. Our hypothesis is that the correlation among sensing process variables could potentially be utilized to characterize process conditions for the purpose of process monitoring and diagnosis. Nevertheless, process variables observed during manufacturing processes (e.g., the temperature distribution and coefficient of friction (COF) on a wafer), vary with time and present in a functional form. This significantly increases the complexity of analyzing correlation patterns and relating them with process conditions.

Previous research focused on analysis of single process variable for online process change detection and diagnosis. However, process variables like COF or temperature distribution alone cannot uniquely distinguish among process conditions. Jointly considering these correlated process variables would assist to discover the "hidden" interaction mechanisms and a fundamental understanding the interaction mechanisms will assist to improve manufacturing processes.

1.2 Related Work and the State of the Art

The study on interaction of FPVs is very limited. This section reviews the related research on process monitoring, diagnosis and control in manufacturing processes.

1.2.1 Related Works in Studying FPVs for Process Control

Due to the wide application of sensor technology in manufacturing, data processing has been significantly increased. Some sensing data collected through multiple sensors simultaneously from manufacturing processes, often arise themselves in the functional form. Examples include tonnage signals from forging processes [20],

assembly force signals during seat and guide assembly in engine machining processes, and quadrupole mass spectrometry samples of rapid thermal chemical vapor deposition process in semiconductor manufacturing [21]. Comparing with observations of scalar or vector characteristics, functional process variables (FPVs) contain richer process information which might potentially provide additionally opportunities for process improvement and quality assurance. It is known that statistical process control (SPC) has been successfully applied to monitor various manufacturing processes where process performance is measured by a scalar or vector characteristics. However, functional data imposes new dimensions and challenges for real-time process control because standard statistical procedures developed in SPC are not directly applicable for continuous sample curves [22].

Two main strategies have been deployed in process control based on FPVs, depending on how the data is summarized. The first one is to extract "features" from functional data (e.g., wavelet coefficients [23-25], or slope and intercept [26]) and apply standard procedures in multivariate statistics (e.g., T^2 control charts) to features for process monitoring [24][26]. The second strategy is nonparametric regression, i.e., to approximate curves with functions non-parametrically. Data collected under different process conditions can then be discriminated by estimated probability density functions [23] or baseline functions [27]. Both of the two strategies assume that the collected functional data is well-summarized by the extracted features, estimated probability density functions, or baseline. Changes in these summaries indicate process changes.

The functional data studied in the SPC literature, however, is mainly univariate, if we treat a curve as one variable in some functional space. Multivariate curves occur from

multi-sensor systems where location (in a general sense) is an important source of variation. The correlation or interaction among these variables will be overlooked if we only investigate the mean curve or study each individual curves separately. Apparently, certain opportunity of variation reduction could be lost if multivariate curves are not modeled simultaneously in process control. Although it is rarely seen in SPC literature, analysis of nonstationary multivariate functional data has been investigated in statistics. The involved modeling techniques include nonparametric regression [28][29], functional data analysis [30], spatial statistics [31], and principal curve [32]. These methodologies were mainly proposed for exploratory study and are not directly applicable for process monitoring and control. Hence, this is a lack of methodology which could integrate both process control methods and modeling techniques for process improvement.

1.2.2 Related Works in Studying Interactions among FPVs

Several approaches have been reported in the literature to model correlations among nonstationary continuous signals. The most commonly used method is the cross-correlogram which measures the cross-covariance of paired FPVs [33][34]. This time-domain method is powerful, but if not used carefully can lead to spurious detection because of artifactual sharp peaks in the signals. Coherence and cross spectrum methods aim to analyze the correlation of paired signals in the frequency domain and are most commonly used with continuous signals [35]. The correlation analysis based on these methods might be affected both by amplitude fluctuations and by phase variability in signals. Hence, the phase synchronization or phase-locking method has recently received increasing attention by studying the timing correlation in the phase domain while discarding the effect of the amplitude [36]. In CMP processes, phase synchronization

modeling provides an effective tool to describe the timing correlation among critical FPVs such as COF and temperature since they show strong cyclic patterns.

To detect the synchrony between these two signals, the phase-locking method and related test statistics focusing on the rhythm of signal pace should be applied while discarding the effect of amplitude. A m:n phase-locking between a pair of signals is that the relative phase $|m\varphi_1\text{-}n\varphi_2|$ is bounded, i.e., $|m\varphi_1\text{-}n\varphi_2|<\text{const}$ [37][38]. There are mainly three phase-locking statistics used to investigate the timing-correlation introduced in following:

- *Phase Coherence* [39]: The phase coherence index is based on the coherence spectral estimator for bivariate time series. The instantaneous phase difference between 2 signals can be obtained from the unit circle representations. Then the time-dependent phase coherence is then defined as

$$\gamma_N(t_k) = \left\| \frac{1}{N} \sum_{j=k-N}^{k} e^{i\Phi_j} \right\|^2 \tag{1.1}$$

  where the t_k are the sampling points, and N is a parameter indicating the number of consecutive data points to be considered. Phase differences are accumulated over a period of time T (a parameter of the statistic). Phase difference vectors are added and phase coherence is the (squared) complex magnitude of the resultant divided by the number of time points. Note that in this computation the phase coherence value at time t_k takes the sum of the N previous points, the number of points in the analysis window. The phase coherence is always less than or equal to 1, taking a value of 1 only when the relative phase Φj remains constant throughout the observation period T.

- *Entropy Index* [40]: This method also first employs the 1:1 phase-locking idea to compute Φj for each time point $t_{k.}$ A histogram of relative phase Φj for j = k-N, …, k , the observation time window is then built . So the entropy of the series is then defined as

$$h_N = -\sum_{j=1}^{L} p_j \log(1/p_j) \tag{1.2}$$

Where L is the number of bins and p_j is the probability corresponding to the jth bin. For example, we could divide the unit circle uniformly into 12 bins and then to calculate the p_j and establish the histogram. Note that the time of observation is implicit in the fact that N prior phase values are used to generate the histogram from which h_n is computed. In order to conveniently evaluate the timing-correlation in this method, the value should be normalized. For example, the maximum value can be obtained through uniform distribution. $p_j = 1/L$ for all j, and h_N reaches its maximum value h_n max = log$L$. The normalized entropy is then

$$h_n^* = (h_n \text{ max} - h_n)/ h_n \text{ max} \tag{1.3}$$

and the value of h_n^* only between 0 and 1. h_n measures the degree of clustering of the angular distribution, and it is therefore different from the phase coherence in that it will achieve high values for multimodal phase distributions as well as for the unimodal case. For example, the phase coherence of a distribution that has two symmetrical opposing lobes in the circle will be zero, whereas the normalized entropy will yield a value closer to 1.

- *Mutual Information Index* [41]: Similarly to entropy index method, mutual information index is defined as following:

$$I_N = \sum_{i=1}^{L} \sum_{j=1}^{L} p_{i,j} \log \frac{p_{i,j}}{p_i p_j} \quad (1.4)$$

where P_i, P_j is probability of corresponding bin from histogram of the individual phase series. Here P_{ij} is probability from joint histogram of the pairs φ_l^1, φ_l^2. This statistic is computed over the histogram of N accumulated phase values at each point in time. N pairs of angular values φ_l^1, φ_l^2 from the phase trajectories of the individual channels are accumulated during an integration window T on a torus. The joint distribution can be obtained from a two dimensional histogram after dividing the torus in L X L bins.

The mutual information can also be normalized to its maximum value, $I_n = \log L$, achieved when the series φ_l^1, φ_l^2 are identical.

$$I_n^* = (I_n\, max - I_n)/\, I_n \max \quad (1.5)$$

Those approaches mentioned above althrough could avoid effects from amplitude of singals or abrupt variations when analyzing interactions, limitation still exists because only paired correlated signals could be studied. Hence, a novel appoach to analyze complex interaction with mulitple variables is essentially required.

1.2.3 Summary of Literature Review

Previous research has been focused on analysis of the individual FPV or coupled FPVs. However, interaction mechanism could not be simply studied by only inspecting individual FPV or paired FPVs. Therefore, process monitoring and improvement may require a thorough understanding of the interrelationship among those FPVs. Moreover, the interaction structure among three or more FPVs has not been thoroughly investigated for the purpose of process control. There is a lack of physical model to describe the interaction structure of multiple FPVs in manufacturing processes.

1.3 Dissertation Outline

The insightful understanding the interaction of FPVs for process improvement requires investigations in the following aspects: (1) experimentally and theoretically study interaction patterns of FPVs under varying process conditions. A method based on functional canonical correlation analysis (FCCA) is employed to extract the interaction patterns between FPVs; (2) further study of interaction among FPVs based on nonlinear dynamics for process condition diagnosis. A novel nonlinear dynamics model is developed to study the strength of interaction and their directionalities, and advanced statistical control charts are established for process monitoring; (3) analysis of complex interaction structures among multiple FPVs based on nonlinear dynamics for a better understanding of process mechanism. An approach with extended nonlinear dynamics model is proposed to characterize process conditions, and complex interaction structure patterns are concluded accordingly for interpretation of process mechanism.

The overall framework of this dissertation is displayed in Fig 1.1.

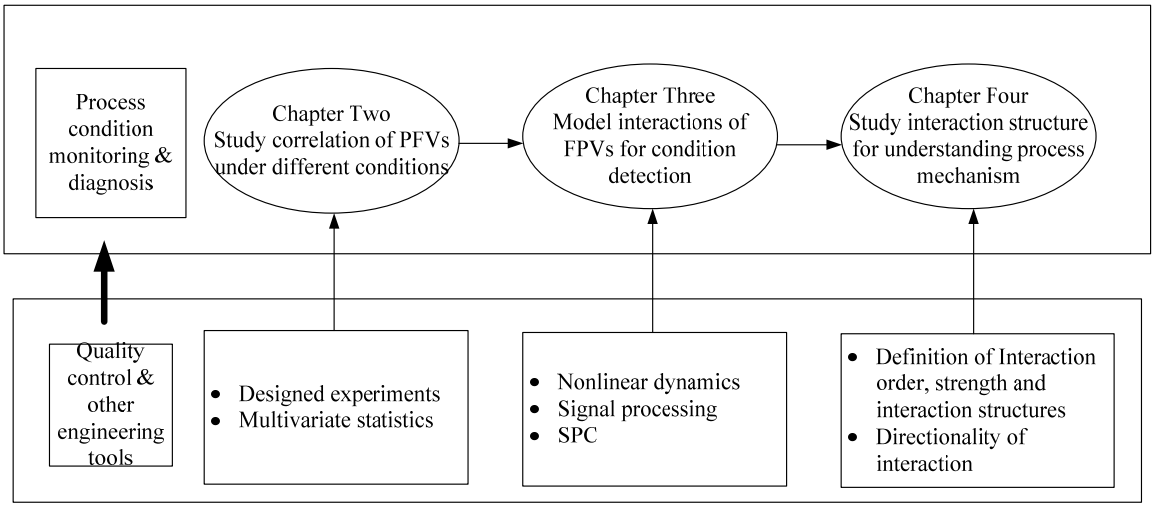


Figure 1.1 The framework of dissertation

Chapter 1 mainly introduces the background of monitoring and diagnosis in semiconductor manufacturing processes and related literature reviews for interaction analysis, process control and CMP studies.

Chapter 2 presents an approach to improve the predication of CMP performance based on the extracted correlation patterns for online process control. The focus of this work is two-fold: (1) experimental investigation of the correlation between process variables and the implication of correlation pattern changes on process conditions, and (2) extraction and statistical analysis of correlation patterns from process variables in functional form.

In Chapter 3, we intend to specifically reveal the timing correlation patterns in CMP. Using nonlinear dynamics, we first established a dynamic phase model to define the strength and patterns of FPV interaction. By monitoring the extracted patterns, we then developed a novel method of detecting CMP condition change and demonstrated the approach via a CMP experiment.

In Chapter 4, we extended our previously developed nonlinear dynamics model by considering the autocorrelation in each FPV to uncover the interaction mechanism of multiple process variables for process condition identification.

Chapter 5 makes the conclusion of this dissertation. Perspectives of future work are also discussed.

Chapter 2

Analysis of Correlated Time-varying Process Variables for Condition Diagnosis in Semiconductor Manufacturing

This chapter mainly focuses on the two aspects: (1) experimental investigation of the correlation between process variables and the implication of correlation pattern changes on process conditions, and (2) extraction and statistical analysis of correlation patterns from process variables in functional form. The ultimate goal is to improve the predication of semiconductor manufacturing performance based on the extracted correlation patterns for online process control. All the experiments and methods are validated through CMP processes.

In this chapter, experimental investigation and statistical modeling of correlation are presented in Section 2.1 and 2.2, respectively, whereby two failure modes (oxidizer and pad failures) are analyzed. A conclusion is given in Section 2.3.

2.1 Experimental Investigation of Slurry Thermal Effects and Correlation among Process Variables

Since the thermal effect of slurry with different percentages of oxidizers and the effect of conditioner-diamond-size induced pad condition change have not been fully explored, we choose slurry and diamond abrasive size of conditioner in our experimental study as experimental factors. Different process conditions will be created to investigate changes in correlation patterns (if there is any). The process variables to be observed

online for correlation analysis include COF of the wafer-pad interface and the temperature distribution over the polishing area.

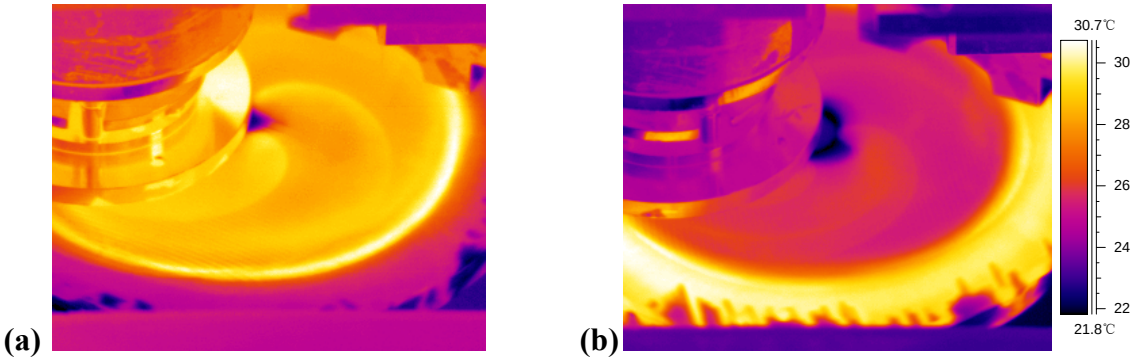


Figure 2.1 The thermal distribution over polishing pad under two polishing conditions

As a critical mechanical variable, the time-varying COF reflects the real-time tribological property at the interface. Changes in COF indicate variations in abrasive performance due to pad failure, large particles on the pads, or underlying barrier layer exposure on the wafer [42]. The thermal distribution over the polishing area is another indicator of process conditions and reflects heat generated through friction and chemical reactions. Figure 1 shows thermal imaging snapshots of two process conditions using different slurries right after 15 seconds of polishing (wafer size 0.7×0.8 inch, slurry temperatures: 30^{o}C). The temperature distribution is relatively uniform under the first process condition (Fig. 2.1a), while there is a bright ring around the polishing zone under the second polishing condition (Fig. 2.1b). Our investigation suggests that the bright ring was due to the heat generated from both chemical reactions and mechanical abrasion, while there was a lack of chemical reaction under the first polishing condition.

2.1.1 CMP Experimental Setup and Design

The experiment was conducted to test the hypothesis that the correlation among time-varying CMP process variables could be utilized to characterize process conditions. To be specific, the experiments aim to investigate (1) thermal effects of slurry with various percentages of oxidizer, and (2) effects of polishing pads generated by conditioners with different diamond abrasive sizes. As shown in Fig. 2.2, the polishing was carried out on a bench-top CMP tester (model CP-4) manufactured by CETR Inc. The real-time shear forces and normal forces at the contact interface were recorded at the frequency of 50Hz. The COF was calculated as the ratio of these two forces. Meanwhile, a FLIR® infrared camera was employed to monitor the temperature distribution in situ on the pad with a sampling frequency of 50 Hz.

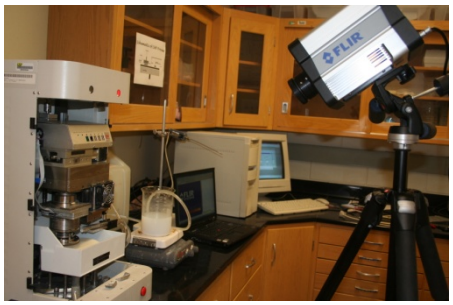


Figure 2.2 Experimental setup

2.1.1.1 Experimental Study of Slurry with Variation of Oxidizer Concentration

In this designed experiment, the 6-inch diameter IC 1000-A4 perforated polishing pad (manufactured by Rodel, Inc.) was attached on the rotating bottom platen of the CMP tester. The 2-inch wafer coupon was attached to the upper polishing head. The upper polishing head and bottom platen rotated in counterclockwise direction in order to let the thermal camera capture the temperature on the polishing zone. The upper slider swung

from side to side to prevent pad from excessive local heat. The slider velocity was maintained at 3 mm/sec during the whole experiment and each run lasted 3 minutes. Cabot copper polishing slurry was mixed with hydrogen peroxide in three designated levels. The slurry was fed into the center of the pad at the rate of 50mL/min. The slurry temperature was maintained at 25^{o}C for each single run by temperature controller (manufactured by Corning, Inc.). The pad was changed in every designated level and was conditioned for two 20-min runs with 1-min polishing of dummy samples in between. The process parameters employed for the experimentation are summarized in Table 2.1.

Table 2.1 Polishing process parameters

Description	**Value**
Wafer coupon	LG Siltron silicon wafer 2 inches
Polishing pad	IC 1000-A4
Slurry	Cabot 5001 Copper slurry
Oxidizer	Hydrogen peroxide
Slider movement	Offset: mm, speed: 3mm/s
Pad conditioning	Pressure: 2psi, Revolution per minute (rpm): 150
Conditioner	Diamond abrasive pad conditioner
Slurry flow rate	50 mL/min
Polishing head	Rpm: 145

LG Siltron silicon wafers were employed during the CMP process to measure the non-uniformity, and the 50μm and 9μm copper lines were chosen. The wafer coupon size was designated 2 inches which contains patterns from four identical dies on the wafer.

Each sample was cut in such a way that the patterns from four dies could rotate on the same circular trajectory during the process, and thereby the patterns could be polished under the same experimental condition (Fig. 2.3).

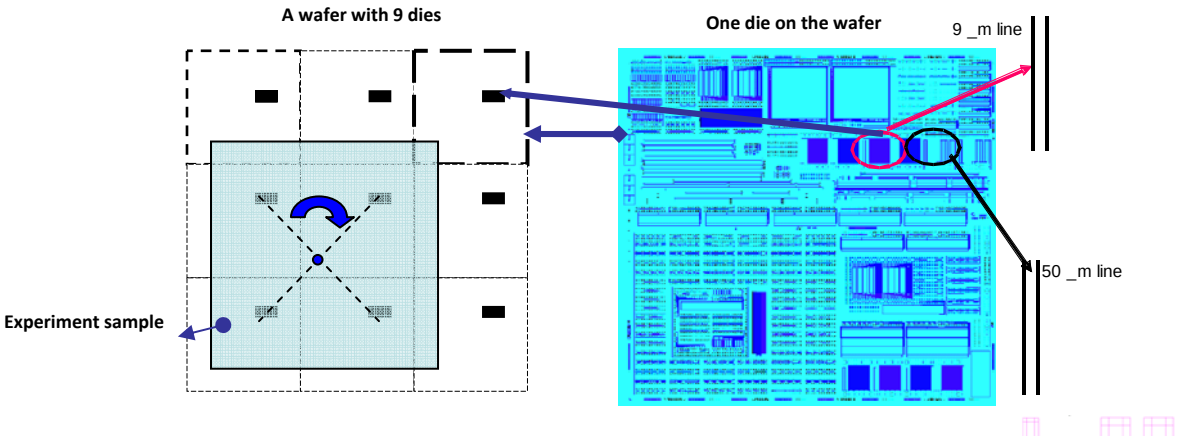


Figure 2.3 Top view of one die on the wafer and sample cutting

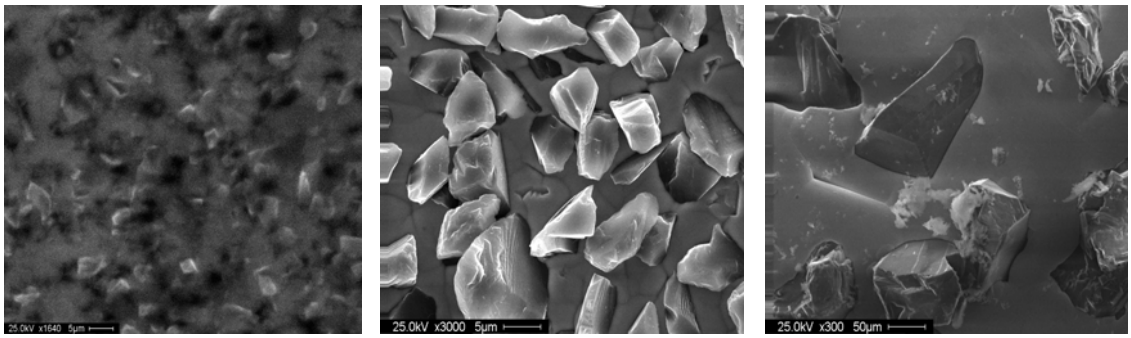
Three different levels (3 samples polished under each level) of slurry condition were designed (see Table 2.2) by adding hydrogen peroxide with different concentration to the Cobot 5001 slurry. The purpose of the design is to simulate the case when oxidizer fails during the polishing process and investigate how the oxidizer levels affect correlation patterns, which are to be discussed in Section III.

Table 2.2 Experimental design

Rotational velocity (Polishing head vs. pad)	Pressure	Polishing time	Slurry Solution : H_2O_2
150 vs. 145 rpm	2 psi	2 minutes	900:150
			900:75
			900:0

2.1.1.2 Experimental Study of Polishing Pad with Varying Diamond Sizes in Conditioner

To study CMP pad failure, the same process parameters as 2.1.1.1 were applied except slurry and wafer samples. Here the Cobot 5001 copper slurry with 2.5% hydrogen peroxide and single-layer copper wafer were used. Three conditioners with different diamond abrasive sizes (0.25μm, 8 μm and 100 μm, Fig. 2.4) were used to generate different morphologies and roughness on polishing pads. In addition, different abrasive size can cause distinct failure patterns on the pad (e.g., scratches). Under each level of conditioners, experiments were replicated three times.



(a) 0.25μm (b) 8μm (c) 100μm

Figure 2.4 Microscopic view of conditioner with different abrasive size

2.1.2 Experimental Results and Post-CMP Studies

The image in Fig. 2.5 shows a snapshot of thermal distribution over the entire pad, while the data at the bottom panel is the time-varying temperature of a selected zone on the pad. The focused thermal zone was selected on the polishing pad adjacent to the wafer-pad interface to record the average temperature. Since the temperature in this interface between wafer and polishing pad cannot be obtained directly, the zone we selected in this way might best indicate the average process temperature 0. The thermal

data were recorded on this selected zone during the whole process, and the data could be observed instantaneously.

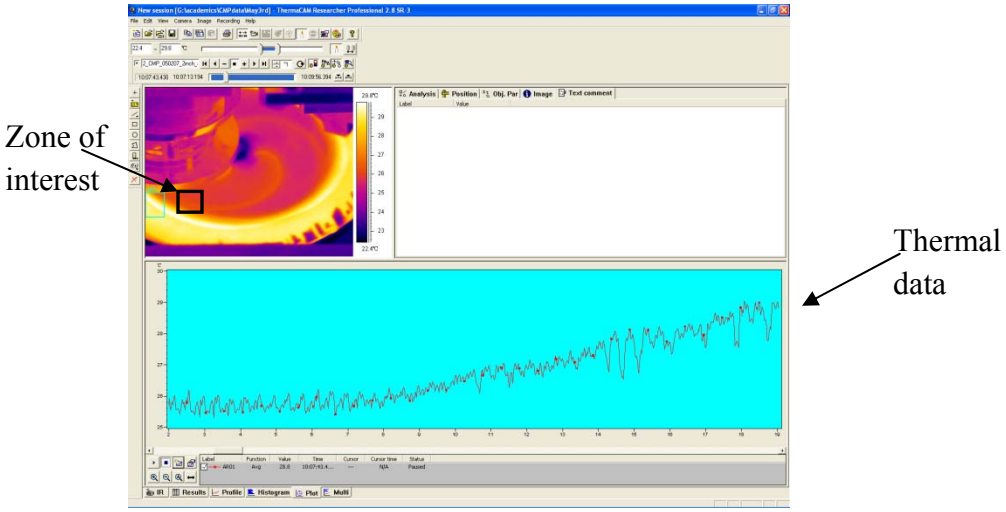


Figure 2.5 An example of recording temperature

In the first designed experiments (Section 2.1.1.1), we collected thermal data and COF during the whole polishing process. For example, Figure 2.6 displays the COF and temperature vs. time under three conditions. Apparently, both signals have the similar patterns in their general trend, i.e., they increase at the beginning of the cycle, decrease after the signals reach the peak values, and eventually asymptotically approach a constant. The decreasing trend in the signals is caused by various factors such as improved lubrication of polishing after the process stabilize or improved interface between the wafer-pad friction pair after certain amount of copper has been removed.

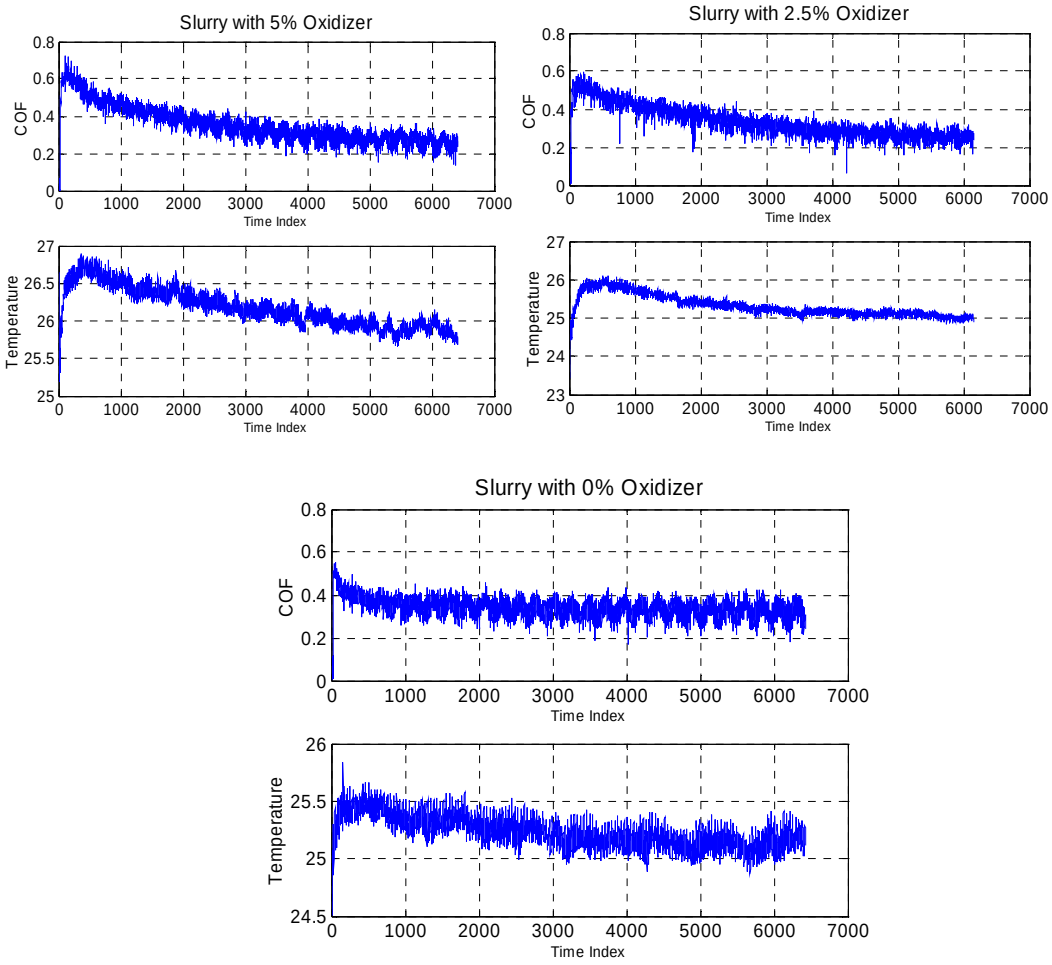


Figure 2.6 Observed functional process variables under 3 conditions

The post CMP surface characterization was carried out on an Atomic Force Microscope (AFM) to study the surface non-uniformity, which is measured in terms of the vertical distance between the different locations on the samples, i.e., the height of "steps" on the surface profile. Since the samples (LG Siltron silicon wafers) we adopted have patterns under the copper layer, we need to evaluate and compare the non-uniformity on the same patterns. In this study, we chose two isolated lines (50μm and 9μm) on each die of the wafer (Fig. 2.7).

The AFM cantilever and tip scanned an area of 30μm×30 μm covering each line to obtain the profile image, on which we measured the "step" heights on sections at three locations. Figure 2.8 shows the examples of measurement and AFM readings for three levels of oxidizer. Table 2.3 gives the statistical summaries of within-sample uniformities (the average values vertical distances μ_h and standard deviation σ_h over 9μm or 50μm lines on four dies).

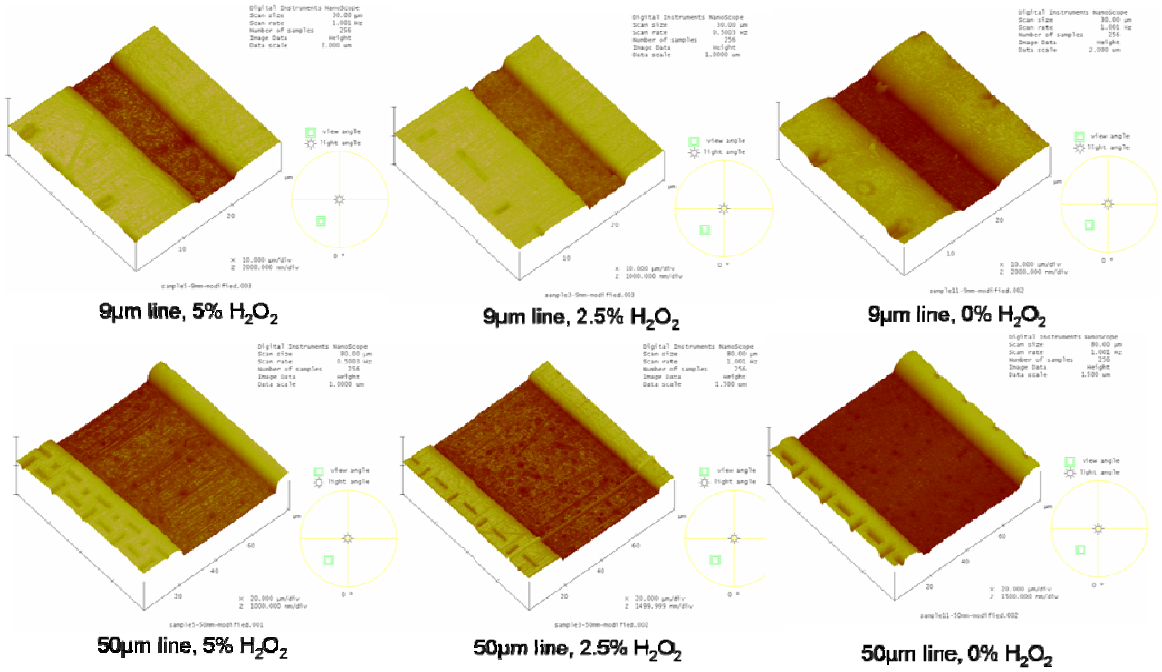


Figure 2.7 AFM 3-D images for three levels of oxidizers

Table 2.3 shows that the second level has the smallest vertical distance and therefore has the lowest non-uniformity. This implies the polishing was more efficient using the slurry with 2.5% H_2O_2 during 2 minutes of polishing. However, under this level, the variation of step height over the same patterns on all samples is the largest, which implies the chemical reaction speed is different over four dies on samples. It is not

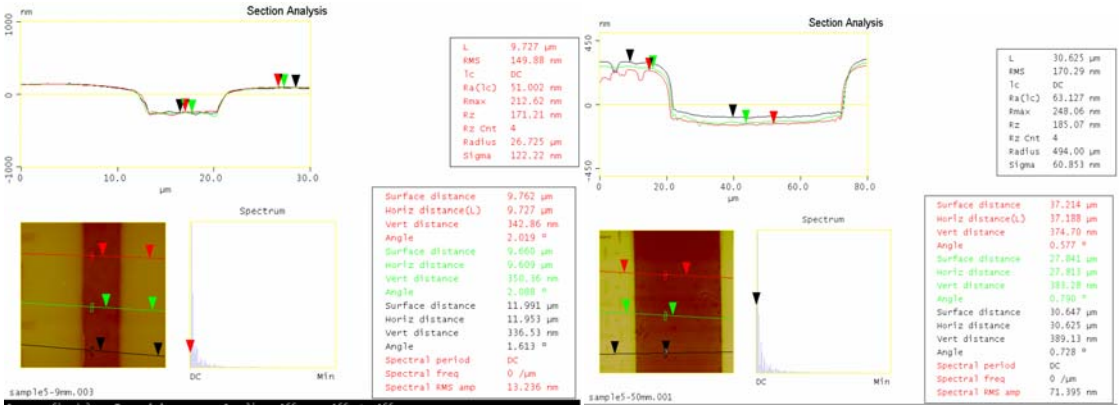
surprising to observe that the third level has smaller variation in "step" height since the most copper remains intact due to a lack of chemical reaction.

Table 2.3 Statistical summaries of within-sample uniformity from AFM (Unit: nm)

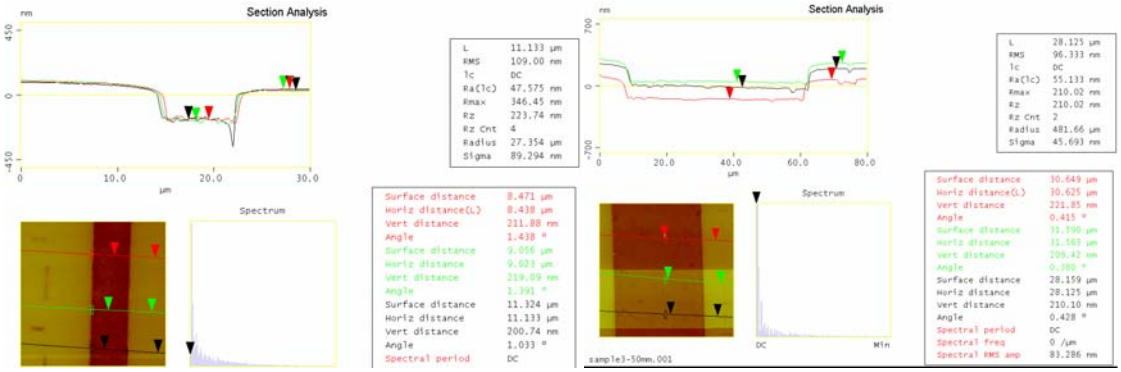
	9 μm line		50 μm line	
Levels	μ_h	σ_h	μ_h	σ_h
5% H_2O_2	319.3933	44.29538	390.7233	25.23637
2.5% H_2O_2	186.5692	51.28105	231.0175	36.73086
0% H_2O_2	573.1422	19.80072	472.2575	28.3114

There is noticeable difference observed between the measurements on 9 μm lines and 50 μm lines. Surface over the 9 μm line has lower non-uniformity than that over the 50 μm line whereas the variation over the 50 μm line is smaller. This fact implies that polishing on thin lines is more efficient than polishing on thick lines in our experiment.

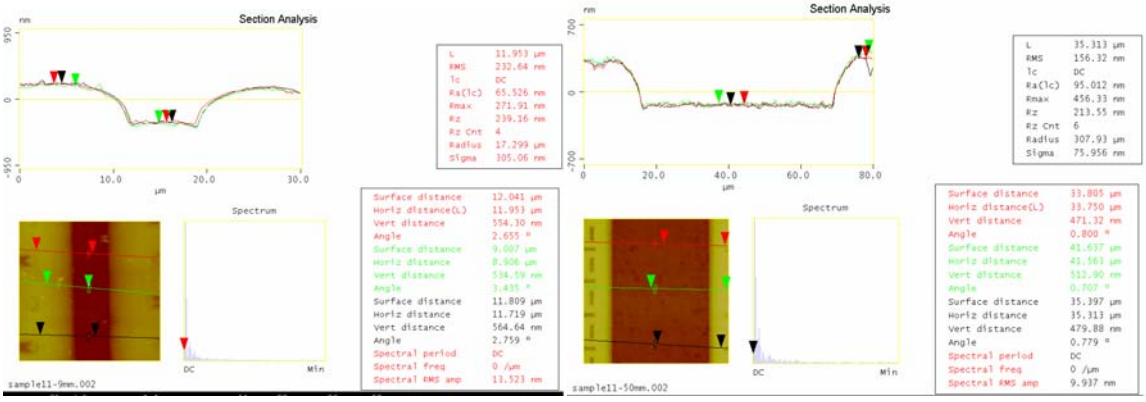
In the second designed experiment (Section 2.1.1.2), the data from both conditioning and polishing process were recorded. Both COF and temperature have similar patterns and trends as in our first designed experiment. The polished wafer surfaces were examined through Leitz Ergolux Optical Microscope. Figure 2.9 shows the optical images of wafer surfaces polished on pads conditioned with different abrasive sizes. It is concluded that the pads conditioned with larger abrasives resulted in more number of scratches compared to the smaller abrasive size.



(a) 9μm line, 5% H_2O_2 (b) 50μm line, 5% H_2O_2

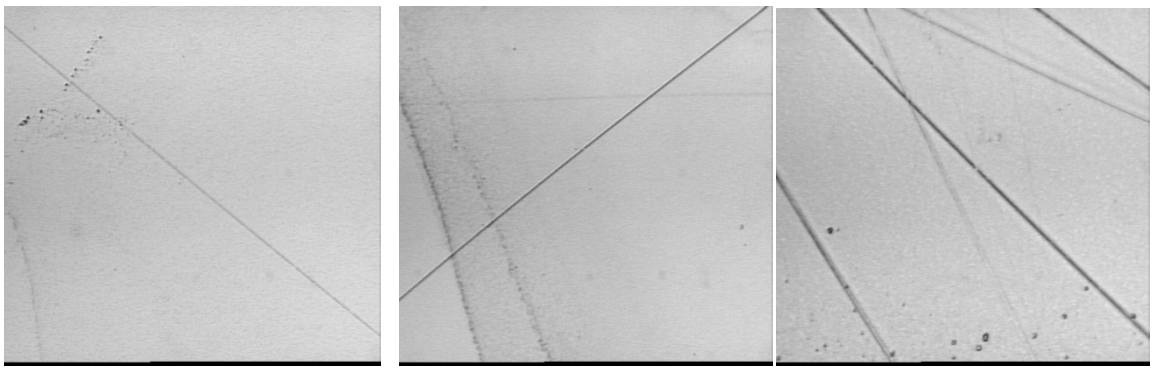


(c) 9μm line, 2.5% H_2O_2 (d) 50μm line, 2.5% H_2O_2



(e) 9μm line, 0% H_2O_2 (f) 50μm line, 0% H_2O_2

Figure 2.8 AFM measurements of wafers under three levels of oxidizers



(a) 0.25μm (b) 8μm (c) 100μm

Figure 2.9 Optical images of wafers polished on pads conditioned with different abrasive sizes

We also collected thermal data and COF during the whole process, and we ran three replicates in each condition. Fig. 2.10 showed one example of COF and temperature in each condition. It can be seen that the signal patterns under each condition are similar and difficult to distinguish with each other.

From the first experiment (Section 2.1.1.1), the results have shown that the levels of oxidizer play a significant role affecting the polishing quality and efficiency. Process with slurry using 2.5% H_2O_2 gives more efficient polishing than 5% H_2O_2. This can be explained by the nonlinear relationship between the removal rate and H_2O_2 concentration, i.e., adding H_2O_2 will significantly increase the polishing rate whereas further increases in H_2O_2 will lower the polishing rate 0. The decrease in polishing rate may be caused by more copper oxides generated by high level of oxidizer, which can reduce or prevent copper layer from further chemical reaction. The second experiment has shown that surface defects (scratches) on polished wafers could be attributed to the pad failure (rough pad) generated by a larger abrasive size of conditioners.

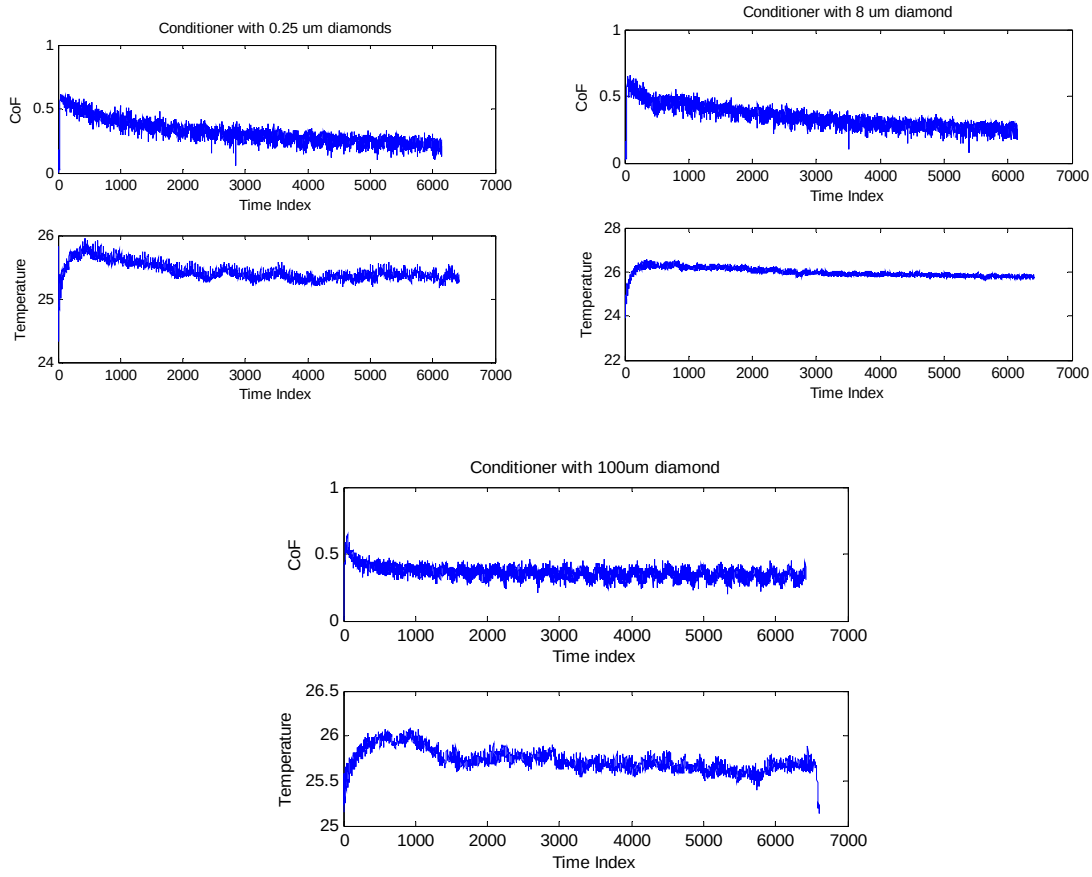


Figure 2.10 Observed functional process variables under 3 different conditioners

It is essential to detect the process change caused by large variations in oxidizer level or pad failure. However, inspecting critical process variables such as COF and pad temperature separately is incapable of distinguishing different conditions. Nevertheless, we could observe that the oscillatory patterns in COF and temperature bear certain similarities under each condition. This fact may lead to a new method to characterize and detect the process condition by uncovering the latent correlation patterns between the critical process variables.

2.2 Statistical Analysis of Correlated Process Variables for Condition Diagnosis

Results from Section 2.1 show that inspection of COF or temperature distribution alone cannot uniquely distinguish all process conditions. In this Section, we will jointly consider these correlated process variables by analyzing their correlation. The objective is to discover the correlation patterns under different process conditions and thereby to detect the condition changes. The modeling approach is based on functional canonical correlation analysis (FCCA), which is to assess the correlation through measuring the statistical similarities among functional data. The statistical correlation patterns under different polishing conditions can be captured.

2.2.1 Process Condition Characterization through FCCA

In applied statistics, canonical correlation analysis (CCA) maximizes the correlation between a linear combination of two random vectors. FCCA is motivated to find functional canonical variates that maximize the covariance function of two process variables (or stochastic processes). To be specific, suppose $x(t)$ and $y(t)$ are two process variables with zero mean and covariance functions $\mathrm{E}[x(t)x(s)]$, $\mathrm{E}[y(t)y(s)]$, and $\mathrm{E}[x(t)y(s)]$. Also assume there are N observed pairs of data curves ($x_i(t)$ and $y_i(t)$) for the two variables. The variance and covariance can then be estimated using these N observations, e.g., $\mathrm{E}[x(t)y(s)]=N^{-1}\sum_i x_i(t)y_i(s)$. The FCCA is to find weight functions $u(t)$ and $v(t)$ to maximize the squared correlation (denoted as $\mathrm{ccorsq}(u,v)$) of $\int u(t)x_i(t)\mathrm{d}t$ and $\int v(t)y_i(t)\mathrm{d}t$, i.e. [0],

$$\max_{u,v} \mathrm{ccorsq}(u,v)=\iint u(t)\mathrm{E}[x(t)y(s)]v(s)\,ds\,dt \tag{2.1}$$

$$s.t.\ \iint u(t)\mathrm{E}[x(t)x(s)]u(s)\,ds\,dt+\lambda\|D^2u(t)\|^2=1 \text{ and}$$

$$\iint v(t)\mathrm{E}[y(t)y(s)]v(s)\,ds\,dt+\lambda\|D^2v(t)\|^2=1,$$

where λ is a positive smoothing parameter, D is a differential operator d/dt, $||.||^2$ is an operator that computes $\int(.)^2 dt$, and $\lambda||D^2(.)||^2$ gives a penalty term that considers the roughness of the functions $u(t)$ and $v(t)$. Appropriately selected smoothing parameter λ can yield fairly smooth weight functions and a correlation that is not unreasonably low. It can be chosen subjectively or according to cross-validation procedures.

Functional curves $u(t)$ and $v(t)$ can be estimated by seeking expansions in terms of a fixed number of basis functions $\Phi_i(t)$, e.g., Fourier basis for periodical signals and B-spline basis for non-periodical signals. Then functional curves are estimated as $\sum_i b_{u,i}\Phi_i(t)$ for $u(t)$ and $\sum_i b_{v,i}\Phi_i(t)$ for $v(t)$, where $b_{u,i}$ and $b_{v,i}$ are coefficients.

It should be noted that the process variables need to be detrended and zero-centered prior to implementing FCCA. The reason is that both COF and temperature have similar general trends and therefore their trends definitely are highly correlated and possess similar correlation patterns. For the purpose of monitoring process condition using correlation analysis, we should look into the details of variations by removing the effect of general trend. Therefore, zero-centering the data is a requirement of FCCA. In this study, cubic spline smoothing is employed to separate the data trend from the process variables. Figure 2.11 gives an example of the detrending results for one replicate of polished samples (5% H_2O_2).

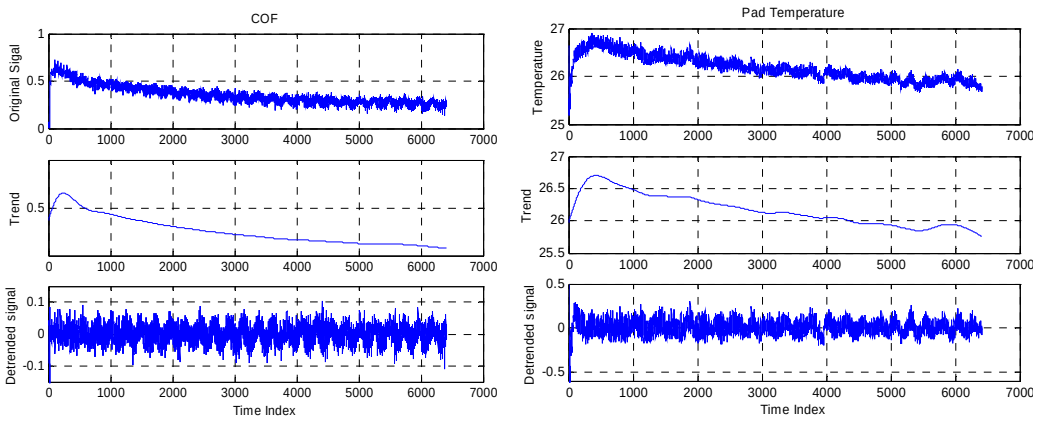


Figure 2.11 An example of detrending (5% H_2O_2)

Solutions to Eq. (2.1) (denoted as u^{I} and v^{I}) give a pair of *leading* canonical variates that maximize ccorsq(u,v). Thus, $u^{I}(t)$ and $v^{I}(t)$ reflect a latent pattern of correlation. Other latent patterns can be identified by the second and higher order canonical variates, denoted as (u^{II}, v^{II}), (u^{III}, v^{III}), and so on. For example, the second pair (u^{II}, v^{II}) is the functions u and v that maximize the same correlation ccorsq(u,v) subject to the constraint that they are to be uncorrelated with the leading pair of canonical variates. The number of potential canonical variates is equal to the number of basis functions in fitting the functional curves.

2.2.2 Results and Analysis

As pointed out in Section 2.2.1, the weight functions $u(t)$ and $v(t)$ determine the correlation patterns, which are able to characterize the process conditions. In 2.1.1.1, we study the correlation patterns under three levels of oxidizer (H_2O_2) by plotting (see Fig. 2.12) the weight functions for the first two pairs of canonical variates, where solid curve gives the weight function of COF and dashed line shows the weight function of temperature. Since the detrended signals oscillate, Fourier basis is chosen to fit the data to

functional curve. Table 2.4 shows the values of corresponding correlations for the first two pairs of canonical variates. Leading pair of canonical variates has very large correlation values and reflects major correlation patterns.

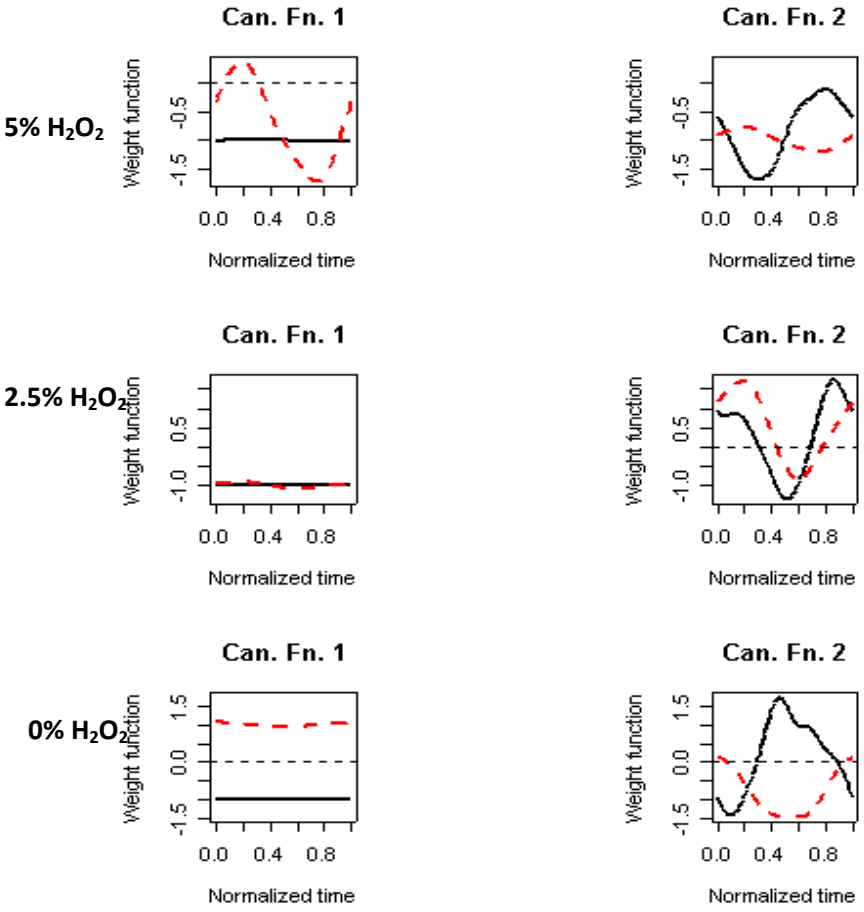


Figure 2.12 Canonical correlation patterns under 3 levels of H_2O_2

Comparing the weight functions for slurries with different oxidizer percentage, we observe that the change in oxidizer will significantly alter the shape of weight function. To interpret such pattern changes, recall that the signals have been detrended and the functional data show a characteristic of periodical oscillation. Corresponding to the slurry with 2.5% H_2O_2 the solid curve and dashed curve of the first pair of canonical

variates have almost constant weight functions. This indicates the original process variables themselves would achieve the maximum correlation (not true in the other two conditions). Since the process with 2.5% H_2O_2 yields the best polishing results, the associated correlation pattern can serve as a baseline pattern for future process diagnosis. A large fluctuation in weight function of temperature is observed under the first level (5% H_2O_2) while the two curves significantly deviate from each other under the third level (0 .0% H_2O_2).

Table 2.4 Canonical correlation of slurry study

Canonical variates	1st	2nd
Slurry w. 5% H_2O_2	0.928	0.259
Slurry w. 2.5%H_2O_2	0.989	0.283
Slurry w/o. H_2O_2	0.996	0.128

The second pair of canonical variates has a relatively small correlation; however, it still provides a reference to distinguish between different levels, especially for the levels of 2.5% H_2O_2 and 0% H_2O_2 since the correlation patterns are very different.

We also analyzed the correlation patterns under three levels of abrasive size of conditioners and the weight functions of canonical variates are given in Fig. 2.13. Table 2.5 gives the correlation for the first two pairs of canonical variates. Prior to the FCCA, the trend is removed and the Fourier basis was chosen in the same way as the case in 2.2.1.1. From Fig. 2.13, we again observed different weighting function patterns under different conditions. When the abrasive size 0.25μm is chosen, the weighting function shapes of both COF and temperature appear almost the same, and this indicated original

two process variables achieved maximum correlation (not true in other two conditions). When 100μm abrasive size is used, in addition to fluctuation in dashed curve, two curves significantly depart from each other.

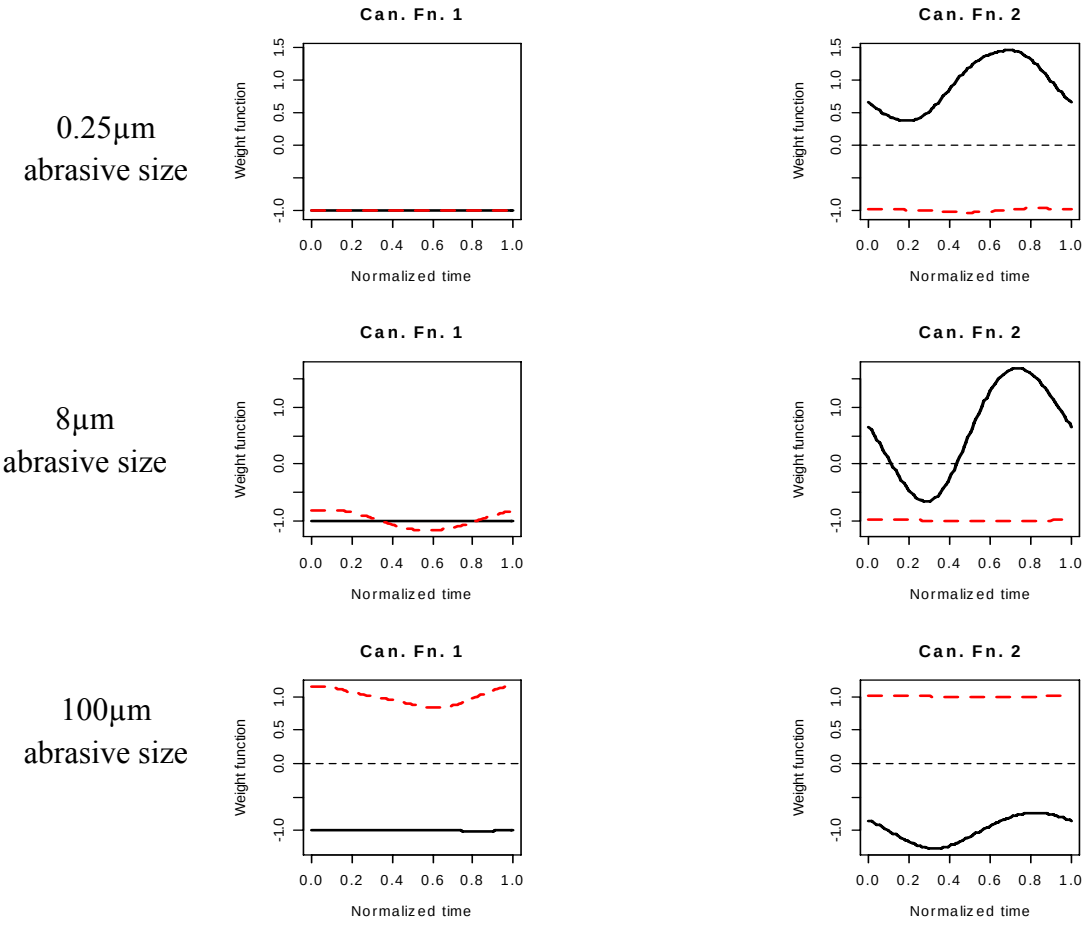


Figure 2.13 Canonical correlation patterns under 3 levels of abrasive sizes

From results of both cases, we may observe that the best process condition seems to be related to the correlation pattern that original process variables achieve the maximum correlation. Further experimental studies are needed to confirm this observation.

Table 2.5 Canonical correlation of conditioner study

Canonical variates	1st	2nd
0.25μm abrasive size	0.992	0.142
8μm abrasive size	0.985	0.466
100μm abrasive size	0.840	0.253

2.3 Summary

This chapter conducted an experimental and statistical analysis of correlation among time-varying process variables for CMP process condition monitoring and process change detection. We proposed that correlation patterns can help to characterize process conditions. Therefore, two experiments were designed to study the correlation between process variables. In the first experiment, we polished silicon wafer using three levels of concentration of oxidizer (H_2O_2) in slurry to simulate oxidizer failure and study its impact on polishing quality and COF-temperature correlation pattern. COF was recorded by the embedded pressure sensors in a CMP machine and temperature on polishing pad (including spatial and temporal variations) was captured by a thermal camera. Post-CMP analysis conducted on an AFM has shown that the level of oxidizer has a huge impact on polishing quality and efficiency. Slurry with 2.5% H_2O_2 yielded the lower non-uniformity than other levels. In a similar fashion, the second experiment investigated pad failures and their impact on COF-temperature correlation. Three types of abrasive sizes of conditioners were used to condition the pad that could cause scratches on the polished wafer. Experimental results show that it is difficult to distinguish among process conditions by investigating sensing variables COF or temperature alone.

The statistical analysis intends to explore the implication of correlation analysis on process condition change detection and monitoring. The functional data (time-varying function) were first expanded under certain functional basis (e.g., Fourier basis). Based on the expansions, the FCCA then measures statistical similarities among time-varying process variables and identified the weight functions that maximize the correlation among variables. The weight functions capture the correlation patterns corresponding to different process conditions. FCCA analysis showed that the slurry problem and pad failure can significantly change the shape of COF-temperature canonical variates. The correlation between the leading pair of canonical variates reveals the major correlation patterns. From both experiments, we found that original COF and temperature signals are more likely to achieve maximum correlation under the baseline conditions and less likely under faulty conditions. This fact leads to a new diagnostic method for abnormal process change caused by certain latent factors (e.g., slurry contamination or pad failure) that cannot be easily detected.

Future research in next chapters involves interaction patterns among multiple process variables and modeling improvement for a better interpretation of correlation patterns.

Chapter 3

Nonlinear Dynamics Modeling of Correlated FPVs for Condition Monitoring in Semiconductor Manufacturing Processes

This chapter aims to investigate interaction between functional process variables (FPVs) for condition monitoring in chemical-mechanical planarization (CMP). During wafer polishing, critical process variables such as coefficient of friction (COF) and pad temperature vary with time and present in the shape of functional curves. In previous chapter, we have demonstrated that correlation patterns among these FPVs could be related to polishing conditions. Since correlation is affected by both amplitude fluctuations and phase variability in FPVs, further study of timing correlation of FPVs measured in different units could bring more insights into physical interactions and thereby enhance CMP condition monitoring. Existing research on FPVs in CMP mainly focuses on individual effects of FPVs and statistical correlations through experimental and theoretical analyses. In this paper, we intend to specifically reveal the timing correlation patterns in CMP. Using nonlinear dynamics, we first established a dynamic phase model to define the strength and patterns of FPV interaction. By monitoring the extracted patterns, we then developed a novel method of detecting CMP condition change and demonstrated the approach via a CMP experiment. The results showed that the proposed method has a promising application in identifying the process changes that may not be easily detected otherwise.

This chapter is organized as follows. Section 3.1 develops methods of nonlinear dynamics modeling of physical interaction. In Section 3.2, a model based process condition change detection/diagnosis method is then proposed by employing statistical process control tools. Section 3.3 applies the methods to a CMP experimental data and discusses the results. Conclusions are given in Section 3.4.

3.1 Nonlinear Dynamics Modeling of FPV Timing Correlation

In CMP processes, phase synchronization modeling provides an effective tool to describe the timing correlation among critical FPVs such as COF and temperature since they show strong cyclic patterns (see Fig. 3.1 as an example). The remaining of this section will establish a model to identify the main effect and interaction effect between COF and temperature. The extracted interaction pattern can facilitate further monitoring and diagnosis of pattern change in timing correlation in Section 3.2.

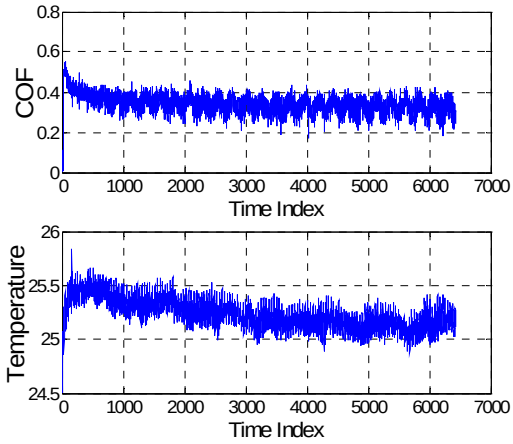


Figure 3.1 Strong cyclic patterns in process variables

Nonlinear dynamics theory suggests that the synchronization among p oscillatory signals can be modeled by [46]

$$d\varphi_k(t)/dt = \omega_k + Q_k[\varphi_1(t), \varphi_2(t), ..., \varphi_p(t)] + \varepsilon_k, \quad k = 1, 2, ..., p, \qquad (3.1)$$

where $\varphi_k(t)$ is the phase variable (a function of time t), ω_k is the base angular frequency (frequency component with the highest power), and ε_k is the white noise of the kth oscillatory signals $x_k(t)$.

The phase variables $\varphi_k(t)$'s can be obtained by constructing an analytical signal obtained from Hilbert transform as follows (For a simple sine signal $x(t)= A \sin(\omega_0 t +\varphi_0)$, the phase $\varphi(t)$ is defined as $\varphi(t) =\omega_0 t +\varphi_0)$,

$$\varphi_k(t) = \arg\{x_k(t) + i[1/\pi \int_{-\infty}^{\infty} x_k(\tau)/(t-\tau) d\tau]\}, \quad i = \sqrt{-1}. \tag{3.2}$$

The term $Q_k(.)$ is defined as a function describing interaction among these phase variable $\varphi_k(t)$ and is approximately periodic. Therefore, $Q_k(.)$ can be approximated by Fourier expansion:

$$Q_k(\varphi_1, \varphi_2, ..., \varphi_p) = \sum_{m_1, m_2, ..., m_p} [a_j^{m_1, m_2, ..., m_p} \cos(\sum_{j=1}^{p} m_j \varphi_j) + b_k^{m_1, m_2, ..., m_p} \sin(\sum_{j=1}^{p} m_j \varphi_j)], \; k=1,2,\ldots p, \tag{3.3}$$

where the superscripts m_j's in the Fourier expansion are integers and cannot be zeros simultaneously. The values of coefficients $a_j^{m_1, m_2, ..., m_p}$ and $b_k^{m_1, m_2, ..., m_p}$ can be estimated through the Ordinary Least Square (OLS) regression method. Prior to fitting the model, any signal trend must be removed from the data and the signal should fluctuate around zero (see an example in Fig 3.2). The trends in the middle graphs were extracted using cubic spline smoothing.

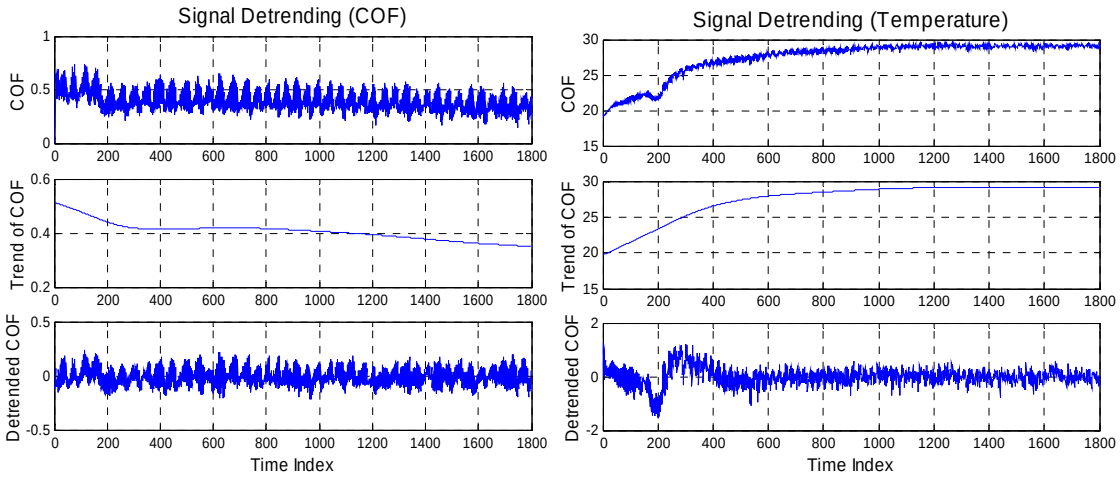


Figure 3.2 An example of signal detrending

The major challenge is to find interaction terms in the model with adequate orders. The model to be developed will not only be statistically adequate to avoid overfitting for robust prediction, but also physically interpretable for a better understanding of the synchronization mechanism. Grounded on these principles, a small Fourier expansion is preferred for a better physical interpretation. Initially we can start with a small series and gradually increase the order of Fourier series through a model adequacy check. The Fourier expansion should be controlled in such a way that the base frequencies roughly go through the middle of the instantaneous frequencies. A model selection procedure, e.g., backward selection method, is adopted to screen the insignificant coefficients in the Fourier series. The statistical significance of model coefficients is determined by a small *p*-value, e.g., 0.01 or less. To obtain a model with better interpretability, the interaction order and strength are defined as follows.

Following the concept of statistical effects in the design of experiments [47], the main effect and interaction effect can be defined in a similar way. For example, the main

effects include $cos[\varphi_1(t)]$, $cos[\varphi_2(t)]$, $cos[\varphi_3(t)]$, $sin[\varphi_1(t)]$, $sin[\varphi_2(t)]$, and $sin[\varphi_3(t)]$. The two-way interaction effects or *the first order interactions* include $cos[\varphi_1(t)+\varphi_2(t)]$, $cos[2\varphi_1(t)+\varphi_3(t)]$, etc. The three-way interaction effects or the second order interactions contain $cos[\varphi_1(t)+\varphi_2(t)+\varphi_3(t)]$, etc. With this definition, basic principles like that of the hierarchical ordering principle are readily applied to the model selection procedure. For instance, the effect heredity principle suggests that in order for the interaction effect $cos[\varphi_1(t)+\varphi_2(t)]$ to be significant, normally at least one of its parent effects should be significant. Its parent effects are the main effects in the trigonometric identity $cos[\varphi_1(t)+\varphi_2(t)] = cos[\varphi_1(t)]cos[\varphi_2(t)]-sin[\varphi_1(t)]sin[\varphi_2(t)]$.

Due to the nature of Fourier expansion, the cosine and sine pairs can be represented in a complex form, e.g., $\sqrt{(a_k^{m_1,m_2,m_3})^2 + (b_k^{m_1,m_2,m_3})^2}$ $\exp\{im_1\varphi_1(t)+im_2\varphi_2(t)+im_3\varphi_3(t)\}$. The term $\sqrt{(a_k^{s,r,w})^2 + (b_k^{s,r,w})^2}$ represents the amplitude in signal processing and $[(a_k^{m_1,m_2,m_3})^2+(b_k^{m_1,m_2,m_3})^2]$ is related to the power of that frequency component. The *strength of each frequency component* in a main effect or an interaction effect can be thus defined using the concept of the power, e.g., $[(a_k^{m_1,m_2,m_3})^2+(b_k^{m_1,m_2,m_3})^2]$. The *strength of the main effects/interaction effects* is defined as the summation of the power of every frequency component in all the main/interaction effects, e.g., $\sum_{m_1,m_2,m_3} [(a_k^{m_1,m_2,m_3})^2+(b_k^{m_1,m_2,m_3})^2]$. The proposed definition provides an opportunity to identify and analyze the important frequency components in each order of interactions.

3.2 Statistical Process Monitoring Based on Nonlinear Dynamics Modeling of FPV Timing Correlation

Statistical process control (SPC) is an effective tool to monitor process changes and reduce process variations. But standard SPC methods could not be directly applied to processes with FPVs [48]. Two main strategies have been deployed for process monitoring using functional data. The first strategy is to extract features from curves, e.g., peak values, wavelet coefficients [49-54], or slope and intercept [53][54]. Then standard procedures developed in multivariate process control (e.g., Hotelling's T^2 control chart) can be applied to monitor those features. The second strategy is nonparametric regression, i.e., to approximate curves with functions. Curves collected under different process conditions can then be discriminated into categories through baseline functions [55][56].

Nevertheless, these approaches mainly focused on one single FPV except the modeling methods reviewed in Section II and a semi-parametric method based on principal curve analysis [58]. In this section, we propose a new statistical method to detect change of timing correlation among FPVs for CMP processes.

Prior to the detection procedures, data for FPVs should be collected in the following manner: sample data under condition 1 (normal condition), and sample data under condition 2 (abnormal condition). The data collected under the normal condition will be used as training data to establish the phase dynamics model proposed in Eq. (3.1). The changes of synchronization pattern include *systematic change* or base frequency ω_k change and *interaction change* or $Q_k(\cdot)$ change. Systematic change in signal base frequency implies that significant process condition changes occur, which can be directly detected by visual inspection on the signal oscillatory pattern. Interaction change is

related to moderate process condition change, which might not be identified by inspecting original signals.

Since the model coefficients ω_k, $a_k^{s,r,w}$'s, and $b_k^{s,r,w}$'s, define patterns that characterize process conditions, they will be used to detect interaction changes. Suppose p synchronized signals are modeled by Eq. (3.1). Denote $\boldsymbol{\omega} = [\omega_1\ \omega_2,\ldots,\ \omega_p]^{\mathrm{T}}$ and $[\boldsymbol{a}_k\ \boldsymbol{b}_k]^{\mathrm{T}}$ as a stackup of coefficients in the interaction function $Q_k(\cdot)$. Given the data collected from the normal process condition, multivariate control charts (e.g., Hotelling's T^2 chart) can be built for $\boldsymbol{\omega}$, and $[\boldsymbol{a}_k\ \boldsymbol{b}_k]^{\mathrm{T}}$ to monitor systematic pattern change and interaction pattern change, respectively.

Due to the irregularity, noise, and complex spatio-temporal patterns in real-time signals, the phase dynamics model may consist of a large set of coefficients $[\boldsymbol{a}_k\ \boldsymbol{b}_k]^T$. The large model dimension will adversely affect the performance of detection procedures. Principal Component Analysis (PCA) is an effective way of dealing with highly correlated parameter estimates. It is implemented along with the development of statistical detection procedures to reduce the model dimension for effective change detection. The basic idea is to monitor the first few principal components of $[\boldsymbol{a}_k\ \boldsymbol{b}_k]^T$ instead of the coefficients themselves. For instance, Hotelling's T^2 statistic in terms of principal components can be $T^2 = \sum_{a=1}^{A} t_a^2 / s_a^2$, where t_a is the ath principal component and s_a^2 is its corresponding sample variance. The number of principal components A can be determined by the amount of total sample variance explained. Thus, the phase I control chart limit to monitor the principal components is UCL=$(m\text{-}1)^2/m\beta_{\alpha,\ A,\ (m\text{-}A\text{-}1)/2}$, LCL=0, where m is the number of samples and $\beta_{\alpha,\ A,\ (m\text{-}A\text{-}1)/2}$ is the upper α percentage point of beta distribution with parameters A and $(m\text{-}A\text{-}1)/2$. In this paper, α is assumed to

be 0.01. Before building a phase II control chart, it is necessary to remove the scores of out-of-control points in the phase I control chart and re-compute the sample variance s_a^2 of the principal components. The phase II control chart to monitor the principal components can be then established by [58],

$$\sum_{a=1}^{A} (t_a^{new})^2 / s_a^2 \le A(m-1)/(m-A)F_{\alpha,A,m-A}, \tag{3.4}$$

where $F_{\alpha,A,m\text{-}A}$ is the upper α percentage point of F distribution with degrees of freedom A and (m-A-1)/2. The principal component t_a^{new} comes from the future observation and s_a^2 comes from the phase I control chart.

3.3 Case Studies

In this part, the real CMP experiments were conducted to validate our approach.

3.3.1 Experiments

To validate the proposed modeling and detection method, we designed validation experiments to generate abnormal process condition changes.

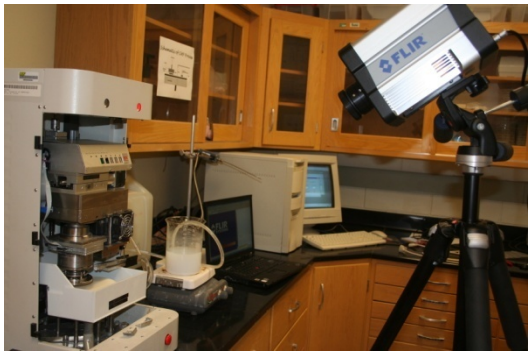


Figure 3.3 Experimental setup used in performing polishing experiments

Setup used in performing the polishing experiments is shown in Fig. 3.3. The polishing process was carried out on a bench-top CMP tester (model CP-4) manufactured by CETR Inc. During polishing, lateral force and normal force of the contact interface

were recorded at frequency 100Hz and COF could be recorded in situ by calculating the ratio of these two forces. Meanwhile, a FLIR® infrared camera was shooting at the polishing area to monitor in situ the temperature distribution on the pad. The average temperature of the polishing zone (Fig. 3.4) on the pad was recorded at a frequency of 30 Hz. The online monitoring of the COF and temperature signals facilitates studying the interaction between chemical and mechanical process variables.

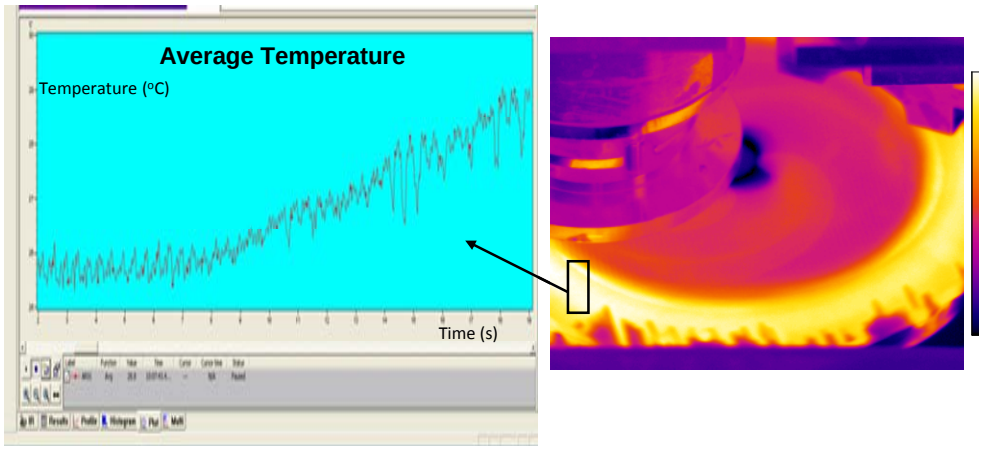


Figure 3.4 An example of recording temperature

The 6-inch diameter IC 1000k grove polishing pad was attached on the rotating bottom platen in CETR and 2-inch wafer coupon was attached to the upper polishing head. The slider velocity was maintained at 3 mm/sec during the whole experiment and the duration of each run was 3 minutes. Planerlite 7105 copper polishing slurry was mixed with 30% hydrogen peroxide in this experiment and was fed onto the center of the pad at the rate of 50mL/min. The slurry temperature was maintained at 30°C using a controllable heater (manufactured by Corning, Inc.). The newly changed pad was conditioned for two 20-min runs with 1-min polishing of dummy samples in between.

During polishing process, the pad was conditioned ex situ after each run. Process parameters of the experimentation are summarized in Table 3.1.

Table 3.1 Wafer polishing process parameters

Description	**Value**
Wafer coupon,	2×2in
Polishing pad	IC-1000k groove
Slurry	Planerlite 7105 copper slurry
Oxidizer	30% hydrogen peroxide
Slider movement	Offset: 5mm, speed: 3mm/s
Pad conditioning	Pressure: 2psi, rpm: 150
Conditioner	Diamond pad conditioner
Slurry flow rate	50mL/min

In the experiment, we simulated the case when the slurry was contaminated by impurities and faucet water during the polishing process (Table 3.2). Four samples (2 in.) were polished with slurry without contamination for 3 minutes, followed by another 3 minutes of polishing using the contaminated slurry.

Figure 3.5 gives the temperature and COF recordings during polishing of one sample. The left panel shows the data when polishing the wafer with normal slurry (350:30:650 for Slurry: H_2O_2: D.I. water) while the right panel displays the signal profile when the slurry was contaminated during polishing. Apparently, visual inspection of these two variables is not easy to identify underlying pattern changes.

Table 3.2 Experimental conditions

Rotational velocity (Polishing head vs. pad)	Pressure	Polishing time	Slurry : H_2O_2 : D.I. water
150 vs. 145 rpm	3 psi	3 min	350:30:650 (mL)
		3 min	Contaminated slurry*

* Contaminated slurry was formulated with 350 mL slurry, 30 mL 30%-H_2O_2, and D.I. water contaminated with 20 mL faucet water.

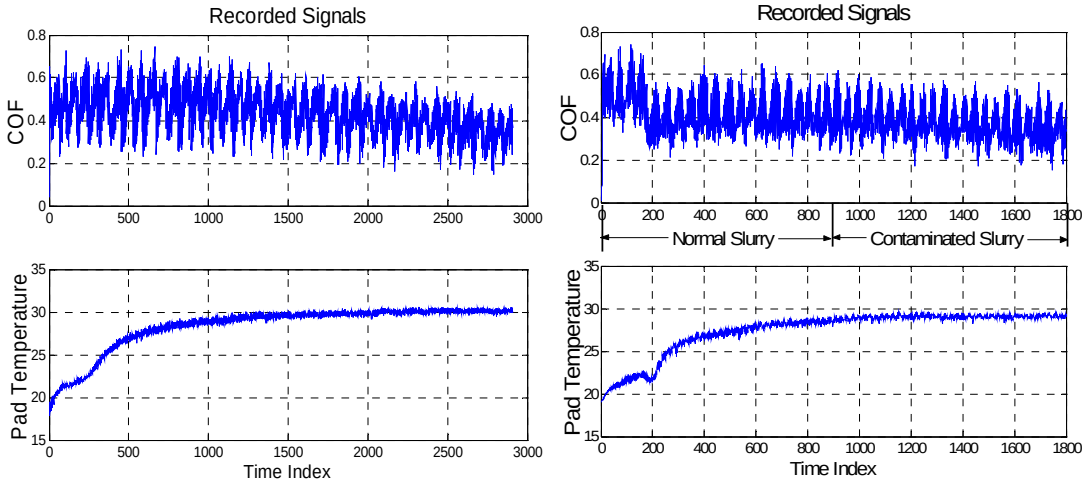


Figure 3.5 An example of signal recordings before and after slurry contamination

3.3.2 Results and Analysis

Instead of visual inspection, nonlinear dynamics modeling of phase synchronization assists statistical detection. This section demonstrates the method based on the data collected in Section 2.2.1.

The nonlinear dynamics modeling (Eq. (3.1)) results after statistical model selection are given in Fig. 3.6, where in the first row the dash line represents instantaneous frequency extracted by Hilbert transform, and the solid line is for instantaneous frequency predicted by the model. The prediction residuals (as shown in the second row) exhibit random patterns and no systematic trend or pattern (e.g. cyclic fluctuation), which implies that the model orders p=14 and q= 14 are adequate.

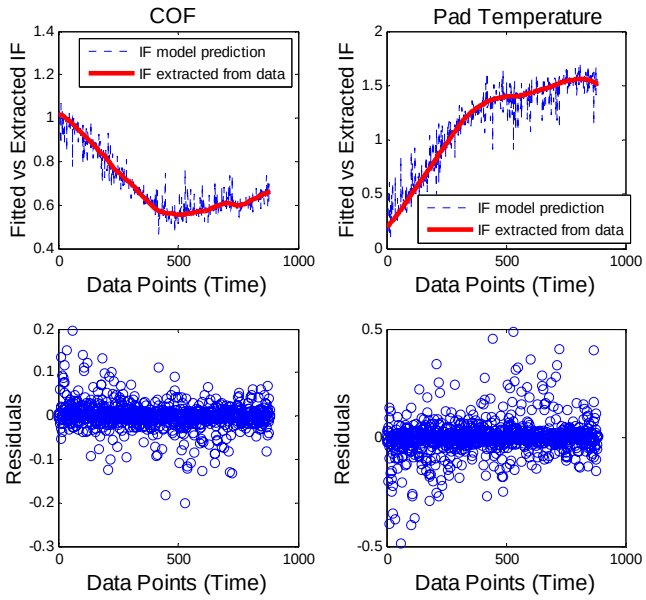
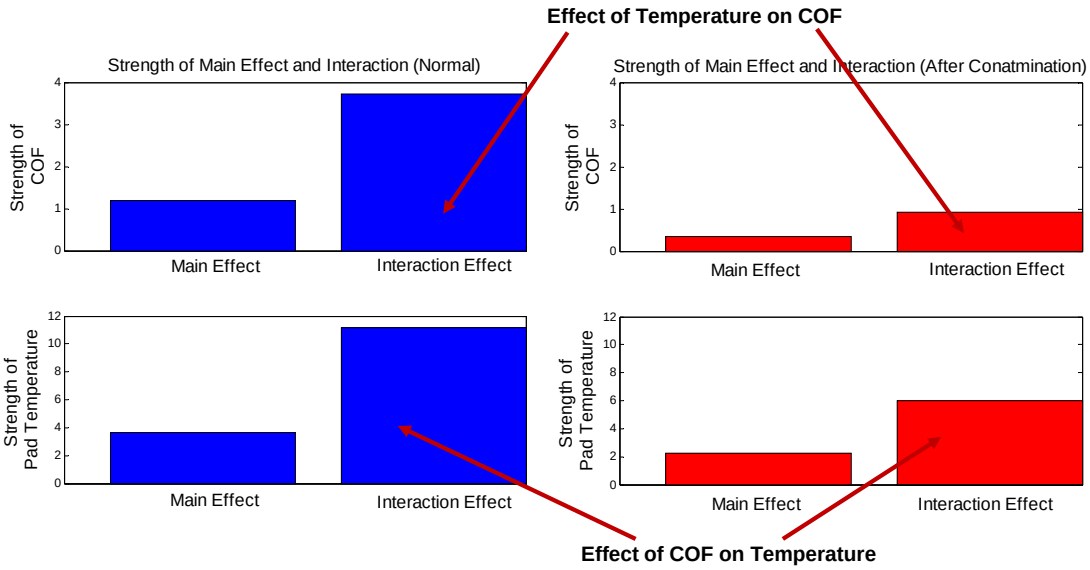


Figure 3.6 Phase nonlinear dynamics modeling results

In Fig. 3.7, the left and right panels compare the interaction strength (defined in Section II) before and after slurry contamination, respectively. The attached table shows the ratio between strengths of interaction and main effects. The interaction effect is significant because the ratio is far larger than 1 under two conditions. Compared to the normal condition (left panel), slurry contamination (right panel) significantly weakens the interaction effect between temperature and COF. Such interaction pattern could reflect process condition changes. For example, strong interaction between temperature and

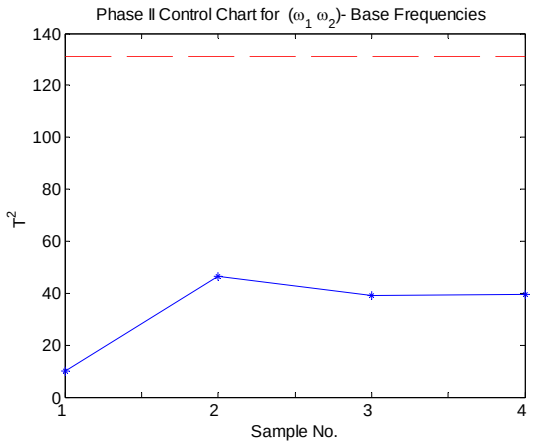
COF could be related to effective chemical reaction while the weakened interaction might indicate the chemical process is jeopardized or changed. Therefore, we propose a statistical detection method to identify the significant interaction pattern change.



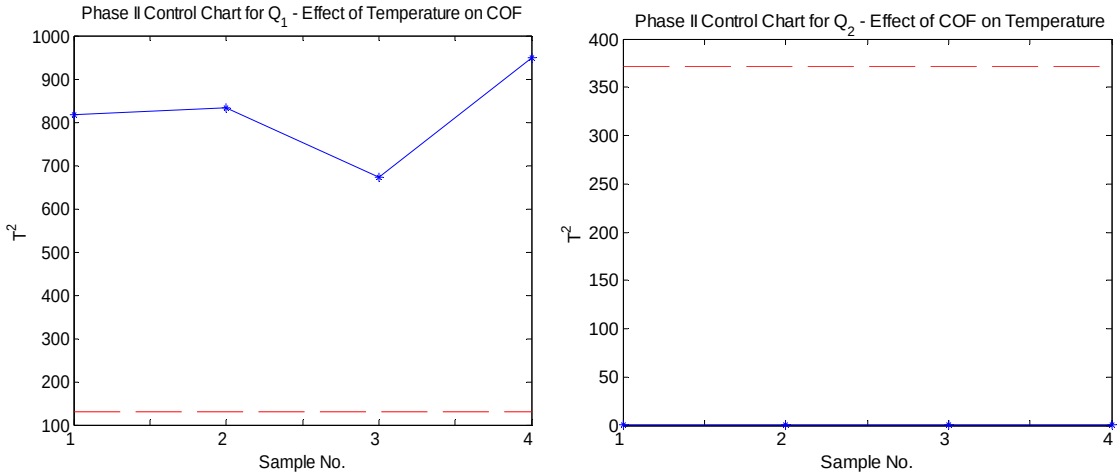
Interaction/Main Effect	Normal	Contamination
Effect of Temperature on COF	3.1222	2.6731
Effect of COF on Temperature	3.0458	2.6561

Figure 3.7 Strength of main and interaction effects

We built phase II T^2-Hotelling charts (Section III) to monitor respectively the change of interaction pattern after the slurry contamination, where Q_1 is the model term that shows the effect of temperature on COF and Q_2 shows the effect of COF on temperature. The base frequency (ω_1, ω_2) that usually corresponds to certain systematic process change is also monitored.



(a) Base frequency monitoring



(b) Interaction effect monitoring

Figure 3.8 Phase II control charts for main and interaction effects

We can see from Fig. 3.8 that all the sample points are under control limit (dash line) for monitoring main effects and Q_2 whereas all the samples of Q_1 are beyond the limits. This implies that the interaction effect of temperature on COF has significantly changed after slurry contamination. However, slurry contamination does not significantly affect the base frequency and the interaction of COF on temperature. Combining the results of interaction strength, we can know that the Q_1 has been

significantly weakened after the slurry contamination. In this study, such type of interaction pattern change can be used as an indicator of less effective chemical reaction that could be related to slurry problems.

3.4 Summary

This chapter studied timing correlation of multiple functional process variables (FPVs) in phase domain for CMP process condition monitoring and diagnosis. Considering the oscillatory pattern in the FPVs, we first established a nonlinear dynamics model to capture the main and interaction effect in the rhythm of cyclic components. It can be estimated through regression analysis followed by statistical screening procedures. Following the concept of statistical effects in design of experiments, we defined the order and strength of interaction, through which the directionality of interaction can be identified. Uncovering interaction directionality will assist to understand physical interaction among multiple FPVs. A statistical method of condition change detection was then developed by monitoring the interaction patterns using statistical process control tools. The extracted interaction patterns are especially helpful for detecting abnormal condition caused by hidden factors that may not be easily identified.

The proposed method was applied to data analysis on CMP experiments, where we generated slurry contamination during CMP polishing. The modeling result showed strong interaction strength on both directions of COF-temperature interaction. Statistical control charts indicated that interaction effects of temperature on COF were significantly changed after slurry contamination, whereas the reversed interaction effect and base frequencies of both signals remained unchanged. Combining the results of interaction strength analysis, we can further conclude that the interaction effect of temperature on

COF has been significantly weakened. In this study, the weakened interaction pattern is an indicator of less effective chemical reaction due to slurry problems. These facts may lead to a new diagnosis method based on interaction modeling for abnormal process change caused by slurry problems.

Chapter 4

Analysis of Interaction Structure among Multiple FPVs for Process Control in Semiconductor Manufacturing

From previous chapters, we have shown that the complex interaction patterns among functional process variables (FPVs) in semiconductor manufacturing processes could indicate process condition changes. We developed a nonlinear dynamics model to describe interactions among FPVs, which was further used to monitor process condition changes. However, the interaction structure among three or more FPVs has not been thoroughly investigated for the purpose of process control. In this work, we first extend our previously developed nonlinear dynamics model by considering the autocorrelation in each FPV. A generalized least square (GLS) method is applied to estimate the extended model. The interaction structure among FPVs is represented as a complex network in which the directionality and strength of interaction are discovered from the extended nonlinear dynamics model. To validate the proposed method, we first conduct simulation study using van der Pol oscillators. Then two sets of real experimental data from chemical mechanical planarization process are used to investigate the interaction structure change over a polishing cycle. The results show that the extracted patterns of interaction structure among FPVs could aid to uncover the polishing mechanisms and provide more insights for condition monitoring and diagnosis.

The chapter is organized as follows. Section 4.1 briefly reviews the work done in previous chapters. In Section 4.2 we first extend our previously developed nonlinear dynamics model by considering the autocorrelation in each FPV. A generalized least square (GLS) method is used to estimate the extended model. The interaction structure among FPVs is represented as a complex network in which the directionality and strength of interaction are discovered from the extended nonlinear dynamics model. We demonstrate the interaction analysis approach through a simulation of van der Pol oscillators. In Section 4.3, two sets of real experimental data from chemical mechanical planarization process are used to investigate the interaction structure change over a polishing cycle. Conclusion is given in Section 4.4.

4.1 Review of Nonlinear Dynamics Model of FPVs

In semiconductor manufacturing many process factors are involved to affect product quality. As an example in chemical mechanical planarization (CMP), factors such as applied force, pad property, and slurry flow rate would jointly (not independently) impact the quality of polished wafers. The interactions among process factors can be very complex. To understand the ways that process factors affect product quality, process variables such as coefficient of friction (COF) and polishing pad temperature are collected online to predict the realtime process conditions. If process variables are continuously observed and vary with time, these functional process variables (FPVs) may contain rich process information. For instance, COF between the wafer and the pad provides information regarding the tribological condition at the interface. An abrupt and large variation in COF could be a realtime indication of pad failure, large particles on the pads, or underlying barrier layer exposure on the wafer. Highly correlated to COF, pad

temperature is another important FPV that depicts heat variations generated through friction and chemical reaction. Previous studies have investigated FPVs both experimentally and analytically [60][61]. Our experience suggested that certain process changes may not be easily detected without collectively studying these FPVs. Simultaneously analyzing these FPVs and their interaction patterns could bring additional insights into process condition changes and new opportunities for process improvement [62][63].

In [62][63] we have shown that the complex interaction patterns among FPVs in semiconductor manufacturing processes could indicate process condition changes. We developed a nonlinear dynamics model to describe interactions among FPVs, which was further used to monitor process condition changes. However, the interaction structure among three or more FPVs has not been thoroughly investigated for the purpose of process control. Below we briefly review related studies.

The common approaches of analyzing correlation or interaction are cross-correlogram and cross-spectrum methods [64-66]. They might be easily affected by artifacts and lead to improper detections, especially for the non-stationary signals collected in CMP process. Coherence and cross spectrum methods aim to analyze the correlation of paired signals in the frequency domain and are most commonly used with continuous signals [65][39]. Phase synchronization was developed to detect timing-correlation in phase domain while discarding the effect of the amplitude of signals [39][40]. Similar to the correlation coefficient in the time domain, coherence, entropy, or mutual information indices in phase domain have been proposed to detect synchrony in

paired signals [70-72]. A nonlinear dynamics model was further developed in [73] to study phase synchronization:

$$\begin{aligned} d\phi_1 / dt &= \omega_1 + q_1(\phi_1) + f_1(\phi_1, \phi_2) + \varepsilon_1(t) \\ d\phi_2 / dt &= \omega_2 + q_2(\phi_2) + f_2(\phi_2, \phi_1) + \varepsilon_2(t) \end{aligned} \tag{4.1}$$

where $\varphi_{1,2}$ are the phases of coupling variables, $\omega_{1,2}$ are the base angular frequencies or natural frequencies, and $\varepsilon_{1,2}$ are the noisy perturbations. $f_{1,2}$ are defined as functions which involve interactions between coupling variables, and $q_{1,2}$ are self-provoked functions. This model can be easily extended to the case with more than two coupling variables.

The major challenge, however, is to find interaction functions $f_{1,2}$ in Eq. (4.1) with adequate orders after Fourier transformation. To improve the mode, we defined main effects and interaction effects of FPVs in [62] and further demonstrated it in CMP process monitoring. However, the temporal patterns, especially the often strong autocorrelation in FPVs were not considered therein. This could lead to inadequate phase dynamics models and provide an incomplete picture of the complex spatio-temporal patterns in FPVs. Furthermore, the complex interaction structure among three or more FPVs has not been thoroughly investigated. As shown in Fig. 4.1, the general correlation analysis could not reveal the directionality and strength of interaction, and network structure among multiple FPVs. As shown in our case study, analyzing interaction structures could assist to understand more insights of the polishing mechanisms.

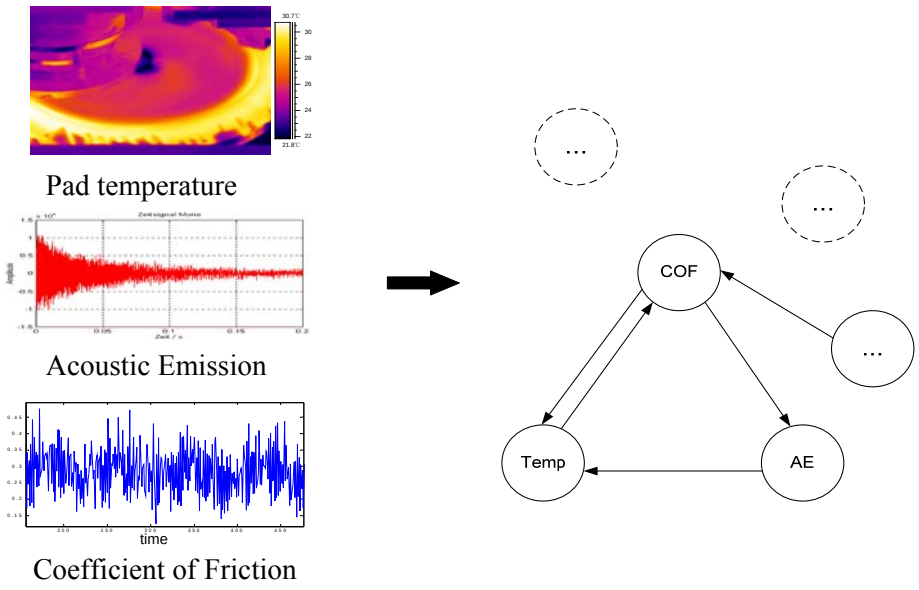


Figure 4.1 Interaction structures represented as a network

Rosenblum *et al.* [74] first investigated interaction structure based on mutual prediction. In their study the canonical structure (three oscillators in a ring) was identified through pairwise analysis of coupled oscillators. More complex structures were investigated using partial directed coherence method [75]. All these methods examine the directionality of two-way interactions to identify interaction structures. Three-way or high order interactions are not considered. Moreover, variables in the interacting network are assumed to be known, and hence hidden influential variables could be missed. Therefore, a new method that could distinguish different orders of interaction would be preferable so as to identify more complex interacting mechanisms.

4.2 Analysis of Interaction Structure among Multiple FPVs

4.2.1 Extended Nonlinear Dynamics Model for Interaction among FPVs

In [62] we characterized the interaction among multiple FPVs by

$$d\varphi_k(t)/dt = \omega_k + Q_k[\varphi_1(t), \varphi_2(t), ..., \varphi_p(t)] + \varepsilon_k, \quad k = 1, 2, ..., p, \qquad (4.2)$$

where phase variables $\varphi_k(t)$'s are obtained through Hilbert transform of FPVs, and $Q_k(.)$ describes interaction among these phase variable $\varphi_k(t)$. $Q_k(.)$ is approximately periodic which can be approximated by Fourier expansion:

$$Q_k(\varphi_1,\varphi_2,...,\varphi_p)=\sum_{m_1,m_2,...,m_p}[a_j^{m_1,m_2,...,m_p}\cos(\sum_{j=1}^{p}m_j\varphi_j)+b_k^{m_1,m_2,...,m_p}\sin(\sum_{j=1}^{p}m_j\varphi_j)],\ k=1,2,\ldots p, \quad (4.3)$$

where the superscripts m_j's in the Fourier expansion are integers and cannot be zeros simultaneously. The values of coefficients $a_j^{m_1,m_2,...,m_p}$ and $b_k^{m_1,m_2,...,m_p}$ were estimated through the Ordinary Least Square (OLS) regression method.

When sampling frequency is high, model (4.2) might overlook the strong autocorrelation or temporal patterns in FPVs. Figure 4.2 shows the fitted results for one segment of COF. The order of the Fourier expansion is m =4. As can be seen, although the fitted model (dash line) could capture the main trend of the original instantaneous frequency (bold line), the residuals show a strong autocorrelation. Without considering the potential autocorrelation could leads to approximating $Q_k(.)$ with many sine and cosine terms, which may be hard to interpret physically. Therefore, we extend the model (4.2) by imposing a structure on residuals,

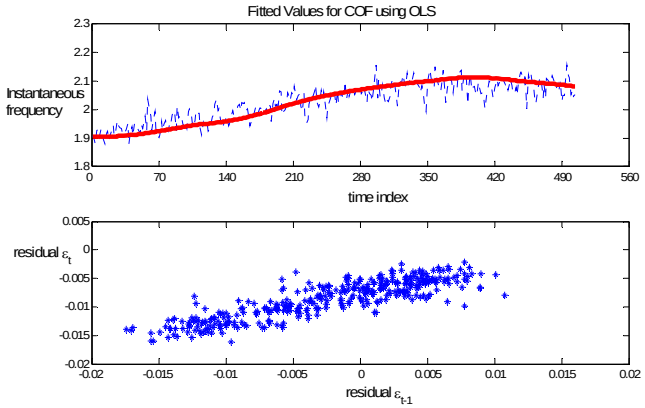


Figure 4.2 Temporal patterns in FPVs

$$\text{Cov}[\varepsilon_k(t), \varepsilon_k(t+1)]=\sigma^2{}_k\rho_l, \text{ time lag } l=0,1,2,\ldots \quad (4.4)$$

Compared with the previous model based on the assumption of noise independence, the noise term in the new model permits an auto-correlated structure, i.e., error terms at time t and $t+l$ have correlation coefficients ρ_l. For example, the error ε_1 of the COF may follow a first-order autoregressive model (AR(1) model), $\varepsilon_1(t)=\eta\varepsilon_1(t-1)+v_1(t)$, where $v_1(t)$ is the white noise, $v_1(t)\sim N(0, \sigma_v^2)$. Then the coefficient $\rho_l=\eta^l$ and $\text{Cov}[\varepsilon_1(t), \varepsilon_k(t+l)]=\sigma_1^2\eta=\sigma_v^2\eta/(1-\eta^2)$. The first derivatives of phase $\varphi_k(t)$'s will serve as response variables while cosine and sine functions in the Fourier expansion of $Q_k(.)$'s will be predictors. The model coefficients ω_i, $a_j^{m_1,m_2,\ldots,m_p}$, $b_j^{m_1,m_2,\ldots,m_p}$, and ρ_l can be estimated using the Generalized Least Squares (GLS) method [76].

4.2.2 Interaction Structure Analysis

The proposed interaction structure analysis approach is outlined in Fig. 3. Since the complexity of interaction analysis increases exponentially with the number of FPVs, we use a network involving three nodes as an example to demonstrate the procedure. With the extended phase dynamics model we will first identify the strength of interactions. As defined in 62, the main effects include $cos[\varphi_1(t)]$, $cos[\varphi_2(t)]$, $cos[\varphi_3(t)]$, $sin[\varphi_1(t)]$, $sin[\varphi_2(t)]$, and $sin[\varphi_3(t)]$. The two-way interaction effects or the first order interactions include $cos[\varphi_1(t)+\varphi_2(t)]$, $cos[2\varphi_1(t)+\varphi_3(t)]$, etc. The three-way interaction effects or the second order interactions contain $cos[\varphi_1(t)+\varphi_2(t)+\varphi_3(t)]$, etc. The term $\sqrt{(a_k^{s,r,w})^2+(b_k^{s,r,w})^2}$ represents the amplitude in signal processing and $[(a_k^{m_1,m_2,m_3})^2+(b_k^{m_1,m_2,m_3})^2]$ is related to the power of that frequency component. The strength of each frequency component in a main effect or an interaction effect can be thus defined using

the concept of the power, e.g., $[(a_k^{m_1,m_2,m_3})^2+(b_k^{m_1,m_2,m_3})^2]$. The strength of the main effects/interaction effects is defined as the summation of the power of every frequency component in all the main/interaction effects, e.g., $\sum_{m_1,m_2,m_3}[(a_k^{m_1,m_2,m_3})^2+(b_k^{m_1,m_2,m_3})^2]$. Then we could construct a bar chart for each FPV which shows the strength of main effect and interaction effects (see an example in Fig. 4.9).

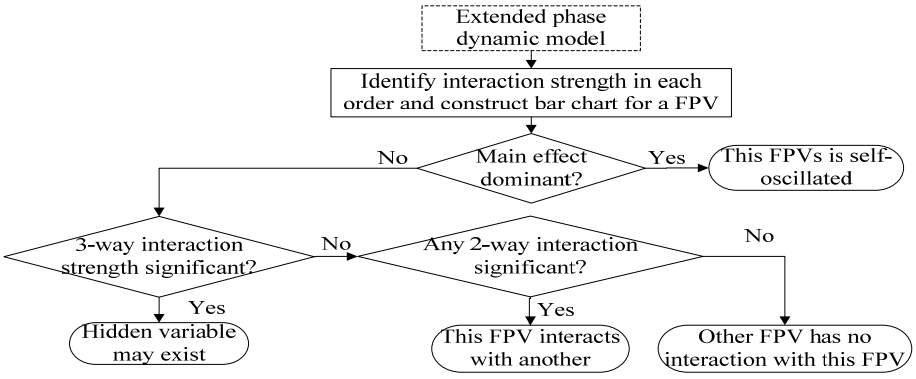


Figure 4.3 Procedure of analyzing interaction structure

From the bar charts we will investigate interaction patterns. We analyze four important cases of interaction structures:

- *Self-oscillated FPV* As illustrated in Fig. 4.4, if main effect of a FPV is dominant and all of the two-way and high order interaction effects are insignificant, the FPV does not interact with others.

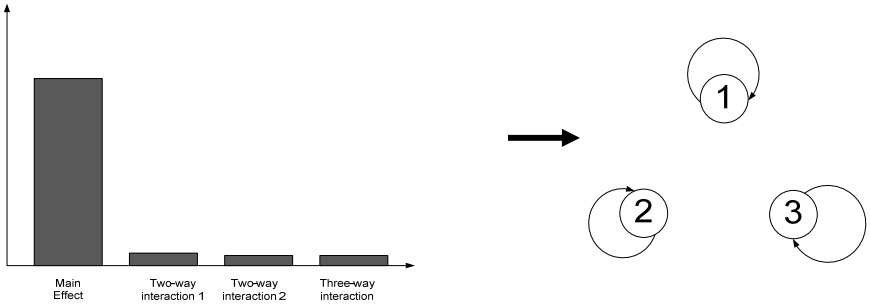


Figure 4.4 Self-oscillated variables in system

- *Clockwise interaction pattern* In this case, each FPV is affected by the other in clockwise manner (Fig. 4.5). From the bar chart, both interaction effects are significant, but one is stronger than the other. Meanwhile, the main effect and three-way or high order interaction are relatively weak.

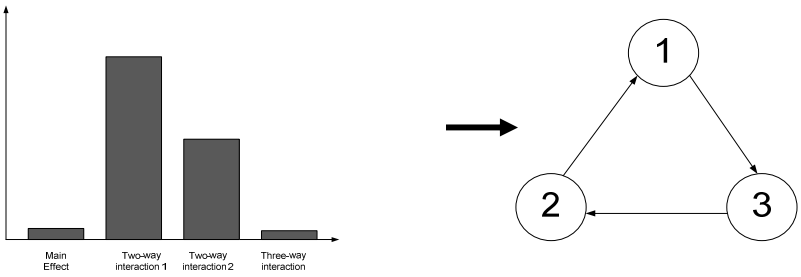


Figure 4.5 Clockwise interactions among FPVs

- *Symmetric interactions among FPVs* If strengths of all two-way interactions are similar and higher order interactions are insignificant, we may encounter symmetric interaction structure (Fig. 4.6).

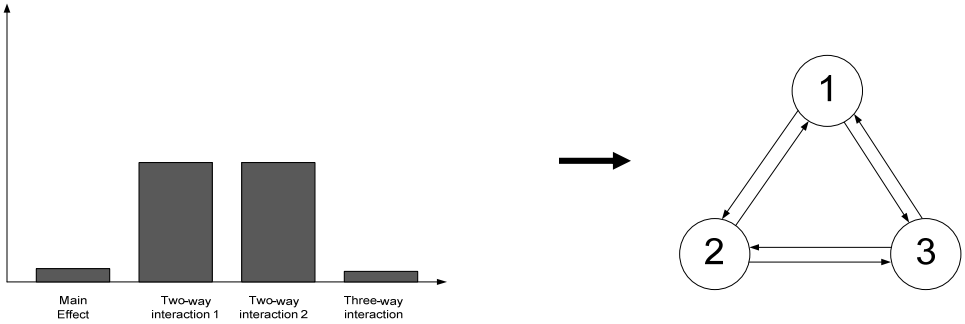


Figure 4.6 Symmetric interactions among FPVs

- *Hidden FPVs in a network* If only one three-way interaction effect is significant, we may suspect that at least one hidden variable exists in the network. As shown in Fig. 4.7, the bar chart displays interaction strength for node 2. If

three-way interaction effects are weak in all the other FPVs, a hidden FPV might interact with the network through node 2. Similar idea for identifying hidden variables could be found in [77].

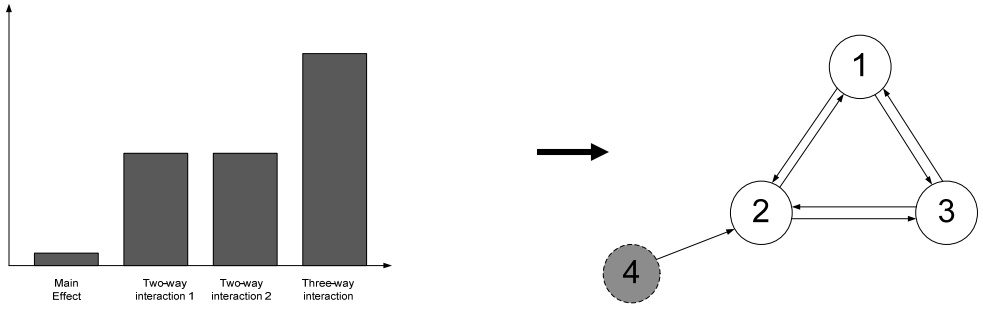


Figure 4.7 Hidden FPVs interact with the network through Node 2

The last issue is to determine the directionality of interaction. If the strength of two-way interaction in node *i* is stronger than that of the corresponding two-way interaction in node *j*, then node *j* has stronger influence on node *i*.

To validate the proposed approach for analyzing interaction structures, four-channel van der Pol oscillator system is simulated using Matlab [78] as follows:

$$\begin{cases} \dfrac{d^2x_1}{dt^2} = u(1-x_1^2)\dfrac{dx_1}{dt} - \omega_1^2 x_1 + \lambda_{1\leftarrow 2}(x_2 - x_1) + \lambda_{1\leftarrow 3}(x_3 - x_1) + \lambda_{1\leftarrow 4}(x_4 - x_1) + \xi_1 \\ \dfrac{d^2x_2}{dt^2} = u(1-x_2^2)\dfrac{dx_2}{dt} - \omega_2^2 x_2 + \lambda_{2\leftarrow 3}(x_3 - x_2) + \lambda_{2\leftarrow 4}(x_4 - x_2) + \lambda_{2\leftarrow 1}(x_1 - x_2) + \xi_2 \\ \dfrac{d^2x_3}{dt^2} = u(1-x_3^2)\dfrac{dx_3}{dt} - \omega_3^2 x_3 + \lambda_{3\leftarrow 4}(x_4 - x_3) + \lambda_{3\leftarrow 1}(x_1 - x_3) + \lambda_{3\leftarrow 2}(x_2 - x_3) + \xi_3 \\ \dfrac{d^2x_4}{dt^2} = u(1-x_4^2)\dfrac{dx_4}{dt} - \omega_4^2 x_4 + \lambda_{4\leftarrow 1}(x_1 - x_4) + \lambda_{4\leftarrow 2}(x_2 - x_4) + \lambda_{4\leftarrow 3}(x_3 - x_4) + \xi_4 \end{cases} \tag{4.5}$$

where $x_{1,2,3,4}$*'s* are four FPVs , u is the nonlinear self-weight parameter, $\omega_{1,2,3,4}$'s are linear self-weight parameter; $\lambda_{i\leftarrow j}$'s are weights from coupled variables, and $\xi_{1,2,3,4}$ are white noises with normal distributions.

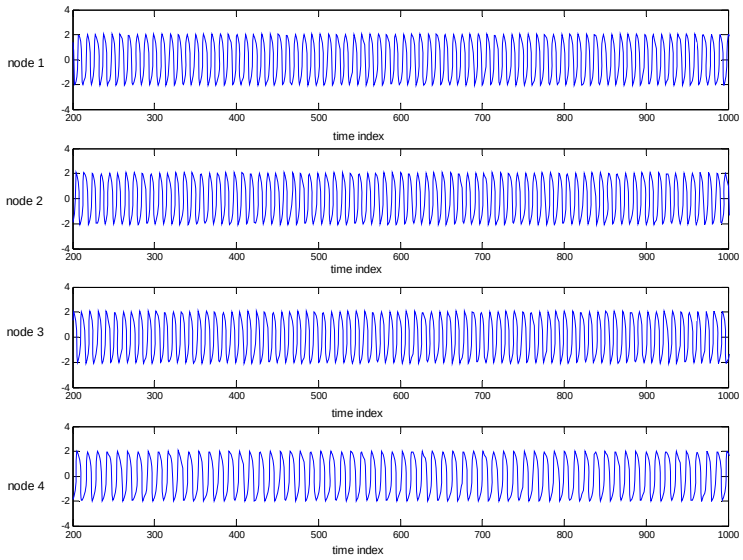
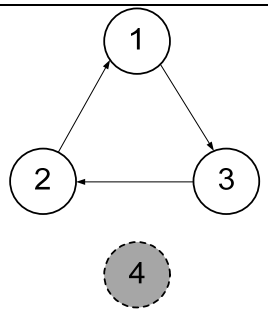
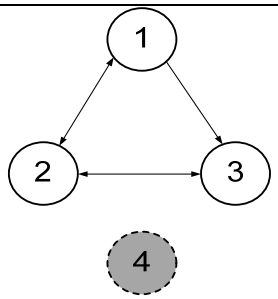
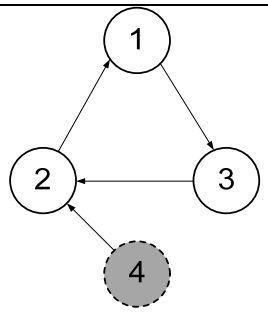


Figure 4.8 Four channel simulated signals via van der Pol oscillators

Three canonical types of interaction structures are generated by varying parameter values (Table. 4.1). The figures on the right column depict interaction structures under variant dynamical systems. Nodes 1, 2, 3 and 4 are represented simulated FPVs, and arrows between nodes represent directed interactions determined by those parameters.

Since some important process variables might be missed in a real dynamical system or process, we choose node 4 as hidden variable to be discovered by the proposed approach. The nonlinear parameter u is fixed to 3 for high nonlinear feature in dynamical systems. To avoid one FPV being modulated by another, small values were given to $\lambda_{i\leftarrow j}$ and ω_i. 5% Gaussian white noise of original signal in decibel was added to each channel.

Table 4.1 Interaction structure types and corresponding parameters

Type A $\omega_1 = 0.5$ $\omega_2 = 0.5$ $\omega_3 = 0.5$ $\omega_4 = 0.5$	$\lambda_{1\leftarrow 2}= 0.5$ $\lambda_{1\leftarrow 3}= 0.5$ $\lambda_{1\leftarrow 4}= 0$ $\lambda_{3\leftarrow 4}= 0$ $\lambda_{3\leftarrow 1}= 0.5$ $\lambda_{3\leftarrow 2}= 0$	$\lambda_{2\leftarrow 3}= 0.5$ $\lambda_{2\leftarrow 4}= 0$ $\lambda_{2\leftarrow 1}= 0$ $\lambda_{4\leftarrow 1}= 0$ $\lambda_{4\leftarrow 2}= 0$ $\lambda_{4\leftarrow 3}= 0$	
Type B $\omega_1 = 0.5$ $\omega_2 = 0.5$ $\omega_3 = 0.5$ $\omega_4 = 0.5$	$\lambda_{1\leftarrow 2}= 0.5$ $\lambda_{1\leftarrow 3}= 0$ $\lambda_{1\leftarrow 4}= 0$ $\lambda_{3\leftarrow 4}= 0$ $\lambda_{3\leftarrow 1}= 0.5$ $\lambda_{3\leftarrow 2}= 0.7$	$\lambda_{2\leftarrow 3}= 0.5$ $\lambda_{2\leftarrow 4}= 0$ $\lambda_{2\leftarrow 1}= 0.7$ $\lambda_{4\leftarrow 1}= 0$ $\lambda_{4\leftarrow 2}= 0$ $\lambda_{4\leftarrow 3}= 0$	
Type C $\omega_1 = 0.5$ $\omega_2 = 0.5$ $\omega_3 = 0.5$ $\omega_4 = 0.5$	$\lambda_{1\leftarrow 2}= 0.5$ $\lambda_{1\leftarrow 3}= 0$ $\lambda_{1\leftarrow 4}= 0$ $\lambda_{3\leftarrow 4}= 0$ $\lambda_{3\leftarrow 1}= 0.5$ $\lambda_{3\leftarrow 2}= 0$	$\lambda_{2\leftarrow 3}= 0.5$ $\lambda_{2\leftarrow 4}= 1$ $\lambda_{2\leftarrow 1}= 0$ $\lambda_{4\leftarrow 1}= 0$ $\lambda_{4\leftarrow 2}= 0$ $\lambda_{4\leftarrow 3}= 0$	

1500 data points were sampled for each FPV. The initial 200 points of data were cut off from original data series to avoid instability caused by initial values. Figure 4.8 displays the simulated signals from each node in type *A* structure. Since the interaction relationship could not be distinguished via visual study, the proposed nonlinear dynamics models are required to unveil the interaction mechanism.

According to the principles for modeling fitting discussed in Section 4.2.1, low order of Fourier expansion was preferred. Stepwise regression was applied to screen out those insignificant terms in the phase dynamics model. Terms with very low strengths of main effects or interactions effect were also dropped off. The order of Fourier expansion was finally set to 2 with *p*-value being 0.05 in model selection. By computing the strength of main effects and interaction effects for each node, we obtain bar charts shown in Figure 4.9.

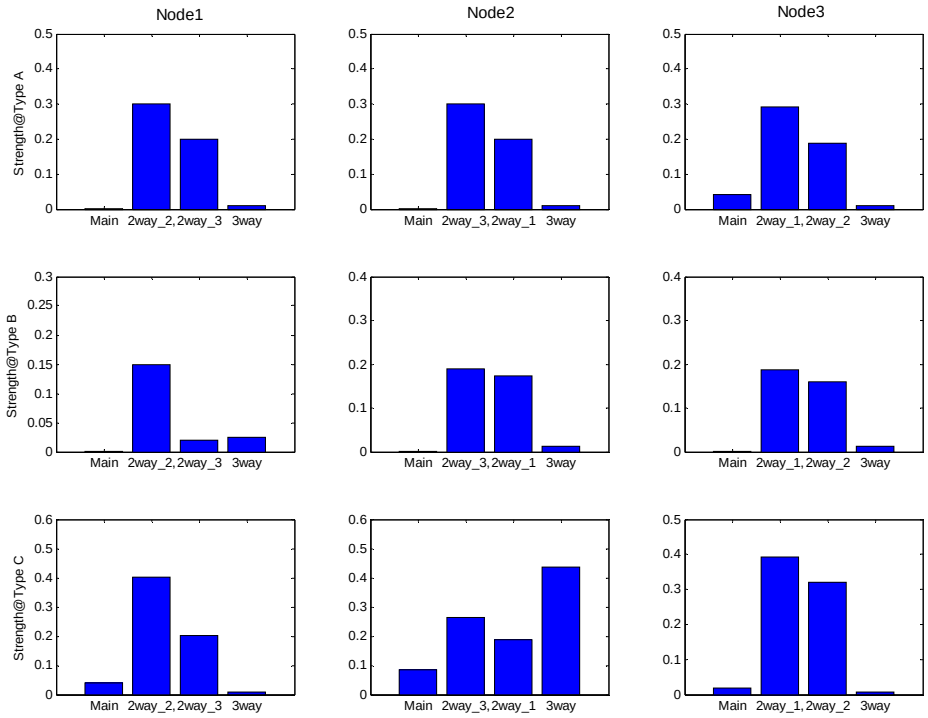


Figure 4.9 Bar charts of interaction strength under three interaction structures

It can be seen clearly in the first row of Fig. 4.9 that both interaction effects are significant, but one is stronger than the other in each node. For instance, in node 2, two-way interaction 2←3 and 2←1 are both significant, but 2←3 is more dominant. The pattern repeats at each node which suggests clockwise interaction structure (nodes in a

ring pattern) in type *A* nonlinear system. Two-way interaction 2←1 in node 2 actually represents an indirect influence between nodes 1 and 2.

In second row of the charts, two-way interactions 1←2 (node 1), 2←3 and 2←1 (node 2), and 3←1, 3←2 (node 3) are significant. This indicates the structure of type B in Table 4.1.

The third row is similar to the first row except that the three-way interaction 2←(1, 3) in node 2 appears significant. Since all the two-way interactions suggest a ring pattern, the unexpected three-way interaction 2←(1, 3) may be due to a hidden variable (node 4). Node 4 influences the network through node 2.

As clearly shown in the simulation study, our method determines the directionality of interaction via the definition of strength. Furthermore, we consider three-way or higher order interactions which assists to understand complex structure. The mutual prediction algorithm [74] constructs an index with value from -1 to 1 using two oscillators each time. Depending on the sign of the index, the directionalities of each oscillator are able to be obtained for structures like Type A. Since high order interactions are not modeled, complex structures with hidden factors are hard to be determined by that method.

4.3 Application to Identification of Interaction Structure Patterns in Real CMP

Understanding the interaction structures has important implication in semiconductor manufacturing process control. Specifically the discovered interaction patterns among process variables over time could assist to understanding the underlying physical mechanisms. In this section we will demonstrate this point using CMP process.

Two CMP process conditions are investigated: diamond conditioners with abrasive size 8μm and 100μm chosen to condition the polishing pads, respectively.

The whole polishing process was conducted on a bench-top CMP tester (model CP-4) manufactured by CETR Inc.. The 6-inch diameter IC 1000-A4 perforated polishing pad (manufactured by Rodel, Inc.) was attached on the rotating bottom platen of the CMP tester. The 2-inch copper wafer coupon was attached to the upper polishing head. Cabot 5003 copper polishing slurry was mixed with 2.5% hydrogen peroxide. The slurry was fed into the center of the pad at the rate of 50mL/min. The pressure were set to 2psi both on conditioning and polishing process, and polishing head rotating speed was 150 rpm.

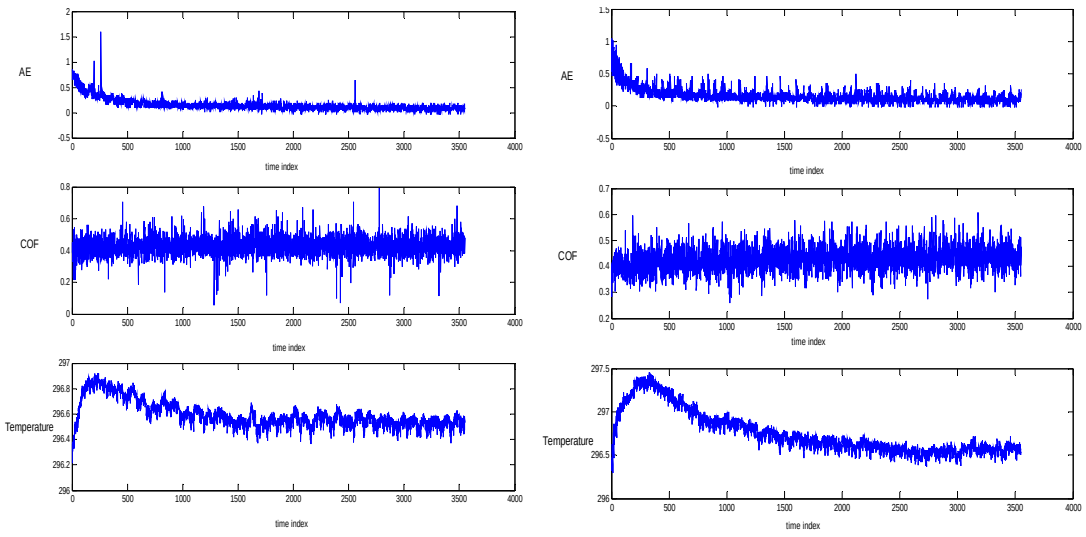


Figure 4.10 FPVs with diamond particle size 8μm(L) and 100μm(R).

We made three replicates under each pad condition and each run lasted 3 minutes. During each polishing process, three FPVs were collected simultaneously at sampling frequency 20Hz. The coefficient of friction could be obtained by gathering shear and normal forces with sensors installed on the polishing head, and acoustic emission (AE) could be collected through AE sensor on the back of wafer holder. A FLIR infrared

thermal camera was used to collect thermal data. The focused thermal zone was selected on the polishing pad adjacent to the wafer-pad interface to record the average temperature. Since the temperature in this interface between wafer and polishing pad cannot be obtained directly, the zone we selected in this way might best indicate the average process temperature [79]. It should be addressed here that these three process variables could properly represent the chemical-mechanical polishing mechanism [80][0]. Figure 10 shows one sample of the original signals collected under each process condition. Directly observing original signals provides limited information.

Since the main trend of each FPV would not be dramatically altered by interactions, we first remove the trend from collected data through cubic-spline smoothing method. Figure 4.11 displays the one sample of temperature data before and after detrending.

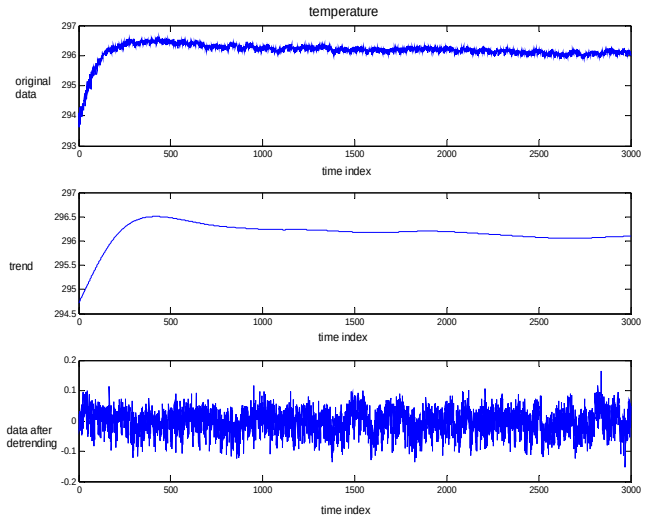


Figure 4.11 Sample data before and after detrending

To understand the dynamics of interaction mechanisms during the whole polishing process, we monitor the process via sliding windows. When determining

window size and length of overlapping period, we intend to (1) maintain a smooth transition between adjacent windows based on the belief that there is no sharp change of polishing mechanism in a short period; and (2) avoid large window size for potential miss-detection of changes. We first start with large window size to test the polishing condition, and reduce the window size until difference is apparent between two windows. Finally, our sliding window size is determined to be 25 seconds (500 data points). We update the sliding window every 5 seconds (100 data points) for good computing efficiency without losing any information of condition change. The overlapping period is 400 points to have smooth transition between adjacent windows. In each window, the dynamics model was fitted by GLS, and the model adequacy was checked. Figure 4.12 shows one sample of the instantaneous frequency extracted from original data versus fitted one in single window, and the residuals are also plotted to check the adequacy of autoregressive terms. It is found that model order $n = 2$ and autoregressive order $l = 5$ would have good prediction.

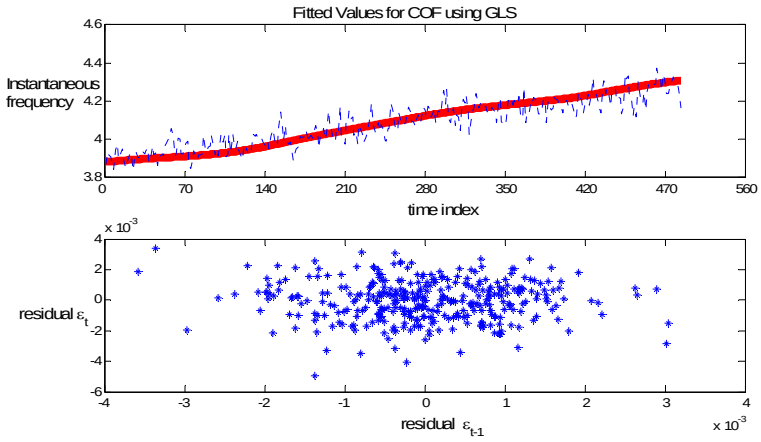


Figure 4.12 One sample of model fitting result in a single window

Strength of main effects, two-way and three-way interaction effects is computed in each window. The interaction structures and their changes are able to be deduced by

analyzing the bar charts. Figure 4.13 shows structure changes under polishing condition using 100μm diamond particle size of conditioner. As can be seen, the two-way interactions AE←temperature and COF←temperature are significant at the beginning, and they decrease gradually over time. Meanwhile, the strength of two-way interactions AE ← COF and temperature ← COF increase. Other interaction effects seem not significant enough to affect interaction structures.

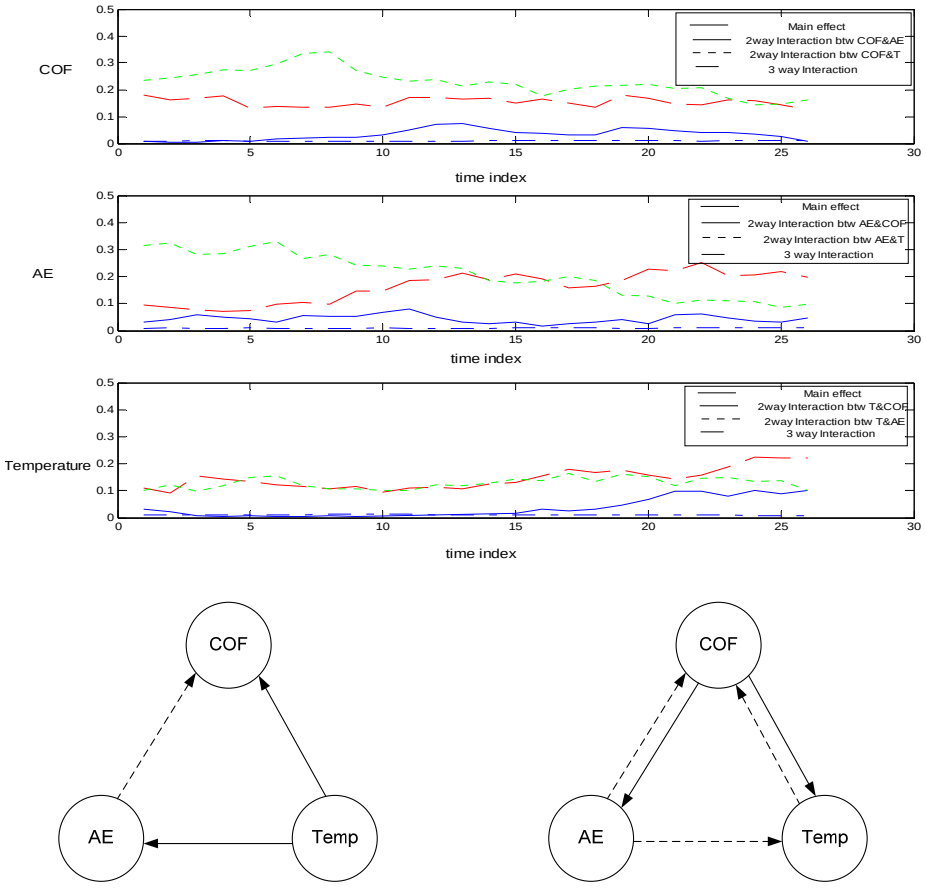


Figure 4.13 Interaction structure analysis: 100μm diamond particle size of conditioner

The dynamics of interaction change are represented by network structure change in Fig. 4.13. The solid line with arrow represents significant effects, and dash line with

arrow represents weak effects in term of strength. From the network, we could see temperature had greater impact on AE and COF initially, and COF influenced AE and temperature more afterwards. It is known that temperature on pad reflects the heat generated by friction and chemical reaction, AE is related to the material deformation, and COF is determined by the interface properties [80][0]. We therefore interpret the network structure change as follows. The chemical reaction was dominant at the beginning and temperature on was mainly affected heat released from wafer surface softening and weakening by slurry. Afterwards, the mechanical friction played a bigger role because of the non-uniformity on the wafer surface in our polishing process.

Figure 4.14 shows the analysis of polishing condition using 8μm diamond particle size of conditioner. Compared with Fig. 13, strength of interaction effects is relatively mild. Initially only the two-way interaction COF←temperature is significant. This interaction also fades as the polishing process goes. The two-way interaction AE←COF appears relative strong compared to other interaction effects during the later period of the polishing process. This pattern might be interpreted as that the smaller diamond particle size will generate smoother the polishing pad. This will lead to a weaker mechanical reaction relative to its chemical reaction.

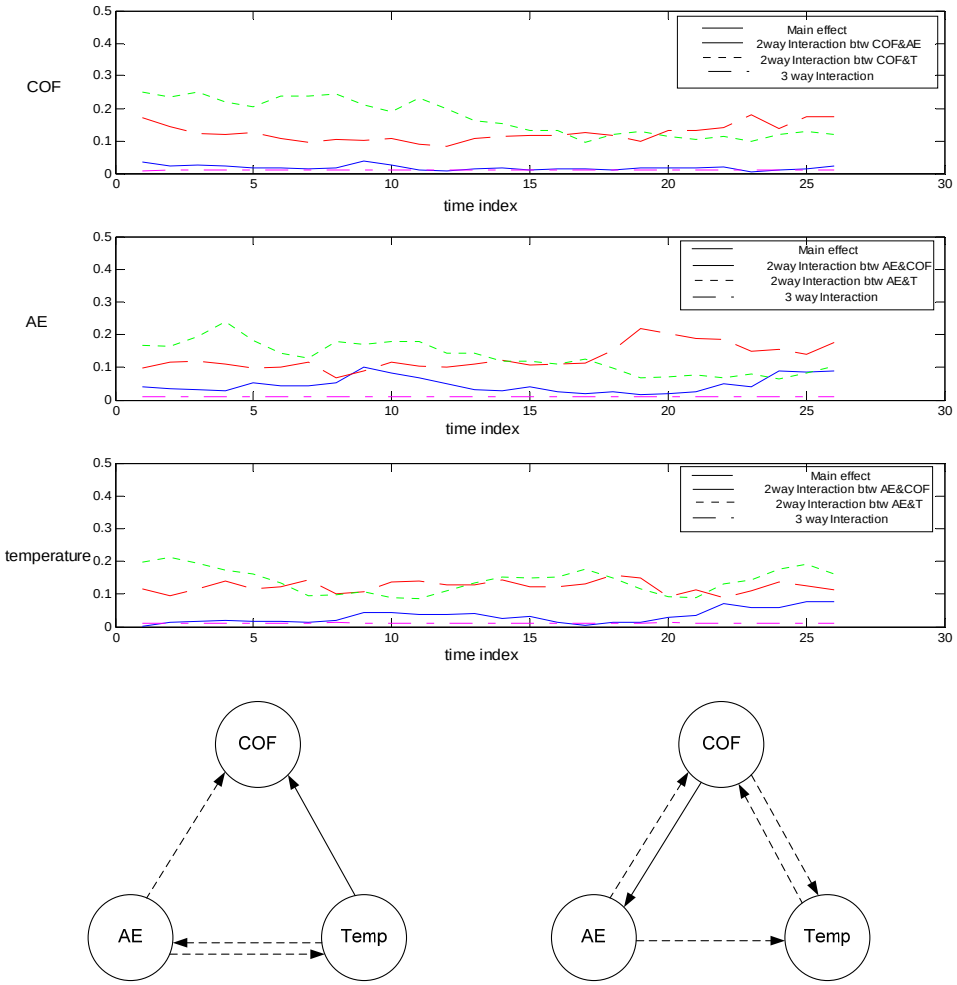


Figure 4.14 Interaction structure analysis: 8μm diamond particle size of conditioner

4.4 Summary

This chapter developed a methodology to analyze the dynamic interaction structures among multiple functional process variables (FPVs). The analysis started with a phase dynamics model extended from our previous work by considering the temporal patterns in FPVs. By analyzing two-way and higher order interactions and their strength, directionality and interaction structures can be deduced from bar charts. The ability to

model high order interaction among three or more variables is one unique feature of the proposed method, which enable us to analyze complex structures.

The method was validated using four-channel van der Pol oscillators. Four important cases of interaction structures were analyzed. The analysis can be extended to network with five more FPVs. To demonstrate the importance of understanding interaction structures in semiconductor manufacturing, we investigated two polishing conditions. The results show that uncovering the interaction patterns among process variables over time brings new insights about the underlying physical mechanisms. These results also provide a new perspective for process monitoring and diagnosis.

Chapter 5

Conclusions and Future Work

5.1 Conclusions

Correlated FPVs usually contain very rich information, and investigation of the interaction among these FPVs could unveil the physical phenomenon through integrating engineering knowledge and statistical approaches. In this research, interactions of FPVs are studied to improve the manufacturing processes. The proposed methodology brings more insights for understanding the interaction mechanism during processes. The major contributions of this dissertation are summarized as follows:

- *Experimentally and theoretically study interaction patterns of FPVs.* A method based on FCCA was borrowed to extract the interaction patterns between FPVs under different process conditions and experiments were conducted to verify the relationship between the interaction patterns and varying process conditions.
- *Establish nonlinear dynamics model for complex interactions among multiple FPVs.* A novel nonlinear dynamics model was developed to describe the interaction by considering not only the complex spatial patterns among multiple FPVs' signals, but also the temporal patterns within individual signals. The proposed model formulation provides better insight into interaction mechanism, such as interaction order and strength. This improved understanding will establish

a solid base for effective detection and diagnosis of failure modes during processes.

- *Analysis of complex interaction structures among multiple FPVs* An approach with extended nonlinear dynamics model is proposed to characterize process conditions and complex interaction structure patterns are concluded accordingly for interpretation of process mechanism. This approach could bring physical understanding of processes, and provide guidance for process diagnosis.

5.2 Future Work

The purpose of this dissertation is to develop a methodology to study interaction of FPVs for process improvement. It is known that the interaction is a common phenomenon existing in many fields. Besides the semiconductor manufacturing processes, healthcare is another applicable area. For instance of Parkinson's disease, the interaction of many neurons results in the tremor activity. Yet diagnosis of medical problems with complex physiologic interactions often relies on either a trial-and-error approach or expensive medical procedures not widely available. Thus, it is highly desirable to develop generic detection and diagnosis techniques for the improved understanding of physiologic interactions and reduction of medical errors in the treatment of a wide range of medical problems. Following case show certain promise as an application of proposed methodology to diagnose disease.

To validate the approach, in vivo experiments have been conducted on adult cats to study interaction of physiologic signals under a variety of perturbations that alter breathing and cardiovascular parameters. In this experiment, process variables phrenic nerve signal and blood pressure are collected simultaneously. Figure 5.1 shows snapshots

of oscillatory phrenic nerve signal and blood pressure of an adult cat before and after introducing intermittent hypoxia, a severe external intervention (sampling frequency: 200Hz). It could not be judged that whether hypoxia occurs on this adult cat only based on these original signals.

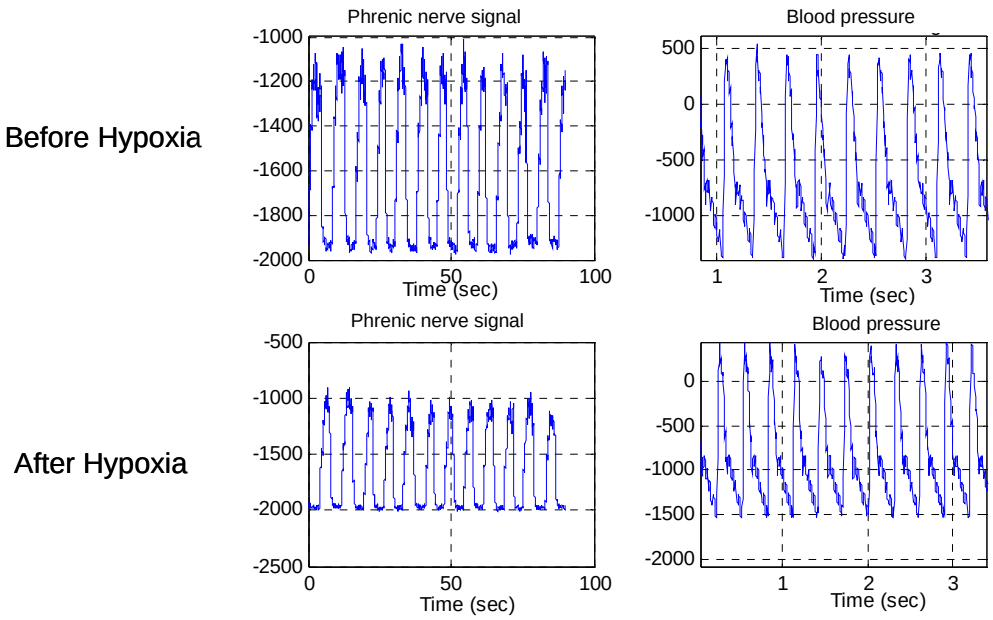


Figure 5.1 Oscillatory integrated phrenic nerve signal and blood pressure

The previously proposed nonlinear dynamics model with incorporation of GLS might be employed to detect the changes before and after hypoxia. Similarly as semiconductor manufacturing processes, through model selection, determination of interaction order and interaction strength definition, we may draw some conclusion to discover the root causes of hypoxia, and provide better guidance for doctors to reduce the medical errors.

In the future, more medical problems are expected to be diagnosed through interaction studies and our approaches. For example, we are attempting to identify obstructive sleeping apnea (OSA) from patients. Hence, studies of interaction based on nonlinear dynamics have shown promise to understand the interaction mechanisms for future complex disease diagnosis. Besides, process control in many other fields may also be attempted through our proposed methodology.

Cited References

1. H. Wang, Q. Huang, K. Morris and B. Lindsey, "A new nonlinear dynamics model to analyze synchronization patterns for physiologic condition monitoring", submitted to *IIE Transactions*, *Special issue on Healthcare Engineering*, 2007.

2. G. Moore, "Cramming more components onto integrated circuits", *Electronics*, vol. 38, No. 8, 1965.

3. K. Witt, and L. Cook, "Getting an Edge with CMP", *Semiconductor International*, pp. 70, 2000.

4. F. Preston, "The theory and design of plate glass polishing machines", *J. Soc. Glass Technol.*, vol. 11, pp. 214-256, 1927.

5. J. Luo, D.A. Dornfeld, "Material removal mechanism in chemical mechanical polishing: theory and modeling*", IEEE Trans. Semiconduct. Mfg.,* vol. 14, No. 2, pp.112-133, 2001.

6. J. Yi, "On the wafer/pad friction of chemical-mechanical planarization (CMP) processes--part I: modeling and analysis", *IEEE Trans. Semiconduct. Mfg*., vol. 18, No. 3, pp. 359-370, 2005.

7. J. Sorooshian, D. DeNardis, L. Charns, Z. Li, F. Shadman, D. Boning, D. Hetherington, and A. Philipossian, "Arrhenius characterization of ILD and copper CMP processes", *J. Electrochem. Soc.* vol.151, pp. G85-G88, 2004.

8. K. Osseo-Asare, "Surface chemical processes in chemical mechanical polishing: relationship between silica material removal rate and the point of zero charge of the abrasive material", *J. of the Electrochem. Soc.,* vol. 149, No. 12, pp. G651-G655, 2002.

9. F. Kaufman. D. Thompson, R. Broadie, M. Jaso, W. Guthrie, D. Pearson and M. Small, "Chemical-mechanical polishing for fabricating patterned with metal features as chip interconnects", *J. of the Electrochem. Soc.,* vol. 138, No. 11, pp. 3460-3465, 1991.

10. S. Mudhivarthi, P. Zantye, A. Kumar, A. Kumar, M. Beerbom and R. Schlaf, "Effect of temperature on tribological, electrochemical, and surface properties during copper CMP", *Electrochemical and Solid-State Letter,* vol. 8, pp.G241-G245, 2005.

11. J. Sorooshian, D.Hetherington and A. Philipossian, "Effect of process temperature on coefficient of friction during CMP*", Electrochem. Solid-State Letter,* vol. 7, pp. G222-G224, 2004.

12. S. Seal, S. Kuiry and B. Heinmen, “Effect of glycine and hydrogen peroxide on chemical mechanical planarization of copper”, *Thin Solid Films*, vol. 423, pp. 243-251, 2004.

13. T. Du, A. Vijayakumar, V. Desai, “Effect of hydrogen peroxide on oxidation of copper in CMP slurries containing glycine and Cu ions”, *Electrochinmica Acta.*, vol. 49, pp.4505-4512, 2004.

14. Z. Li, K. Ina, P. Lefevre, I. Koshiyama and A. Philipossian, “Determining the effects of slurry surfactant, abrasive size, and abrasive content on the tribology and kinetics of copper CMP”, *J. of The Electrochem. Soc.,* vol. 4, G299-G304, 2005.

15. Z. Li, L. Borucki, I. Koshiyama and A. Philipossian, “Effect of slurry flow rate on tribological, thermal, and removal rate attributes of copper CMP*”, J. of Electrochem. Soc.,* vol. 151, pp. G482-G487, 2004

16. H. Hocheng and Y. Huang, “In situ endpoint detection by pad temperature in chemical-mechanical polishing of copper overlay*”, IEEE Trans. Semiconduct. Mfg.,* vol. 17, No.2, pp. 180-187, 2004.

17. R. Ganesan, T. Das, A. Sikder and A. Kumar, “Wavelet-based identification of delamination defect in CMP (Cu-low k) using nonstationary acoustic emission signal”, *IEEE Trans. Semiconduct. Mfg.*, vol. 16, No. 4, pp. 677-685, 2003.

18. R. Ganesan, T. Das, A. Sikder and A. Kumar, “Online end point detection in CMP using SPRT of wavelet decomposed sensor data”, *IEEE Trans. Semiconduct. Mfg.*, vol. 18, No.3, pp. 440-447, 2005.

19. H. Wang, X. Zhang, A. Kumar, Q. Huang, “Nonlinear dynamics modeling of correlated functional process variables for condition monitoring in chemical-mechanical planarization“, *IEEE Trans. Semiconduct. Mfg.*, vol. 22, No.1, pp. 188-195, 2009.

20. J. Kim, Q. Huang, and J. Shi, “Online multi-Channel forging tonnage monitoring and fault pattern classification using principal curve”, Submitted to *ASME Transactions, J. of Mfg. Science and Engineering*, 2003.

21. E. Lada., J. Lu, and J. Wilson, “A Wavelet-based procedure for process fault detection”, *IEEE Transactions on Semiconduct. Mfg.*, vol. 15, pp. 79-90, 2003.

22. P. Castro, W. Lawton, and E. Sylvester ,“Principal modes of variation for processes with continuous sample curves”, *Technometrics*, vol. 28, pp. 329-337, 1986.

23. R. Hall, D. Poskitt and B. Presnell “A functional data-analytic approach to signal discrimination”, *Technometrics*, vol. 43, pp. 1-9, 2001.

24. J. Jin, and J. Shi, “Feature-preserving data compression of stamping tonnage information using wavelets”, *Technometrics*. vol. 41, pp. 327-339, 1999.

25. J. Jin, and J. Shi, “Automatic feature extraction for in-process diagnostic performance improvement”, *J. of Intelligent Manufacturing*, vol. 12, pp. 267-268, 2000.

26. J. Jin, and J. Shi, “Diagnostic feature extraction from stamping tonnage signals based on design of experiment”, *ASME Transactions, J. of Mfg. Science and Engineering*, vol. 122, No. 2, pp. 360-369, 2000.

27. L. Kang and S. Albin, “On-line monitoring when the process yields a linear profile”, *J. of Quality Technology*, vol. 32, pp. 418-426. 2000.

28. M. Garder, J. Lu, R. Gyurcsik, J. Wortman, B. Hornung, H. Heinisch, E. Rying, S. Rao, J. Davis and R. Mozumder, “Equipment fault detection using spatial signatures”, *IEEE Transactions on Components, Packaging, and Mfg. Tech.*, vol. 20, pp. 295-304, 1997.

29. M. Schimek, “Smoothing and regression: approaches, computation, and application”, *Wiley*, 2000.

30. J. Ramsay and B. Silverman, “Functional data analysis”, *Springer Verlag*, *New York*, 1997.

31. N. Cressie, “Statistics for spatial data, (revised edition)”, *Wiley*, 1993.

32. T. Hastie and W. Stuetzle, “Principal curves”, *J. of the American Statistical Association*, vol. 84, pp. 502-516, 1989.

33. P. Kirkwood, “On the use and interpretation of cross-correlations measurements in the mammalian central nervous system”, *J. of Neuroscience Methods*, vol. 1, pp. 107-132, 1979.

34. C. Brody, “Correlation without synchrony”, *Neural Computation*, vol. 11, pp. 1537-1551, 1999.

35. M. Jarvis and P. Mitra, “Sampling properties of the spectrum and coherency in sequences of action potentials,” *Neural Computation*, vol. 13, pp. 717-749, 2001.

36. J. Hurtado, L. Rubchinsky, and K. Sigvardt, “Statistical method for detection of phase-locking episodes in neural oscillations”, *J. of Neurophysiology*, vol. 91, pp. 1883-1898, 2004.

37. A. Pikovsky, M. Rosenblum and J. Kurths, “Synchronization: A Universal Concept in Nonlinear Sciences. Cambridge Nonlinear Science Series 12”, *Cambridge University Press, New York*, 2003.

38. P. Tass, M. Rosenblum, J. Weule, J. Kurths, A. Pikovsky, J. Volkmann, A. Schnitzler, H. Freund, "Detection of n:m phase locking from noisy data: application to magnetoencephalography", *Physical Review Letters*, vol.81, pp.3291-3294, 1998.

39. J. Rosenberg, A. Amjad, P. Breeze, D. Brillinger and D. Halliday, "The Fourier approach to the identification of functional coupling between neuronal spike trains", *Prog. Biophys. Mol. Biol.* vol. 53, pp. 1–31, 1989.

40. J. Gross, P. Tass, S. Salenius, R. Hard, H. Freund, and A. Schnitzler "Cortico-muscular synchronization during isometric muscle contraction in humans as revealed by magnetoencephalography", *J. of Physiol.,* vol. 527, pp. 623–631, 2000.

41. M. Palus and D. Hoyer "Detecting nonlinearity and phase synchronization with surrogate data", *IEEE Eng Med. Biol. Mag.* vol. 17, pp. 40–45, 1998.

42. P. Zantye, "Processing, reliability and integration issues in chemical mechanical planarization", *dissertation of doctor of philosophy, University of South Florida*, 2005.

43. J. Sorooshian, D. DeNardis, L. Charns, Z. Li, F. Shadman, D. Boning, D. Hetherington, and A. Philipossian, "Arrhenius characterization of ild and copper CMP processes", *J. of the Electrochem. Soc*., vol. 151, pp. G85-G88, 2004.

44. S. Aksu, L. Wang, and F.M. Doyle, "Effect of hydrogen peroxide on oxidation of copper in CMP slurries containing glycine", *J. of the Electrochem. Soc.,* vol. 150, pp. G718-G723, 2003.

45. J. Ramsay and B. Silverman, "Functional data analysis 2nd Ed", *Springer, New York*, 2005.

46. A. Pikovsky, M. Rosenblum, and J. Kurths, "Synchronization*:* A universal concept in nonlinear sciences"*, Cambridge University Press, New York*, 2003.

47. C. Wu and M. Hamada, "Experiments: Planning, Analysis, and Parameter Design Optimization", *Wiley, New York*, 2000.

48. P. Castro, W. Lawton, and E. Sylvester, "Principal modes of variation for processes with continuous sample curves", *Technometrics,* vol. 28, pp.329-337, 1986.

49. J. Jin and J. Shi, "Feature-preserving data compression of stamping tonnage information using wavelets", *Technometrics,* vol. 41, pp. 327-339, 1999.

50. E. Lada., J. Lu, and J. Wilson, "A wavelet-based procedure for process fault detection", *IEEE Transactions on Semiconduct. Mfg.*, vol. 15, pp.79-90, 2003.

51. R. Ganesan, T. Das, A. Sikder, and A. Kumar, "Wavelet-based identification of delamination defect in CMP (Cu-low k) using nonstationary acoustic emission signal", *IEEE Transactions on Semiconduct. Mfg.* vol. 16, pp.677- 685, 2003.

52. M. Jeong, J.- Lu, X. Huo, B. Vidakovic and D. Chen, "Wavelet-based data reduction techniques for process fault detection", *Technometrics*, vol. 48, pp. 26-40, 2006.

53. L. Kang and S. Albin, "On-line monitoring when the process yields a linear profile", *J. of Quality Technology*, vol. 32, pp. 418-426, 2000.

54. M. Mahmoud and W. Woodall, "Phase I analysis of linear profiles with calibration applications", *Technometrics*, vol. 46, pp.380-391, 2004.

55. R. Hall, D. Poskitt, and B. Presnell, "A functional data-analytic approach to signal discrimination", *Technometrics*, vol. 43, pp.1-9, 2001.

56. M. Garder, J. Lu., R. Cyurcsik, J. Wortman, B Hornung, H.Heinisch, E. Rying, S. Rao, J.Davis, and P. Mozumder, "Equipment fault detection using spatial signatures", *IEEE Transactions on Components, Packaging, and Mfg. Tech.*, vol. 20, pp. 295-304, 1997.

57. J. Kim, Q. Huang, J. Shi, T.-S. Chang, "Online multi-channel forging tonnage monitoring and fault pattern discrimination using principal curve", *ASME Transactions, J. of Mfg. Science and Engineering,* vol. 128, pp. 944-950, 2006.

58. T. Hastie, and W. Stuetzle, "Principal Curves", *J. of the American Statistical Association*, vol. 84, pp. 502-516, 1989.

59. R. Johnson and D. Wichern, "Applied Multivariate Statistical Analysis, 4th Ed.", *Prentice-Hall, New Jersey*, 1998.

60. J. Luo, D. Dornfeld, "Material removal mechanism in chemical mechanical polishing: theory and modeling", *IEEE Trans. Semiconduct. Mfg.*, vol. 14, No. 2, pp.112-133, 2001.

61. S. Bukkapatnam, P. Rao, W. Lih, N. Chandrasekaran and R. Komanduri "Process characterization and statistical analysis of oxide CMP on a silicon wafer with sparse data" , *Applied Physics A: Science & Processing*, vol. 88, No.4, pp. 785-792, 2007.

62. H. Wang, X. Zhang, A. Kumar, and Q. Huang, "Nonlinear dynamics modeling of correlated functional process variables for condition monitoring in chemical mechanical planarization", *IEEE Trans. Semiconduct. Mfg.,* vol. 22, No. 1, pp.188-195, 2009.

63. X. Zhang, H. Wang, Q. Huang, A. Kumar, J. Zhai, "Statistical and experimental analysis of correlated time-varying process variables for conditions diagnosis in chemical-mechanical planarization", *IEEE Trans. Semiconduct. Mfg.,* in press, 2009.

64. P. Kirkwood, "On the use and interpretation of cross-correlations measurements in the mammalian central nervous system", *J. Neurosci Methods* vol. 1, pp. 107-132, 1979.

65. M. Javis, P. Mitra, "Sampling properties of the spectrum and coherency in sequences of action potentials", *Neural Computation.* vol. 13, pp. 1717-749, 2001.

66. C. Brody, "Slow covariations in neural resting potential can lead to artificially fast cross-correlations in their spike trains", *J. of Neurophysiology*, vol. 80, pp. 3345-3351, 1998.

67. J. Rosenberg, A. Amjad, P. Breeze, D. Brillinger, and D. Halliday. "The Fourier approach to the identification of functional coupling between neuronal spike trains", *Pro.g Biophys. Mol. Biol.* Vol. 53*, pp.* 1–31*,* 1989.

68. J. Gross, P. Tass, S. Salenius, R. Hari, H.J. Freund, and A. Schnitzler. "Cortico-muscular synchronization during isometric muscle contraction in humans as revealed by magnetoencephalography", *J. Physiol.* vol. 527, pp. 623–631*,* 2000.

69. M. Palus and D. Hoyer. “Detecting nonlinearity and phase synchronization with surrogate data”, *IEEE Eng. Med. Biol. Mag.*vol. 17, pp. 40–45, 1998.

70. J. Hurtado, L. Rubchinsky, and K. Sigvardt, “Statistical method for detection of phase-locking episodes in neural oscillations*”, J. of Neurophysiology*, vol. 91, pp. 1883-1898, 2004.

71. B. Feige, A. Aertsen, and R. Kristeva-Feige, “Dynamic synchronization between multiple cortical motor areas and muscle activity in phasic voluntary movements”, *J. of Neurophysiology*, vol. 84, pp. 2622-2629, 2000.

72. P. Tass, M. Rosenblum, J. Weule, J., Kurths, A. Pikovsky, J. Volkmann, A. Schnitzler, and H. Freund, “Detection of n:m phase locking from noisy data: application to magnetoencephalography”, *Phys. Review Letter*, vol. 81, pp. 3291-3294, 1998.

73. M. Rosenblum and A.. Pikovsky, “Detecting direction of coupling in interacting oscillators”, *Physical Review E.*, vol. 64, pp. 0452021-0452024, 2001.

74. M. Rosenblum, L Cimponeriu, A. Bezerianos, A. Patzak and R. Mrowka, “Identification of coupling direction: application to cardiorespiratory interaction”, *Physical Review E*, vol. 65, pp. 0419091-04190911, 2002.

75. B. Schelter, M. Winterhalder, M. Eichler, M. Peifer, M. Peifer, B. Hellwig, B. Guschlbauer, C. Lucking, R. Dahlhaus and J. Timmer, "Testing for directed influences among neural signals using partial directed coherence", *J. of Neuroscience Methods*, vol. 152, pp. 210-219, 2005.

76. N. Draper, H. Smith, "Applied Regression Analysis, 3rd Ed.", *Wiley,* 1998.

77. L. Martgnon, G. Laskey, M. Diamond, W. Freiwald, E. Vaadia, "Neural coding: higher-order temporal patterns in the neurostatistics of cell assemblies", *Neural Computation*, vol. 12, pp. 2621-2653, 2000.

78. E. V. Appleton and B. van der Pol, "On a type of oscillation-hysteresis in a simple triode generator", *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science Ser.6*, vol. 43, pp. 177-193, 1922.

79. J. Sorooshian, D. DeNardis, L. Charns, Z. Li, F. Shadman, D. Boning, D. Hetherington, and A. Philipossian, "Arrhenius characterization of ILD and copper CMP processes", *J. Electrochem. Soc.* vol. 151, pp. G85-G88, 2004.

80. D.E. Lee, I. Hwang, C.M.O. Valente, J.F.G. Oliveira and D.A. Dornfeld, "Precision manufacturing process monitoring with acoustic emission", *International J. of Machine Tools and Mfg.*, vol. 46, pp. 176-188, 2006.

81. M. Moinpour, A. Tregub, A. Oehler, and K. Cadien, “Advances in characterization of CMP consumables”, *Material Research Society Bulletin*, vol. 27, pp. 766-771, 2002.

About the Author

Xi Zhang received the B.S. degree in mechanical engineering from Shanghai Jiao Tong University, China, in 2006. He is currently working towards the Ph.D. degree at the University of South Florida, Tampa. While in the Ph.D. program at the University of South Florida, Mr. Zhang focused on the research of modeling and diagnosis for process quality control on semiconductor manufacturing systems. He also has two publications in IEEE transactions on Semiconductor Manufacturing and one journal paper submitted. He also made several paper presentations at annual meetings of INFORMS. He is a member of INFORMS, IIE.