BAYESIAN PREDICTION OF FUNCTIONS WITH APPLICATIONS TO MANUFACTURING AND MARKETING

BY

MICHAEL GARY TRAVERSO

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Industrial Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2010

Urbana, Illinois

Adviser:

Associate Professor Ali Abbas

ABSTRACT

Bayesian predictive methods have a number of advantages over traditional statistical methods. For one, Bayesian methods allow one to aggregate information from multiple sources (theoretical models, experimental data, and expert opinions). In addition, a Bayesian prediction can be updated dynamically as new data becomes available. Lastly, Bayesian methods are compatible with decision analytic methods such as value of information calculations. Hence, when using a Bayesian approach, there is no guesswork in optimal (profit maximizing) design of experiments.

Bayesian methods do however have one primary drawback over traditional statistical methods and that is that they tend to be more mathematically complex. As a result, especially in complex problems, Bayesian methods see much less use than traditional statistical methods. One example of this is in the prediction of functions.

From a mathematical perspective, a function is simply a list of numbers, or in other words a vector. Thus, when predicting a function, one is really just assigning a multivariate probability distribution. What makes this problem fundamentally difficult, however, is that functions are generally defined over a continuous space. Thus, when predicting a function, one must define an uncountably infinite dimensional probability distribution.

Because of the high dimensionality, the general problem of Bayesian prediction of functions is very far from feasible. However, certain families of these probability distributions can be treated. In this work, functional probability distributions which satisfy a certain condition on their dependence structure (a chainlike dependence structure) will be considered. In the first chapter, Bayesian updating of these probability distributions will be discussed. In particular, we will show that the updated marginal distributions for any prediction which satisfies the condition

on the dependence structure can be calculated numerically using particle filtered Markov chains. In addition, analytic solutions for the updated marginals will be given for two families of functional probability distributions. In the remaining chapters, these results will be applied to real world problems in manufacturing and marketing. In chapters 2 and 3, stability limit prediction in high speed machining will be considered while in chapter 4 the focus will be on demand curve prediction.

ACKNOWLEDGEMENTS

The author gratefully acknowledges financial support for this work from the National Science Foundation (CMMI-0927051, and CMMI-0641827) as well as J.-S. Pang. Any opinions, findings, and conclusions expressed in this material are those of the authors and do not necessarily reflect the views of the NSF. The author also wishes to thank his parents, Gary and Barbara Traverso, as well as his adviser Ali Abbas.

TABLE OF CONTENTS

CHAPTER 1

MATHEMATICAL DEVELOPMENT

Bayesian predictive methods have a number of advantages over traditional statistical methods. For one, Bayesian methods allow one to aggregate information from multiple sources (theoretical models, experimental data, and expert opinions). In addition, a Bayesian prediction can be updated dynamically as new data becomes available. Lastly, Bayesian methods are compatible with decision analytic methods such as value of information calculations. Hence, when using a Bayesian approach, there is no guesswork in optimal (profit maximizing) design of experiments.

Bayesian methods do however have one primary drawback over traditional statistical methods and that is that they tend to be more mathematically complex. As a result, especially in complex problems, Bayesian methods see much less use than traditional statistical methods. One example of this is in the prediction of functions.

From a mathematical perspective, a function is simply a list of numbers, or in other words a vector. Thus, when predicting a function, one is really just assigning a multivariate probability distribution. What makes this problem fundamentally difficult, however, is that functions are generally defined over a continuous space. Thus, when predicting a function, one must define an uncountably infinite dimensional probability distribution.

Because of the high dimensionality, the general problem of Bayesian prediction of functions is very far from feasible. However, certain families of these probability distributions can be treated. In this work, functional probability distributions which satisfy a certain condition on their dependence structure (a chainlike dependence structure) will be considered. In the first chapter, Bayesian updating of these probability distributions will be discussed. In particular, we

will show that the updated marginal distributions for any prediction which satisfies the condition on the dependence structure can be calculated numerically using particle filtered Markov chains. In addition, analytic solutions for the updated marginals will be given for two families of functional probability distributions. In the remaining chapters, these results will be applied to real world problems in manufacturing and marketing. In chapters 2 and 3, stability limit prediction in high speed machining will be considered while in chapter 4 the focus will be on demand curve prediction.

1.1 MODEL ASSUMPTIONS

The model used throughout this paper will be assumed to satisfy the following:

Assumption I – Locality Property. Given X_i and X_k, X_j is irrelevant to X_h and X_l for all $h < i < j < k < l$

$$P(X_j | X_h, X_i, X_k, X_l) = P(X_j | X_i, X_k) \quad \forall (h < i < j < k < l)$$

Assumption II – Separable Measurement. All information relevant to the prediction of function can be separated into information directly relevant to the stability limit prediction at a single spindle speed.

These assumptions can be illustrated using an influence diagram (Fig. 1.1). Here, X_i denotes the dependent variable depth when the independent variable is equal to t_i and I_i denotes the information directly relevant to X_i, and I denotes the set of all I_i. Thus, the prediction for the axial $X(t_i)$ distribution $P(X_i|I)$.

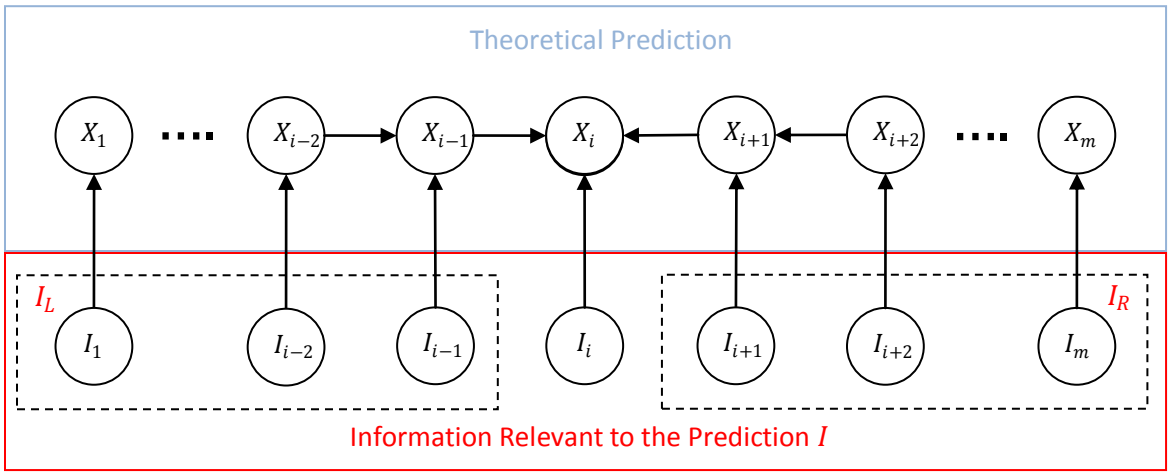

Figure 1.1: Influence diagram depicting the dependence structure assumed throughout this work.

1.2 EXPRESSING $P(X_i|I)$ IN TERMS OF EXPECTATIONS OF MARKOV CHAINS

The prediction for $X(t_i)$ conditioned on all relevant information $P(X_i|I)$ can be calculated by first applying Bayes' rule (Eq. 1.1). For brevity, normalization will be neglected throughout this chapter.

$$P(X_i|I) = P(I|X_i)P(X_i) \qquad (1.1)$$

Noting the dependence structure of Figure 1.1, this can be decomposed as

$$P(X_i|I) = P(I_i|X_i)P(I_L|X_i)P(I_R|X_i)P(X_i).$$

By again applying Bayes' rule and rearranging, $P(I_L|X_i)$ and $P(I_R|X_i)$ can be expressed as

$$P(I_L|X_i) = \frac{P(X_i|I_L)}{P(X_i)} \qquad (1.2)$$

$$P(I_R|X_i) = \frac{P(X_i|I_R)}{P(X_i)}. \qquad (1.3)$$

Inserting Eq. 1.2 and 1.3 into Eq. 1.1 and simplifying, $P(X_i|I)$ is given by

$$P(X_i|I_1, I_2, \dots, I_m) = \frac{P(I_i|X_i)P(X_i|I_L)P(X_i|I_R)}{P(X_i)}. \qquad (1.4)$$

Applying the law of total probability, $P(X_i|I_L)$ and $P(X_i|I_R)$ can be expanded as

$$P(X_i|I_L) = \int_{\mathbb{R}^{i-1}} P(X_1, X_2, .., X_i|I_L) \prod_{j=1}^{i-1} dX_j \qquad (1.5)$$

$$P(X_i|I_R) = \int_{\mathbb{R}^{m-i-1}} P(X_{i+1}, X_{i+2}, .., X_m|I_R) \prod_{k=i+1}^{m} dX_k \quad (1.6)$$

By again noticing the dependence structure of Fig. 1.1, these expressions can be simplified to

$$P(X_i|I_L) = \int_{\mathbb{R}^{i-1}} \left(P(X_1) \left[\prod_{j=1}^{i-1} P\left(X_{j+1}\middle|X_j\right) P\left(I_j\middle|X_j\right) dX_j \right] \right) \quad (1.7)$$

$$P(X_i|I_R) = \int_{\mathbb{R}^{m-i-1}} \left(P(X_m) \left[\prod_{k=i+1}^{m} P(X_{k-1}|X_k) P(I_k|X_k) dX_k \right] \right) \quad (1.8).$$

At this point, it is useful to introduce some alternate notation:

$$V_i(x_i) \equiv \log P(I_i|X_i = x_i) \ \forall i$$

$$\varphi_1(x_1) \equiv P(X_1 = x_1)$$

$$\varphi_i(x_i) \equiv P(X_i = x_i|I_L) \ \forall i$$

$$\hat{\varphi}_m(x_m) \equiv P(X_m = x_m)$$

$$\hat{\varphi}_i(x_i) \equiv P(X_i = x_i|I_R) \ \forall i.$$

Using this notation, $P(X_i|I)$ can now be expressed as

$$P(X_i = x_i|I) = e^{V_i(x_i)} \frac{\varphi_i(x_i)\hat{\varphi}_i(x_i)}{P(X_i)} \qquad (1.9)$$

$$\varphi_i(x_i) = \int_{\mathbb{R}^{i-1}} \left(\exp\left(\sum_{j'=1}^{i-1} V_{j'}(x_{j'}) \right) \left[\varphi_1(x_1) \prod_{j=1}^{i-1} P\left(X_{j+1}\middle|X_j\right) dX_j \right] \right) \qquad (1.10)$$

$$\hat{\varphi}_i(x_i) = \int_{\mathbb{R}^{m-i-1}} \left(\exp\left(\sum_{k'=i+1}^{m} V_{k'}(x_{k'}) \right) \left[\hat{\varphi}_m(x_m) \prod_{k=i+1}^{m} P(X_{k-1}|X_k) dX_k \right] \right) \quad (1.11)$$

From these equations, it follows that $\varphi_i(X_i)$ and $\hat{\varphi}_i(X_i)$ can each be expressed as the expectation of a markov chain. In particular, suppose one defines Markov chains $\{Y_t | t \in 1,2,\dots,m\}$ and $\{\hat{Y}_t | t \in m, m-1,\dots,1\}$ (reversed indexing denotes m is the initial state and 1 as the final state) such that the transition probabilities are given by

$$P(Y_{t+1}|Y_t) = P(X_{t+1}|X_t) \quad 1 \le t < m \qquad (1.12)$$

$$P(\hat{Y}_{t-1}|\hat{Y}_t) = P(X_{t-1}|X_t) \quad 1 < t \le m. \qquad (1.13)$$

Then $\varphi_i(X_i)$ and $\hat{\varphi}_i(X_i)$ can be rewritten as an expectation of the processes Y_t and $\hat{Y}_t$ respectively (Eq. 1.14- 1.15).

$$\varphi_i(x_i) = \mathbb{E}^{Y_t}\left[\varphi_1(Y_1)\exp\left(\sum_{j=1}^{i-1} V_j(Y_j)\right)\middle| Y_i = x_i\right] \qquad (1.14)$$

$$\hat{\varphi}_i(x_i) = \mathbb{E}^{\hat{Y}_t}\left[\hat{\varphi}_m(\hat{Y}_m)\exp\left(\sum_{k=i+1}^{m} V_k(\hat{Y}_k)\right)\middle| \hat{Y}_i = x_i\right] \qquad (1.15).$$

Here, the expectation is understood to be taken over all instances of the process. In other words, $\varphi_i(x_i)$ (or $\hat{\varphi}_i(x_i)$) can be found by generating a large number of instances of Y_t (or $\hat{Y}_t$), calculating $\exp\left(\sum_{k=i+1}^{m} V_k(\hat{Y}_k)\right)$ for each instance, and then summing over all instances which satisfy $Y_i = x_i$ (or $\hat{Y}_i = x_i$).

While the iterated integral representation of Eq. 1.10-1.11 cannot be extended to the continuous case (as this would require a continuum of iterated integrals), the Markov chain representation (Eq. 1.14-1.15) holds when a continuum of $X_i's$ are considered. However, a slight change of notation is required:

$$X_i \to X(t_i)$$

$$\{Y_t | t \in 1,2,\dots,m\} \to \{Y_t | t_1 \le t \le t_m\}$$

$$\{\hat{Y}_t | t \in m, m-1,\dots,1\} \to \{\hat{Y}_t | t_m \ge t \ge t_1\}$$

$$\varphi_i(x_i) \to \varphi(t_i, x)$$

$$\hat{\varphi}_i(x_i) \to \hat{\varphi}(t_i, x)$$

$$V_i(x_i) \to V(t_i, x)$$

In the continuous case, Equations 1.14-1.15 then take the form

$$\varphi(t,x) = \mathbb{E}^{Y_t}\left[\varphi\left(t, Y_{t_1}\right) \exp\left(\int_{t_1}^{t} V(\tau, Y_\tau)\, d\tau\right)\middle| Y_t = x\right] \quad (1.16)$$

$$\hat{\varphi}(t,x) = \mathbb{E}^{\hat{Y}_t}\left[\hat{\varphi}\left(t, \hat{Y}_{t_m}\right) \exp\left(\int_{t_m}^{t} V\left(\tau, \hat{Y}_\tau\right) d\tau\right)\middle| \hat{Y}_t = x\right] \quad (1.17)$$

1.3 ANALYTIC SOLUTIONS FOR $\varphi(t,x)$ AND $\hat{\varphi}(t,x)$ FOR ITŌ PROCESSES

The advantages of expressing $\varphi(t,x)$ and $\hat{\varphi}(t,x)$ as expectations of Markov chains are at least twofold. First, regardless of the transition probabilities, the expectations can be approximated using either Markov chain Monte Carlo methods or particle filtered Markov chains. Secondly, for certain classes of transition probabilities, equations (1.16) and (1.17) have known analytic solutions.

One such example is when the transition probabilities are defined such that Y_t and $\hat{Y}_t$ are Itō processes (Fig. 1.2). Itō processes are essentially the continuous version of random walks (Fig 1.2) and if and only if $\{Y_t | t_1 \le t \le t_m\}$ is an Itō process, it can be expressed in the form

$$Y_t = Y_{t_1} + \int_{t_1}^{t} \mu(\tau, Y_\tau)\, d\tau + \int_{t_1}^{t} \sigma(\tau, Y_\tau)\, dB_\tau. \quad (1.18)$$

From equations (1.12-1.13), it follows that for $\{Y_t | t_1 \le t \le t_m\}$ defined by Eq. (1.18), $\left\{\hat{Y}_t \middle| t_m \ge t \ge t_1\right\}$ must take the form

$$\hat{Y}_t = \hat{Y}_{t_m} - \int_{t_m}^{t} \mu\left(\tau, \hat{Y}_\tau\right) d\tau + \int_{t_m}^{t} \sigma\left(\tau, \hat{Y}_\tau\right) dB_\tau. \quad (1.19)$$

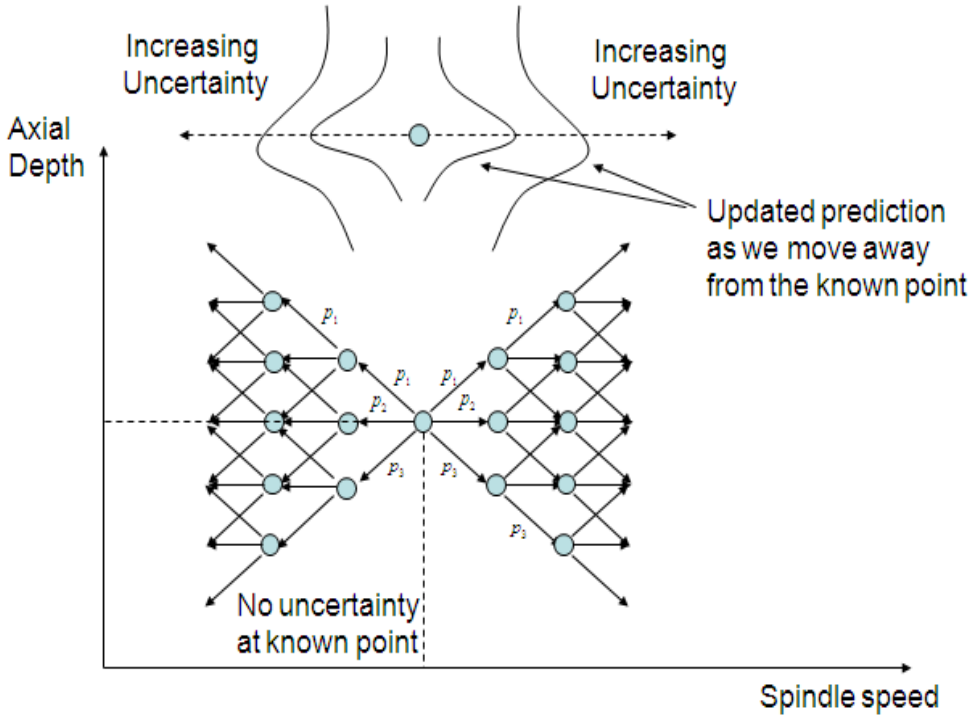

Figure 1.2: Discretized representation of an Itō process (random walk). At each step, the process moves one step in the +x direction with probability p_1, doesn't move at all with probability p_2, or moves one step in the –x direction with probability p_3=1- p_1- p_2. The drift and diffusion fields are determined by the choice for p_1, p_2 and p_3. I would like to thank Ali Abbas for allowing me to use this figure.

In these equation, B_t denotes a standard Brownian motion (or wiener process) and an integral with respect to B_t is understood in the stochastic sense (as an Itō integral). The field $\mu(t,x)$ is referred to as the drift field and is formally related to Y_t by the expression

$$\mu(t,x) = \lim_{\varepsilon \downarrow 0} \left\{ \frac{\mathbb{E}^x[Y_{t+\varepsilon} - Y_t | Y_t = x]}{\varepsilon} \right\}. \quad (1.20)$$

Informally, $\mu(t,x)$ can be thought of as the expected slope of $X(t)$ given that $X(t)=x$. The field $\sigma(t,x)$ is referred to as the diffusion field and is formally related to Y_t by the expression

$$\sigma(t,x)=\sqrt{\lim_{\varepsilon\downarrow 0}\left\{\frac{\mathbb{E}^x[(Y_{t+\varepsilon}-Y_t)^2|Y_t=x]}{\varepsilon^2}\right\}}. \qquad (1.21)$$

Informally, $\sigma(t,x)$ can be thought of as the standard deviation of the slope of $X(t)$ given that $X(t)=x$. Thus, when $\sigma(t,x)=0$, there is no uncertainty in the slope of $X(t)$ and when $\sigma(t,x)$ is large there is a lot of uncertainty in the slope of $X(t)$.

When the transition probabilities are defined such that Y_t is an Itō process, the expectations in Eq. 1.18 and Eq. 1.19 can be evaluated analytically by applying the Feynman-Kac formula. The Feynman-Kac formula states that for an Itō process defined by drift field $\mu(t,x)$ and diffusion field $\sigma(t,x)$, $\varphi(t,x)$ as defined by Eq. 1.16 is given by the solution to the initial value problem

$$\left[\frac{\partial}{\partial t}\right]\varphi(t,x)=\left[\frac{1}{2}\sigma(t,x)^2\left(\frac{\partial}{\partial x}\right)^2+\mu(t,x)\frac{\partial}{\partial x}+V(t,x)\right]\varphi(t,x) \qquad (1.22)$$

$$\varphi(t_1,x)\equiv P(X(t_1)=x).$$

Similarly, $\hat{\varphi}(t,x)$ as defined by Eq. 1.17 is given by the solution to the initial value problem:

$$\left[-\frac{\partial}{\partial t}\right]\hat{\varphi}(t,x)=\left[\frac{1}{2}\sigma(t,x)^2\left(\frac{\partial}{\partial x}\right)^2-\mu(t,x)\frac{\partial}{\partial x}+V(t,x)\right]\hat{\varphi}(t,x) \qquad (1.23)$$

$$\hat{\varphi}(t_m,x)\equiv P(X(t_m)=x).$$

1.4 ANALYTIC SOLUTIONS FOR $\varphi(t,x)$ AND $\hat{\varphi}(t,x)$ FOR RELATIVISTIC ITŌ PROCESSES

Another example yielding analytic solutions is when the transition probabilities are defined such that Y_t and $\hat{Y}_t$ are given by what will be referred to as relativistic Itō processes. Relativistic Itō processes differ from regular Itō processes in that a relativistic Itō process Y_t satisfies the constraint

$$\left|\frac{Y_{t+\delta} - Y_t}{\delta}\right| < c \quad \forall \delta \, .$$

In other words, the slope of all sample paths of a relativistic Itō process are bounded by some fixed c. Thus, when predicting a function $X(t)$ which is known to satisfy some constraint on the slope, such as monotoncity, a relativistic Itō process can be used to enforce this constraint while a standard Itō process cannot.

Formally, the construction of a relativistic Itō process differs from the construction of a standard Itō process in how distances and angles in the $\{t,x\}$ plane are treated. For a standard Itō process, the plane is treated as a Euclidean (classical) space-time as for a relativistic Itō process, the plane is treated as Minkowskian (relativistic) space-time. For the purposes of taking the expectation of the process, the difference between the two is that for a Minkowskian space-time, the Laplacian operator must be changed to the D'Alembertian operator:

$$\left(\frac{\partial}{\partial x}\right)^2 \rightarrow \left(\frac{\partial}{\partial x}\right)^2 - \left(\frac{1}{c}\frac{\partial}{\partial t}\right)^2 .$$

Thus, when the transition probabilities are defined such that Y_t and $\hat{Y}_t$ are given by relativistic Itō processes, Eq. 1.18 and 1.19 are given by the initial value problems

$$\left[\frac{\partial}{\partial t}\right]\varphi(t,x)=\left[\frac{1}{2}\sigma(t,x)^2\left(\left(\frac{\partial}{\partial x}\right)^2-\left(\frac{1}{c}\frac{\partial}{\partial t}\right)^2\right)+\mu(t,x)\frac{\partial}{\partial x}+V(t,x)\right]\varphi(t,x) \qquad (1.24)$$

$$\varphi(t_1,x)\equiv P(X(t_1)=x).$$

$$\left[-\frac{\partial}{\partial t}\right]\hat{\varphi}(t,x)=\left[\frac{1}{2}\sigma(t,x)^2\left(\left(\frac{\partial}{\partial x}\right)^2-\left(\frac{1}{c}\frac{\partial}{\partial t}\right)^2\right)-\mu(t,x)\frac{\partial}{\partial x}+V(t,x)\right]\hat{\varphi}(t,x) \qquad (1.25)$$

$$\hat{\varphi}(t_m,x)\equiv P(X(t_m)=x).$$

1.5 REVIEW OF MATHEMATICAL DEVELOPMENT

To review, for any problem satisfying assumptions I and II, the prediction for X_i (or $X(t_i)$ in the continuous case) conditioned on all relevant information $P(X_i|I)$ ($P(X(t_i)|I)$) is given by equation 1.9 where the terms $\varphi_i(x_i)$ and $\hat{\varphi}_i(x_i)$ ($\varphi(t,x)$ and $\hat{\varphi}(t,x)$) can each be written as the expectation of a Markov process (Eq. 1.14-1.15 (1.16-1.17)). Regardless of the specific choice of transition probabilities, these expectations can be approximated using Markov chain Monte Carlo methods or particle filtered Markov chains. In addition, when that the transition probabilities are defined such that the resultant Markov processes are Itō processes or relativistic Itō processes, the expectations can be evaluated analytically as solutions to initial value problems (Eq. 1.22-1.25).

REFERENCES

[1.1.] Kleinart, H., 2004, Path Integrals in Quantum Mechanics, Statistics, Polymer Physics, and Financial Markets, 4th edition, World Scientific, Singapore.

[1.2.] Taira, K., 1988, Diffusion Processes and Partial Differential Equations, Academic Press Inc., Boston, MA.

[1.3.] Nagasawa, M., 1993, Schrödinger Equations and Diffusion Theory, Basel, Birkhäuser Verlag.

CHAPTER 2

OPTIMAL EXPERIMENTATION FOR SELECTING STABLE MILLING PARAMETERS: A BAYESIAN APPROACH

In the remaining chapters, the focus will shift from mathematical development of the Bayesian model to how the model can be applied to real world problems. In this chapter, the problem of stability limit prediction in high speed machining will be introduced. Primarily, this chapter will be concerned with illustrating the advantages of Bayesian prediction of functions and how one could actually go about applying the methods in a relatively simple setting. In particular, this chapter will first cover how a simple experiment (performing a test cut and manually inspecting it to determine whether the test cut was stable or unstable) can be used to update one's prediction of the stability limit. Next, the concept of *value of information* will be introduced and how it can used to determine an optimal sequence of experiments to perform will be discussed. Lastly, a simulated case study is presented.

2.1 INTRODUCTION

Both physically and economically motivated uncertainties are often present in manufacturing decision situations. To deal with these uncertainties, predictive models can be used. Most methods for predicting functions are statistically based, including regressions and curve fits. However, these methods may be inadequate for the prediction of complex functions, such as those that arise in machining dynamics, since there is no rigorous method of aggregating information from multiple sources (experimental data, simulation data, and theoretical predictions) and they do not quantify the value of information gathering activities.

Although Bayesian methods for predicting functions can be mathematically complex, they have two major advantages over other approaches. First, any information known (or learned through experiment) about the function can be incorporated. This makes these methods more versatile than statistical methods, which are most compatible with simple sets of observables. Second, Bayesian methods enable information to be valued based on its influence on profit, which makes it possible to systematically choose an information gathering scheme that maximizes profit.

Since Bayesian methods have not been widely applied in manufacturing, the focus of this work is to explore their use in this field. To aid in the explanation, stability limit prediction is examined. In Section 2.2, general Bayesian probability assignment is discussed. In Section 2.3, Brownian distributions are introduced. In Section 2.4, a case study involving stability limit prediction is presented. In this case study, the stability limit is modeled using Brownian distributions. These predictions are then used to approximate an optimal experimental design. Section 2.5 introduces the basics of building complex correlative structures from an underlying dynamical model.

2.2 BAYESIAN PREDICTION OF FUNCTIONS

These methods are fundamentally different from curve-fitting methods in that they treat a set of predictions rather than a single prediction. Methods which use a single curve as a prediction only answer a single question about an unknown function $g(x)$:

Given an arbitrary knowledge state, what is the most likely g(x)?

In contrast, Bayesian methods assign a probability to every function. In other words, for any $g(x)$ the following question can be answered:

Given an arbitrary knowledge state, how likely is it that an arbitrary function = g(x)?

The advantage of this approach is that arbitrary information regarding the function of interest can be incorporated into the predictive model. For example, abstract functional assumptions that can't be handled by statistical methods, such as monotonicity, differentiability class, or periodicity, can be incorporated. In addition, multiple informative statements can be aggregated in a mathematically rigorous manner.

Mathematically, probability assignment over a functional space is equivalent to assigning a multivariate distribution over infinite dimensions. This assignment can be simplified by assigning a distribution only over the set of all *measurable quantities.* The most general set of measurable quantities is the set of all functionals, i.e., any operator which takes a function as an input and produces a scalar, such as a definite integral. However, in most situations, this can be expressed in terms of a finite number of dimensions. The set of all observables is denoted $\vec{O}$ and individual observable is denoted as O_i.

Probability density functional (pdf) $f_{\vec{O}}(g(x))$

For an indexed set of *n* observables $\vec{O}$ and any values of o_i, $1<i<n$, the probability density functional answers the question:

What is the probability that $\{O_1, O_2, \ldots O_n\} = \{o_1, o_2, \ldots o_n\}$*?*

Marginal probability distribution $f_{O_j}(o_j)$

The marginal probability distribution answers the question:

Given no information about any other observables, what is the probability that observable J = j?

In terms of the pdf, it is given by:

$$f_{O_j}(o_j) = \int_{R^{n-1}} f_{\vec{O}}(g(x)) \prod_{i \leq n, i \neq j} do_i \qquad (2.1)$$

Marginal distributions of multiple observables are defined similarly.

Conditional distributions $f_{O_1 \ldots O_m | O_{m+1} \ldots O_n}(o_1, .. o_m)$

A conditional probability distribution answers the question (note that the O_i can be indexed arbitrarily):

Given that $\{O_1, O_2, \ldots, O_m\} = \{o_1, o_2 \ldots o_m\}$, what is the probability that $\{O_{m+1}, O_{m+2}, \ldots, O_n\} = \{o_{m+1}, o_{m+2} \ldots o_n\}$?

Mathematically, the conditional probability distribution is given by the pdf divided by a marginal distribution:

$$f_{O_1 \ldots O_m | O_{m+1} \ldots O_n}(o_1, .. o_m) = \frac{f_{\vec{O}}(g(x))}{f_{O_{m+1}, \ldots, O_n}(o_{m+1})} \qquad (2.2)$$

2.3 BROWNIAN DISTRIBUTIONS

Suppose that the set of observables is chosen as the values of $g(x)$ for all values of x. Furthermore, assume that the value of $g(x)$ for an arbitrary x depends only on the value of g in a small neighborhood of x. In terms of the conditional probabilities, this is given by:

$$\text{for any } x_1 < x_2 < x_3 < x_4 < x_5$$

$$f_{g(x_3)|g(x_2),g(x_4)} = f_{g(x_3)|g(x_1),g(x_2),g(x_4),g(x_5)} \qquad (2.3)$$

With these assumptions, $g(x)$ can instead be represented in terms of its derivative with respect to x, $\dot{g}(x,g)$ [2.1]. The advantage of this representation is that by Eq. (2.3) all $\dot{g}(x,g)$ are independent of each other. If it is assumed that $\dot{g}(x,g)$ are generated from distributions in the exponential family, the pdf can be written in terms of a definite integral. If $\dot{g}(x,g)$ is assumed to be normally distributed for every $\{x,g\}$, the probability density functional takes the form:

$$f_{g(x)}(g(x)) = exp\left(- \int_{-\infty}^{\infty} k(x,g)\big(\dot{g}(x,g) - \mu(x,g)\big)^2 \quad + \quad c(x,g)\, dx \right) \quad (2.4)$$

where $k(x,g)$ is the diffusion field, $\mu(x,g)$ is the drift field, and $c(x,g)$ is the creation and killing field.

Distributions which take this form are referred to as Brownian distributions. Brownian distributions are ideal for predicting and updating when little is known about the function in question. Provided that any information gained through experiment can be expressed in terms of g or its slope at each point, any distribution updated from a Brownian distribution will also be a Brownian distribution. If the experiment measures g directly, the creation and killing field is updated. If an experiment measures $\dot{g}$, the drift field is updated. The diffusion field may be updated in either situation and can be thought of as a measure of uncertainty at $\{g,x\}$.

2.4 STABILITY LIMIT PREDICTION USING BROWNIAN DISTRIBUTIONS

To illustrate how Bayesian methods can be used in a manufacturing application, consider the prediction of the speed-dependent stability limit in milling. The stability limit has a particularly complex structure, since it generally has multiple non-differentiable points (Fig. 2.1). For the numerical example provided here, the true stability limit is generated using the algorithm described in [2.2]. The input parameters used for the algorithm are summarized in Table A.1 of Appendix A.

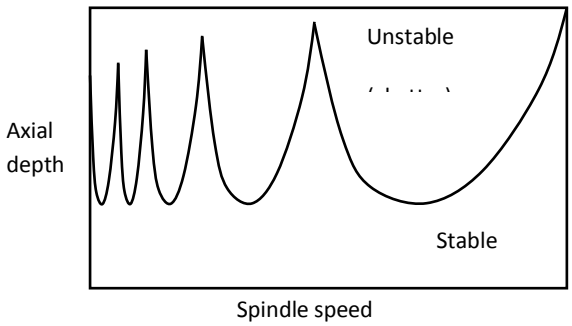

Figure 2.1: Example stability lobe diagram.

For this study, a Brownian distribution is used to describe the initial knowledge state (Table A.2 gives the parameter fields). In addition, it is assumed that the stability limit is known to occur below an axial depth g_{max} known *a priori*. While this may be an unrealistic assumption in practice, the bound provides an intuitive link between the imposed constraint and $c(x,g)$.

Here, new knowledge is gathered using only simple stability testing. In stability testing, a machinist selects an axial depth-spindle speed combination, performs a test cut, and examines the cut to determine whether the parameters provided stable or unstable operation. For this numerical evaluation, the "testing" is simulated by comparing the selected test parameters$\{x_i^t, g_i^t\}$ to the predefined/reference stability limit (via the stability lobe diagram generated using the parameters listed in Table A.1).

The knowledge gained from a simple stability test can be expressed as an inequality constraint on $g(x)$. If a test is stable, the stability limit must occur above that point; see Eq. (2.5a). If a test is unstable, the stability limit must occur below that point; see Eq. (2.5b).

$$g(x_i^t) < g_i^t \; if \; test \; i \; stable \qquad (2.5a)$$

$$g(x_i^t) > g_i^t \; if \; test \; i \; unstable \qquad (2.5b)$$

Since simple stability testing updates knowledge of $g(x)$ directly, the probability density functional for any updated knowledge state can be expressed by updating the creation and killing field. The constraints of Eqs. (2.5a) and (2.5b) can be imposed by "killing" all $g(x)$ which do not satisfy these constraints. For an arbitrary number of simple stability tests, the updated creation and killing field is given by Eq. (2.6).

$$c_{up}(x,g) = \begin{cases} \infty\ for\ g < 0 \\ \infty\ for\ g > g_{max} \\ \infty\ for\ g(x_i^t) < g_i^t\ if\ test\ i\ stable \\ \infty\ for\ g(x_i^t) > g_i^t, if\ test\ i\ unstable \\ 0\ otherwise \end{cases} \qquad (2.6)$$

In order to gauge how performing an experiment will affect profit, the decisions that are affected by knowledge of the stability limit should be considered. In this study, the decision situation consists of selecting an axial depth-spindle speed combination $\{x_{op}, g_{op}\}$ at which to mill away a cube of material (see Table A.3 in Appendix A for additional parameters). Profit for stable operation, $P(x_{op},g_{op})$, is calculated using the approach provided in [2.3]. Profit when the operating parameters are unstable is assumed to be zero here, although the effects of chatter can generally be removed by additional finishing operations in practice.

For a fixed knowledge state, the probability of stability for an arbitrary spindle speed-axial depth combination $P_{stab}(x_{op}, g_{op})$ can be expressed in terms of a marginal distribution; see Eq. (2.7).

$$P_{stab}(x_{op}, g_{op}) = \int_{x_{op}}^{\infty} f_{g_{op}}(\varepsilon) d\,\varepsilon \qquad (2.7)$$

For an arbitrary test history, defined in Eq. (2.8), the expected profit of a production run, $P_{exp}(x_{op}, g_{op})$, can be calculated using the probability tree shown in Fig. 2.2. The resultant expression for $P_{exp}(x_{op}, g_{op})$ is given in Eq. (2.9).

$$T_i \coloneqq \{x_i^t, g_i^t, R_i^t\} \quad 1 < i < \#\ of\ tests \qquad (2.8)$$

$$P_{exp}(x_{op}, g_{op}; T) = P_{stab}\{x_{op}, g_{op}; T\} P(x_{op}, g_{op}) \qquad (2.9)$$

$$where\ R_i^t\ denotes\ the\ result\ of\ stability\ test\ i$$

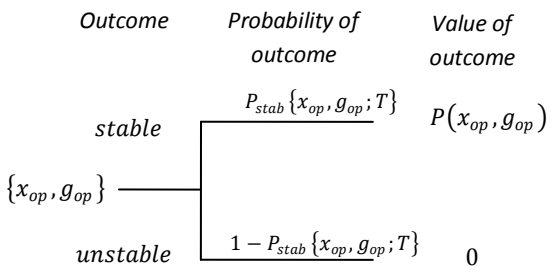

Figure 2.2: Probability tree for calculating the expected profit of a production run $\mathbf{P_{exp}(x_{op}, g_{op}; T)}$.

Assuming that the operating conditions are chosen to maximize the expected profit, the value of performing the production run given an arbitrary test history, $V_T\{T\}$, is defined using Eq. (2.10).

$$V_T\{T\} := \sup_{x_{op}, g_{op}} \{P_{exp}(x_{op}, g_{op}; T)\} \tag{2.10}$$

For a stability test at an arbitrary $\{x^t, g^t\}$, the probability of each outcome (Eq. (2.7)), as well as the expected value of each outcome (Eqs. (2.9) and (2.10)) can be calculated before actually knowing the outcome of the test. Using the probability tree shown in Fig. 2.3, the expected value of the production run after the experiment's result may be calculated. Subtracting the value of the production run prior to the result of the experiment being known yields $V_E\{x^t, g^t\}$, the expected value gained from performing that experiment; see Eq. (2.11).

$$V_E\{x^t, g^t\} = \left(1 - P_{stab}\{x^t, g^t; T^{initial}\}\right) V_T\{T^{unstable}\} + P_{stab}\{x^t, g^t; T^{initial}\} V_T\{T^{stable}\} - V_T\{T^{initial}\} \tag{2.11}$$

where $T^{initial}$ is the current knowledge state, T^{stable} denotes $T^{initial}$ augmented with an additional row vector $\{x^t, g^t, stable\}$, and $T^{unstable}$ denotes $T^{initial}$ augmented with an additional row vector $\{x^t, g^t, unstable\}$.

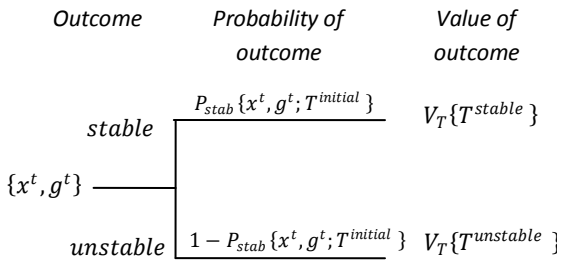

Figure 2.3: Probability tree for calculating the value of performing an experiment, $\mathbf{V_E}\{\mathbf{x^t}, \mathbf{g^t}\}$.

Using larger probability trees, the value gained through an arbitrary number of experiments can be calculated. As a result, it is possible to assign a value to any arbitrary sequence of experiments, which can then be maximized to determine the optimal testing policy.

Since stability tests can be chosen sequentially (i.e., the results of the first test is known when choosing the testing parameters for the second test), the number of test parameters grows exponentially with the number of tests in the sequence (because the parameters for the second test depend on the result of the first test). To reduce the computational complexity of this optimization, a greedy heuristic was used. In particular, test *i* was chosen by maximizing $V_E\{x_i^t, g_i^t\}$ (the tests are treated independently, instead of together as a sequence). Using this heuristic, a sequence of twelve experiments were chosen and tested against the reference stability lobe diagram. Figure 2.4 shows the progression of $P_{stab}\{x, g; T\}$ (the probability that $\{x,g\}$ is stable), as the results of a sequence of experiments are incorporated into the prediction.

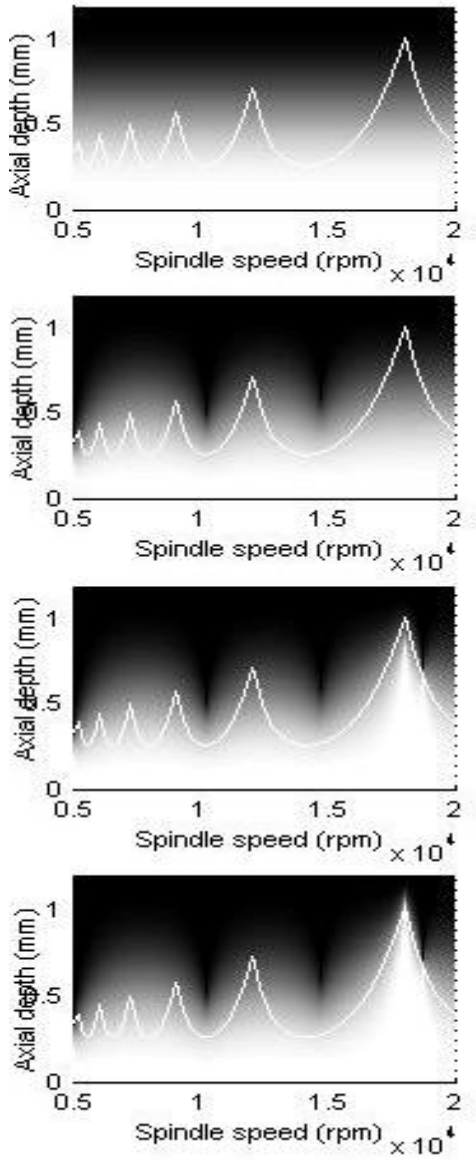

Figure 2.4: Progression of $P_{stab}\{x, g; T\}$ as experimental data is gathered. The top panel shows the prediction prior to any testing. From top to bottom, the remaining panels show the updated predictions after 4, 8, and 12 tests, respectively.

2.5 BUILDING GENERAL DISTRIBUTIONS BY TRANSFORMING A SET OF OBSERVABLES

While Brownian distributions offer the advantage of simplicity, in many cases it is advantageous to describe the function of interest using an underlying model. For milling dynamics, the stability limit can be inferred from knowledge of the frequency response function (FRF) as reflected at the tool point and a cutting force model, which relates the force to prescribed chip area [2.4]. As a result, the stability prediction problem can be treated using the FRF and force model as the set of observables. Since the relevant information can be expressed in terms of a limited number of parameters (cutting force model coefficients; and modal stiffness, damping, and natural frequency values for the tool point FRF), a distribution of possible stability limits can be obtained by transforming a distribution assigned over these parameters, instead of assigning the distribution in terms of the stability limit directly.

The advantage of this treatment is two-fold. First, a complex correlative structure is implied directly from the underlying dynamical model. Second, experiments which measure the inputs (such as using a dynamometer to establish the force model or an impact test to identify the tool point FRF) can be incorporated into optimal experimental design. This provides an important alternative to simple stability testing, since a single FRF measurement, for example, enables much more accurate predictions to be made (see Fig. 2.5).

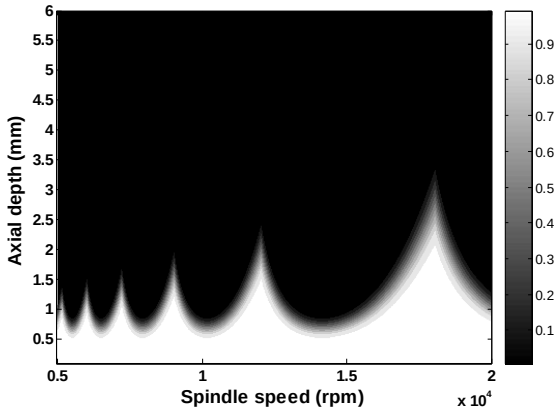

Figure 2.5: Contour plot of the probability of stability calculated assuming uncertain force model coefficients [2.3].

While the Brownian correlative structure works best when little is known about the function of interest, a correlative structure built from an underlying model enables predictions to be made when relationships governing the underlying dynamics are known. Current work is still explorative, but future work will include developing a rigorous framework for an arbitrary underlying model. One goal, in particular, is to incorporate an underlying dynamical model which is itself probabilistic, instead of a deterministic model with probabilistic inputs.

2.6 CONCLUSION

As efficiency and productivity demands on the manufacturing industry increase, the need for a systematic approach to information gathering is growing. Due to the inherent complexity involved in many manufacturing decision problems, a general treatment of optimized experimentation for an arbitrary pay-off and knowledge state is typically beyond the scope of traditional statistical methods. However, Bayesian methods are a natural candidate for this task. Not only can they treat arbitrarily complex mathematical objects, including spaces of functions, but they can also associate a value with each knowledge state. As a result, information gathering can be optimized in a manner similar to production planning or process control.

Since Bayesian methods have not previously been widely applied in manufacturing decision making, the primary motivation for this work was to explore possibility of using Bayesian methods to predict functions arising from the complex dynamics involved in milling. First, Brownian distributions were used to explore how Bayesian methods could be used in cases where little is known about the function of interest. Second, the process of incorporating underlying dynamical models into Bayesian prediction was explored. While these methods are still in their infancy, they offer sufficient analysis capability for a systematic treatment of the prediction of complex milling dynamics.

REFERENCES

[2.1] Nagasawa, M., 1993, Schrödinger Equations and Diffusion Theory, Basel, Birkhäuser Verlag.

[2.2] Altintas, Y. and Budak, E., 1995, Analytical Prediction of Stability Lobes in Milling. Annals of the CIRP. 44/1: 357-362.

[2.3] Zapata, R., Schmitz, T., Traverso, M., and Abbas, A., 2009, Value of Information and Experimentation in Milling Profit Optimization, International Journal of Mechatronics and Manufacturing Systems (in press).

[2.4] Altintas, Y., 2000, Manufacturing Automation Metal Cutting Mechanics, Machine Tool Vibrations, and CNC Design, Cambridge University Press.

CHAPTER 3

A SEQUENTIAL GREEDY SEARCH ALGORITHM WITH BAYESIAN UPDATING FOR TESTING IN HIGH SPEED MILLING OPERATIONS

In this chapter, the problem of stability limit prediction will again be considered. While the previous chapter used stability limit prediction primarily as a tool to illustrate how Bayesian methods can be used, in this chapter the problem is considered much more seriously through a realistic case study. As a result, a good portion of this chapter will be devoted to describing the particular tooling setup, how the tests were physically performed, and practical considerations. Results will then be presented for a number of simulated and real world testing scenarios. Finally these results and the overall effectiveness of the model will be critiqued.

3.1 INTRODUCTION

In the past several decades, high-speed machining has become a staple in discrete part production. Improved high-speed machining technology has made increased spindle speeds and axial depths of cut possible. The availability of these new parameter choices is due in large part to the increased stability at tooth passing frequencies which are substantial integer fractions of the dominant system natural frequency. However, regenerative chatter can still occur. Early work in describing chatter as a self-excited vibration was completed by Tobias, Arnold, Tlusty, and Merrit [3.1-3.8] and more recently Altintas and Budak [3.9-3.10], for example.

Although methods are available for predicting the spindle speed-dependent stability limit, the requirement for knowledge of the tool point frequency response function (for each tool-holder-spindle-machine combination) can impose a significant obstacle in some production facilities. The purpose of this study is to characterize the uncertainty about the stability limit using a

probability distribution, rather than a deterministic boundary. The characterization is based on using initial belief about the stability limit, and updating belief using Bayesian analysis through limited experiments. This work builds on recent efforts for applying decision analysis to milling [3.11-3.12]. Concepts in Bayesian inference are provided in the following section.

3.2 BAYESIAN INFERENCE

Regardless of a model's rigor, uncertainty due to factors that are unknown or not included and imperfect knowledge of the input data always exists. This uncertainty may be reduced by experimentation. However, this experimentation may be costly and is not, in general, a realistic substitute for the model. Therefore, a predictive model should address theoretical considerations, incorporate uncertainty, and update uncertainty as new information is made available (from experiments, for example). In this work, the stability limit for milling is treated as the unknown quantity with inherent uncertainty.

Bayesian inference models, which form a normative and rational method for updating belief, are applied here. In the case of the frequency-domain stability algorithm [3.9] used in this work, a single (deterministic) prediction for the stability limit is made. A Bayesian model, on the other hand, assigns a probability distribution over the set of all possible stability limits and updates this probability distribution with new information (experimental results).

Let the prior distribution about an uncertain event, A, at a state of information, &, be $\{A|\&\}$, the likelihood of obtaining an experimental result B given that event A occurred be $\{B|A,\&\}$, and the probability of receiving experimental result B (without knowing A has occurred) be $\{B|\&\}$. Bayes' rule determines the posterior belief about event A after observing the experiment results, $\{A|B,\&\}$ as show in Eq. 3.1. Using Bayes' rule, information gained through experimentation can

be incorporated with the prior prediction about the stability limit to obtain a posterior distribution.

$$\{A|B,\&\} = \frac{\{A|\&\}\{B|A,\&\}}{\{B|\&\}} \qquad (3.1)$$

An important requirement for applying Bayes' rule in this case is selecting the initial belief (prior distribution) for the stability limit. In general, this initial prediction: 1) can be constructed from any combination of theoretical considerations, previous experimental results, and expert opinions; and 2) should be chosen to be as informative as possible regarding experimenter's belief. In this work, a worst case scenario is considered where inference of the stability limit must be completed without knowledge of machining dynamics. However, the general understanding that it is more likely to obtain an unstable result at increased axial depths of cut is applied.

Bayesian inference models offer several advantages. When using a Bayesian inference model, experiments can be chosen such that the expected value added by performing the experiment is maximized. This enables the best selection of experiments to be performed. Although the mathematics is complicated, the basic idea is presented as follows. Prior to performing a test, the following information is known: 1) the probability of each outcome of a test (this can be determined from the prior distribution); and 2) the updated prediction given each possible test outcome. For example, in the case of stability limit prediction, prior to performing a test it can be determined how likely it is that the test will give a stable result versus an unstable result, as well as what parameters should be selected given a stable versus unstable result (this provides the value of performing a cut given a stable versus unstable result.) Using this reasoning, the expected value added by performing a stability test at a given spindle speed-axial depth combination can be determined. It can then be decided whether to perform the stability test

using the spindle speed-axial depth combination which maximizes the value added; note that this is the most profitable experiment to perform.

3.3 BAYESIAN INFERENCE CASE STUDY

To illustrate how the Bayesian inference model can be applied to milling, a case study is presented. Two scenarios are considered: 1) in simulated stability testing, the result of a test is determined by comparing the test point to a reference stability limit generated using the Altintas and Budak [3.9] algorithm; and 2) stability is determined through experiments on a high-speed machining center (Mikron UCP 600 Vario). While the second scenario is more realistic, the disadvantage is that, since the true stability limit is unknown, the model performance is difficult to quantify. Therefore, the first scenario is provided to qualitatively examine the model's capability. In both cases, the same tool and workpiece combination are used. See Table 3.1.

Prior to performing the stability tests, the force model coefficients and the tool point frequency response function (FRF) were measured. The force model coefficients for the 6061-T6 workpiece material-tool combination were calculated using a linear regression to the mean values of X (feed) and Y direction cutting forces measured at a range of feed per tooth values [3.13-3.14]. The tool point FRF was also measured using impact testing; see Fig. 3.1.

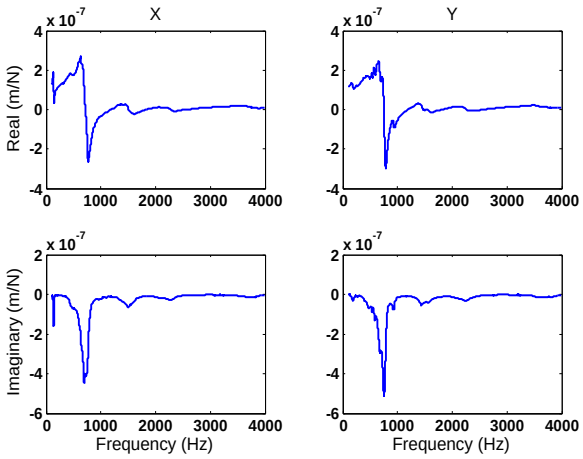

Figure 3.1: Tool point FRF for experiments. I would like to thank Raúl Zapata for allowing me to use this figure.

For the simulated testing scenario, the following conditions were applied. First, a large spindle speed domain was selected (5000 rpm to 30000 rpm); this domain was limited to the tool manufacturer's recommendations for the actual testing (4000 rpm to 7250 rpm). In addition, a symmetric, single degree-of-freedom FRF was used. The stiffness was 20 MN/m, the viscous damping coefficient was 0.05, and the natural frequency was 2667 Hz. The additional parameters used to generate the reference stability limit (shown in Fig. 3.2) are summarized in Table 3.1.

Table 3.1: Parameters used to determine the reference stability limit for the simulated testing scenario.

Parameter	Value	Units
Radial depth	6.35	mm
Feed per tooth	0.15	mm/tooth
Tool radius	6.35	mm
Number of teeth	4	teeth
Helix angle	30	deg
Tangential coefficient	613	N/mm^2
Normal coefficient	149	N/mm^2
Tangential edge coefficient	7.0	N/mm
Normal edge coefficient	6.0	N/mm

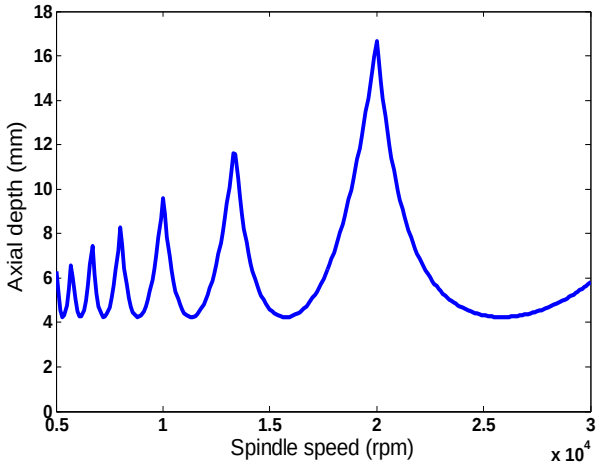

Figure 3.2: Reference stability limit for simulated testing scenario. I would like to thank Raúl Zapata for allowing me to use this figure.

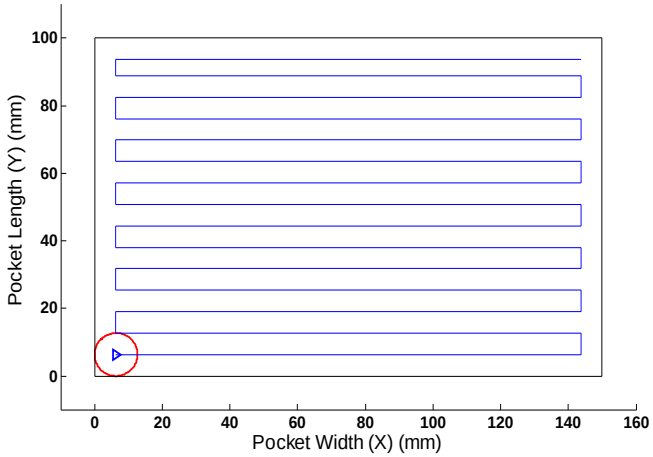

Figure 3.3: Tool path for pocket milling. I would like to thank Raúl Zapata for allowing me to use this figure.

Because the sequence of tests to be performed is dependent on the machining cost, the feature being machined must be specified. In both scenarios, a pocket with dimensions of 100 mm in the X direction, 150 mm in the Y direction, and 25 mm deep was used. The tool path is shown in Fig. 3.3.

3.3.1 Cost Estimation and Stability Limit

In order to perform the updating procedure, the following is required: 1) the cost of performing the desired operation for selected operating conditions; and 2) the stability condition must be determined from experimental data. These issues are addressed in the following sections.

The cost model for deterministic setups used here selects experiments that maximize expected profit (because revenue was considered constant, minimizing the expected cost is equivalent). The cost function for this exercise does not include the effects of tool wear; it was neglected for the 6061-T6 workpiece/TiCN-coated carbide tool combination. The simplified cost, C, (Eq. 3.2) is based on the machining cost per minute (r_m = \$2) and machining time (t_m, which depends on the part path geometry and machining parameters).

$$C = r_m t_m \tag{3.2}$$

Due to the nature of the part path, for any selected spindle speed the cost function is stepped; see Fig. 3.4. These steps occur at an integer fraction of the pocket depth. As the maximum permitted axial depth increases, the number of steps required to complete the pocket decreases until the permissible axial depth equals or exceeds the pocket depth, where only a single pass is required to produce the desired pocket.

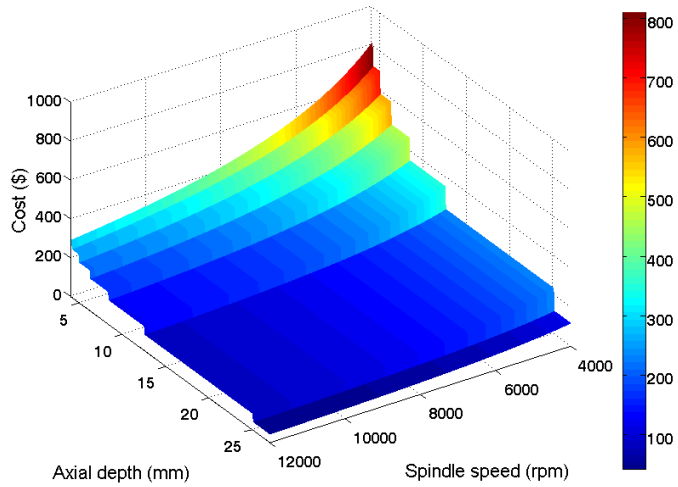

Figure 3.4: Cost of machining at axial depth-spindle speed combinations given that the resultant cut is stable. I would like to thank Raúl Zapata for allowing me to use this figure.

Stability analysis for the simulated testing scenario was straightforward. A set of stability lobes was calculated and the test point was compared to the stability boundary for the "unknown" system behavior. If the test point was located at a spindle speed axial depth combination that appeared below the boundary, the test was stable. Otherwise, the test was considered unstable.

For experiments, however, stability identification is more involved. Considering the machined surface, stable milling behavior results in a repeatable surface with small feed marks

which depend on the feed per tooth value. The self-excited vibrations which occur during unstable cutting modulate the cutting forces and resulting vibrations due to fluctuations in chip thickness. These vibrations occur at or near the one of the system's natural frequencies and result in a rough surface. The exact boundary between stable and unstable machining based on the surface finish is somewhat subjective.

The frequency content of the milling signal can also be used to identify chatter [3.14]. For stable cutting, only forced vibrations are present. Content is expected only at the tooth passing frequency and its harmonics in this case. In the unstable case, both forced and self-excited vibrations occur. Although the self-excited vibrations occur near a system natural frequency, the magnitude at which this spurious frequency content is identified as chatter is also subjective. Additionally, if the chatter frequency is near a tooth passing frequency, interaction between the two can occur. In order to make the best decisions on stable versus unstable behavior, both the appearance of the cut surface and the frequency content of the cutting force data were used.

3.3.2 Modeling Uncertainty in Stability Limits

In this section, the inference model and Bayesian updating of the model from stability testing is presented. In addition, the experimental setup used for the case study is described.

In the absence of other information, a simple model is to assume that, given knowledge of the stability limit at a spindle speed *s* and axial depth *b*, the uncertainty about the stability limit increases as the distance from the operating point increases; see Fig. 3.5. At a slightly higher (or lower) spindle speed than the known point, the axial depth of the stability limit will be:

- at a slightly higher axial depth with probability p_1;
- at the same axial depth with probability p_2; or
- at a slightly lower axial depth, with probability $p_3 = 1 - p_2 - p_1$.

A three-outcome explanation is used to explain this model, but, in principle, any number of outcomes can be imposed to obtain a higher resolution and the propagation of uncertainty depends on the decision maker's beliefs. While this assumption does not directly imply the actual location of the stability limit, it does imply something about its structure: knowledge of the stability limit at one spindle updates the prediction of the stability limit at nearby spindle speeds, i.e., knowledge can be propagated from one spindle speed to another. Although the details have been omitted for brevity, the updated prediction at any spindle speed for the selected propagation model can be calculated by solving a system of parabolic partial differential equations [3.15-3.16]. The extent to which the prediction at nearby spindle speeds is updated can be controlled by changing p_1 and p_2. In addition, it was assumed that for all spindle speeds, the stability limit occurs between 0 mm and the maximal axial depth defined by the flute length (14 mm). With these two assumptions, the initial probability distribution (i.e., the prior) is defined as shown in Fig. 3.6.

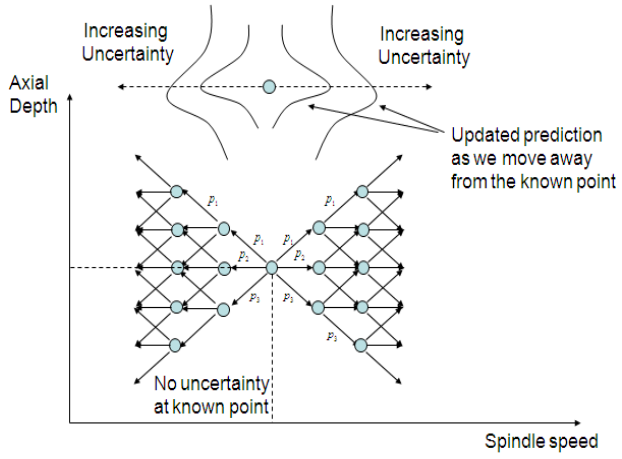

Figure 3.5: Discretized representation of the initial stability model. Knowledge of the stability limit at one spindle speed updates the prediction at nearby spindle speeds. As the distance from the known point is increased, the extent to which the prediction is updated decreases. I would like to thank Ali Abbas for allowing me to use this figure.

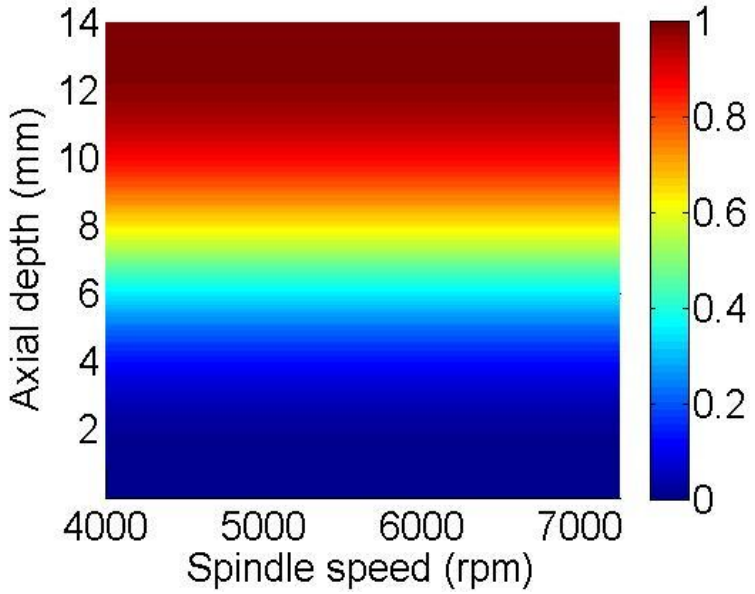

Figure 3.6: Initial prediction for the stability limit. Color denotes the probability that an axial depth-spindle speed combination will result in an unstable cut (0 – not likely, 1 – very likely).

3.3.3 Updating the Prediction from Experiments

Once the initial prediction is made, it can be updated whenever new information is observed, such as through experimentation. For the case study, experimentation was performed by selecting a particular spindle speed-axial depth combination, performing a test cut, and identifying stable or unstable behavior. For both stable and unstable results, the information gained from this experiment was expressed as an inequality. For a test at axial depth B and spindle speed S, a stable cut implies that the axial depth of the stability limit at spindle speed S must be greater than B, while an unstable cut implies that the axial depth of the stability limit at spindle speed S must be less than B.

Applying Bayes' rule, updating the prediction of the stability limit at the tested spindle speed was fairly straightforward. See Fig. 3.7, which shows probability density functions before and after updating. This knowledge was propagated to other spindle speeds to determine the updated

predictions there as well. Figure 3.8 shows the results of a sample update for stable and unstable cases.

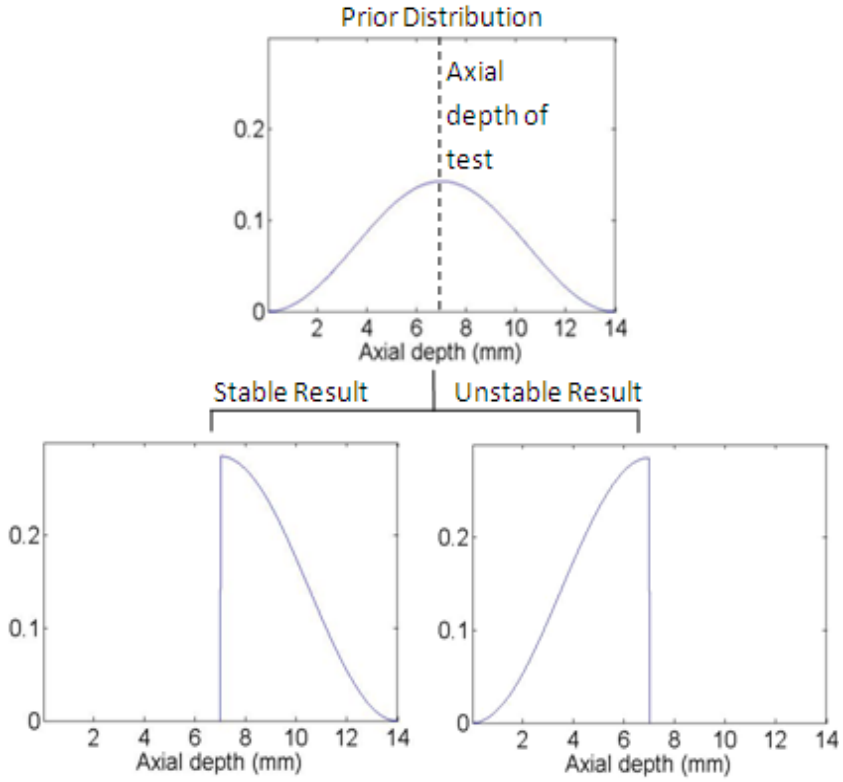

Figure 3.7: Updating the prediction at the tested spindle speed. The likelihood function is constant where the inequality constraint is satisfied and zero where it is violated. The updated distribution is therefore given by truncating and normalizing the prior distribution as a PDF.

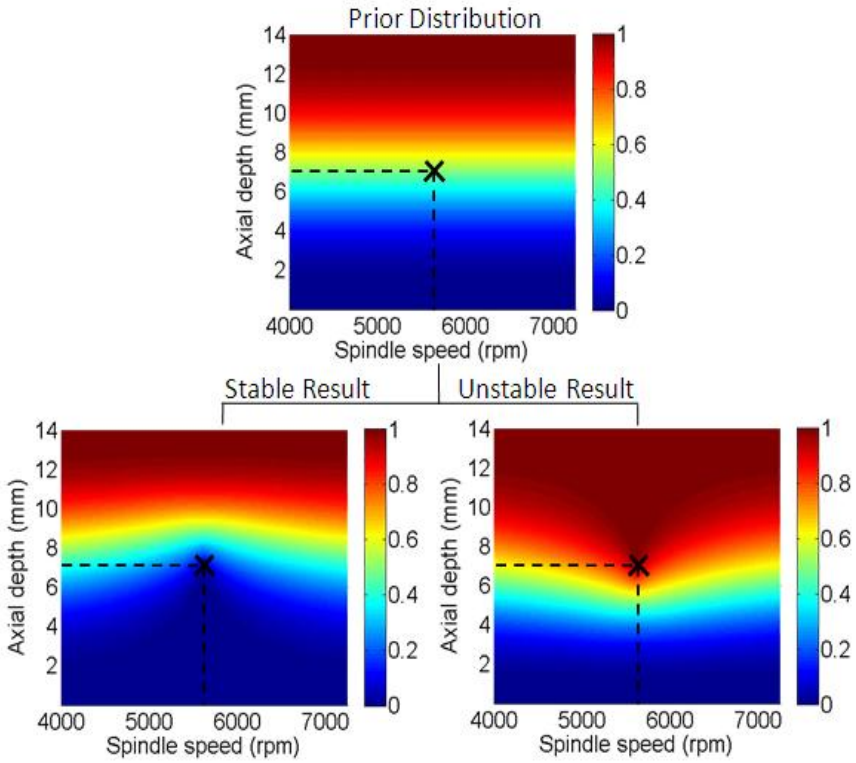

Figure 3.8: Updates for a single stability test for stable and unstable results within the probability of instability CDF field. The test point is denoted with an X.

3.3.4 Experimental Setup for Case Study

Stability testing was performed using a three-component force dynamometer as a rigid base for the 6061-T6 aluminum test piece and a compliant tooling system, realized using a long collet holder and stub length carbide tool as shown in Fig. 3.9.

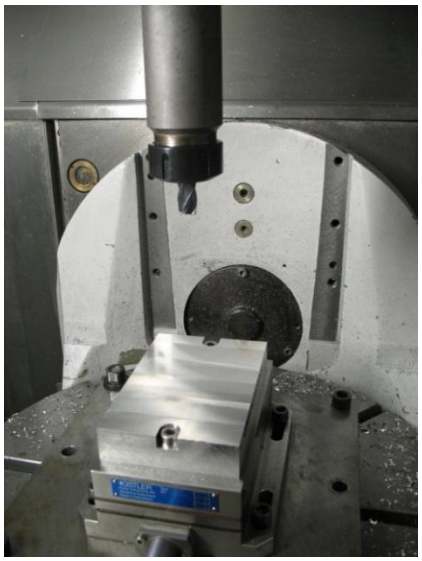

Figure 3.9: Tool and holder combination used in the experiments; the workpiece is mounted on the dynamometer. I would like to thank Raúl Zapata for allowing me to use this figure.

The workpiece geometry is shown in Fig. 3.10. In preparation for the test cuts, the top of the part was faced and a 14 mm deep slot was made with the test tool. The side of the slot that was to be machined during the subsequent test cut was finished machined with passes 1 mm deep axially and 0.1 mm deep radially. Both the slot and the (post-test) clean-up passes were performed at 7250 rpm with a feed per tooth of 0.15 mm/tooth.

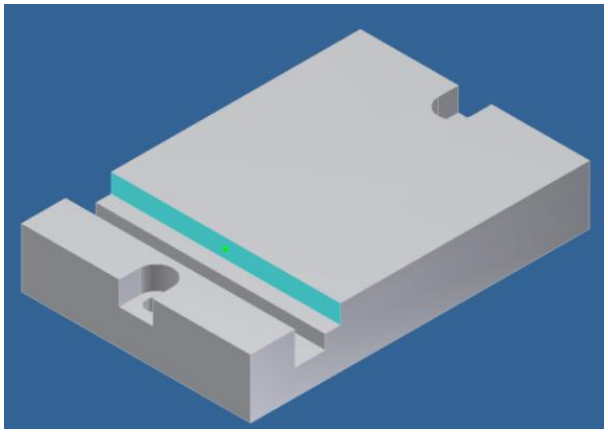

Figure 3.10. Schematic of the workpiece with the wall of the test cut surface highlighted. I would like to thank Raúl Zapata for allowing me to use this figure.

Stability of a selected point was determined experimentally by completing 50% radial immersion (6.35 mm radial depth) down milling cuts at the selected axial depth of cut and spindle speed. The highlighted surface in Fig. 3.10 is the sidewall of the test cut. The surface texture of the machined wall was used in combination with the frequency content of the force data obtained from the dynamometer in order to determine if chatter (unstable behavior) existed. After the first test, the surface was prepared using a cleaning pass and the testing process was repeated for a new spindle speed-axial depth of cut combination. For the given geometry, this enabled up to 14 test cuts on each block. The updating scheme was implemented over a series of up to 14 test cuts, where the test cuts were selected according to the programmed algorithm and the stable/unstable test results were determined as described previously. The spindle speed and axial depth parameters were selected by the updating algorithm as it searched for points with maximum value of experimentation within the selected domain.

3.3.5 Value of Information and Experiment Design

Bayesian inference combined with decision analysis models enables a dollar value to be place on the information gained from an experiment prior to performing it. To illustrate this point, consider a simple situation where only three spindle speed-axial depth combinations are available (A, B, and C). Suppose it is initially predicted that A is definitely stable, while B and C each have a 50% chance of being stable. In addition, suppose that the cost of machining (assuming the cut is stable) is $100 using A, $50 using B, and $30 using C and that only stable operating points will be used (in other words, the cost of performing an unstable cut is very large).

Prior to performing the stability test, only A can be chosen as the operating point and, therefore, if no testing is performed the cost of machining will be $100. However, suppose the option of performing a single stability test at either A, B, or C was given; how can the proper test be selected? Because it is already known that a test at A will have a stable result, no test should be completed at A because no new information will be obtained. However, if it was possible to test at B, there is a 50% chance that the result is unstable, in which case the choice will still be A and the cost will be $100. On the other hand, there is also a 50% chance that the test will be stable, in which case B will be selected and the cost will only be $50. Thus, the expected cost of machining given the result of a test at B is $75. The value gained by testing at B (defined as the cost prior to testing minus the expected cost after testing) is therefore $25. Similarly, the value gained by testing at C can also be calculated. There is a 50% chance that the result will be unstable, in which case machining will be completed at A and the cost will be $100. There is also a 50% chance that the test will be stable and then machining will be completed at C and the cost is only $30. Thus, the expected cost given the result of a test at C is $65 and the value gained by testing at C is $35. Now (assuming the goal is to maximize profit), the question of which test to perform has a straightforward answer: choose the test which adds the most value. For this example, testing would be completed at C.

This example was simplistic. However, the implications are powerful: when using a Bayesian inference model, it is possible to calculate the expected value that will be gained from performing any experiment. In other words, by using a Bayesian inference model, the guesswork that is normally involved in experimental design is eliminated. It is only necessary to perform the sequence of tests which adds the most value.

From a practical point of view, however, this still leaves the question of how to determine which sequence of experiments is optimal (adds the most value). In the previous example, simply calculating the value of each possibility and comparing them (optimization by exhaustive enumeration) was computationally simple. However, in more complicated situations, computational complexity can become a significant issue.

In the case study, the computational complexity is particularly problematic since sequences of experiments are being optimized as opposed to a single experiment. The reason for this is that the number of decision variables increases exponentially with the number of experiments in the sequence. For example, when optimizing a sequence of two stability tests, three sets of decision variables are required (as opposed to one for a single test). The reason is that the value of the second test depends on the result of the first test. Therefore, the optimal parameter combination to use in the second test if the result of the first test is stable will, in general, be different from the optimal parameter combination to use in the second test if the result of the first test is unstable. Consequently, to truly optimize the two-test sequence, simultaneously optimization over all three sets of parameters (parameter combination for the first test, parameter combination for the second test if the first test result is stable, and the parameter combination for the second test if the first test result is unstable) must be completed.

In the case study, the calculation of the optimal sequence of experiments (via exhaustive enumeration) is made computationally feasible by using a greedy heuristic. In this case, each test is treated separately, rather than optimizing the sequence of tests as a whole. In terms of a two-test sequence, this means that, instead of simultaneously optimizing over all three parameter sets, the parameters for the first test are optimized first by assuming that no tests will be performed afterward. Then, after this test is performed and the result is known, the second test is optimized.

With this heuristic, the computational complexity increases only linearly with the number of tests.

Overall, while this greedy heuristic makes the optimization computationally feasible, the behavior of the sequence of experiments may deviate from the behavior expected from the truly optimal sequence of experiments. It is likely that with an intelligent optimization algorithm the computational complexity problem could be alleviated without a sacrifice in performance. However, this particular optimization problem has a number of properties which make it challenging (the probability of stability is not convex or continuous everywhere and is expressed in terms of a Fourier series). Therefore, the greedy heuristic is a reasonable choice for providing a baseline application of Bayesian methods.

3.3.6 Results of Simulated Testing Scenario

Using the methods described in the previous sections, a sequence of 13 stability tests was generated. The result of each test was determined by comparing the test point to the reference stability limit. The parameters for each test, as well and the results of the test, are provided in Table 3.2. The updated prediction after all of the tests were performed is shown in Fig. 3.11 (the prior distribution is shown in Fig. 3.6). To gain some insight into how the performance is affected by the input parameters, two additional test sequences were generated. In the first case, the maximum spindle speed was reduced from 30000 rpm to 25000 rpm; see Fig. 3.12. In the second case, the stiffness of the system (an input parameter for the FRF) was changed from 20 MN/m to 30 MN/m; see Fig. 13. The CDFs are shown in Figs. 11 through 13 where the color scale indicates probability of instability (0 indicates definite stability). In each case, the reference stability lobe is shown superimposed for comparison.

Table 3.2. Sequence of tests performed in simulated testing scenario and their results.

Test number	Spindle speed (rpm)	Axial depth (mm)	Stability condition
1	30000	0.42	stable
2	30000	2.82	stable
3	30000	5.04	stable
*4	30000	6.30	stable
5	27100	8.34	unstable
6	25000	8.34	unstable
7	28600	8.34	unstable
8	26200	8.34	unstable
9	29200	8.34	unstable
10	27900	8.34	unstable
11	21200	12.54	unstable
12	20000	12.54	stable
13	20300	12.54	stable

*Due to a numerical error, point 4 was recommended as a test point twice. The error resulted from expressing the probability of stability in terms of a Fourier series (Gibbs phenomenon [3.17]).

From these results, a number of insights can be drawn regarding the capability of the model. First, because the greedy heuristic was used, there is a bias towards selecting a test point relatively close to the point which has the least cost of all points previously determined as stable and the performance in identifying a local maximum is very good. However, as the tests tend to cluster around a local maximum, the performance in moving from one stability peak to another is poor (this degrades the performance in the increased stiffness case). In contrast, for the true optimal sequence of experiments, a better balance between exploratory tests (tests which serve

the purpose of determining viable places to look for the global maximum in future tests) and tests which attempt to achieve the global maximum would be expected. This would likely lead to improved performance because more of the domain would be explored.

Second, it is seen that the performance of the model is somewhat dependent on the axial depth-spindle speed domain. The more centered the reference stability limit is in terms of maximum axial depth, the better the model performs (this results from the prior distribution choice). Simply put, the better the initial prediction, the better the inference model will perform.

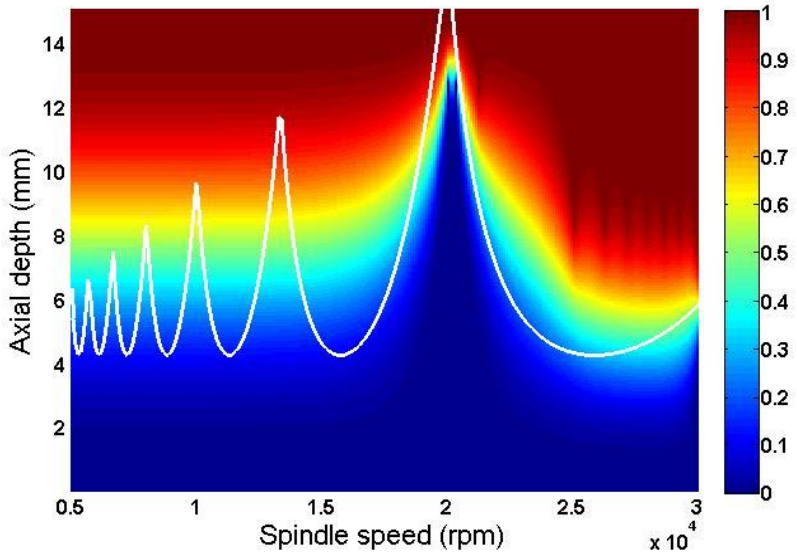

Figure 3.11: Final CDF for the first numerical test set. Reference stability lobes are superimposed in white.

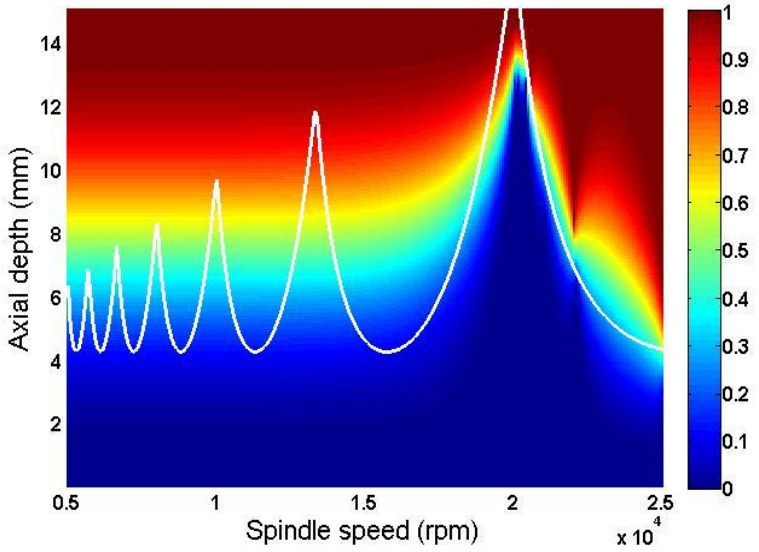

Figure 3.12: Final CDF for the reduced maximum spindle speed numerical test set.

Third, it is seen that the sequence of experiments performed is fundamentally tied to the cost function. This occurs in this case due to the stepped nature of the cost function. Comparing the sequence of tests performed to the cost function at these points, an interesting pattern is observed: every test is selected at a step in the cost function.

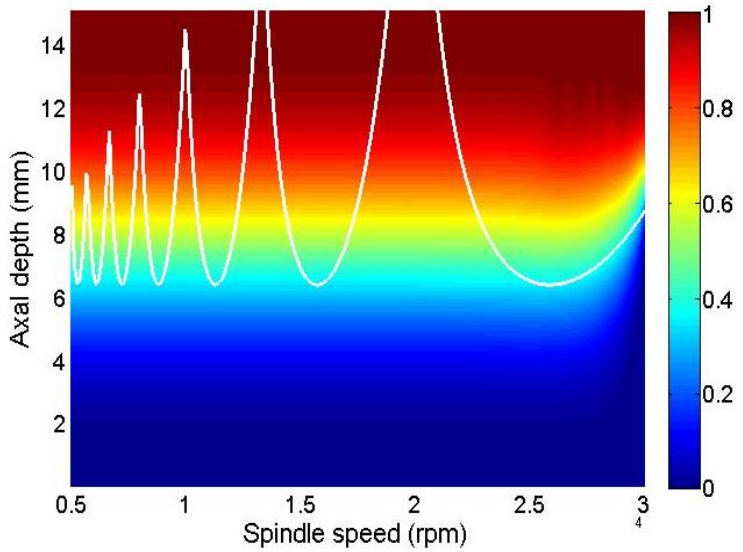

Figure 3.13: Final CDF for the numerical tests with increased tool stiffness (20 MN/m to 30 MN/m).

Overall, although the performance is somewhat dependent on the particular input parameters (reference stability limit and test domain), in many cases the cost achieved using the inference model is not significantly greater than the true minimal achievable cost. It is also important to note that the limitations of the model are in no way related to the underlying methodology of Bayesian inference, but rather result from the choice of assumptions. Specifically, the assumption of no prior knowledge of machining dynamics and the heuristic optimization algorithm caused the most difficulty.

3.3.7 Results of Experimental Testing Scenario

To provide a realistic application of the inference model, the same methods were again applied to a sequence of 13 milling tests. The parameters of each test, as well as the result of the test, are shown in Table 3.3. The updated prediction after all of the tests were performed and the location of these tests are shown in Fig. 3.14.

Table 3.3: Experimental test points and results.

Test number	Spindle speed (rpm)	Axial depth (mm)	Stability result
1	7250	0.51	Stable
2	7250	2.80	Stable
3	7250	5.04	Stable
4	7250	6.27	Stable
*5	7250	8.34	Stable
6	7250	12.54	Unstable
7	6977	12.54	Unstable
8	6743	12.54	Unstable
9	7120	12.54	Unstable
10	6509	12.54	Unstable
11	6275	12.54	Unstable
12	6860	12.54	Unstable
13	6626	12.54	Unstable

*Point 5 was recommended as a test point twice due to Gibbs phenomenon [3.17].

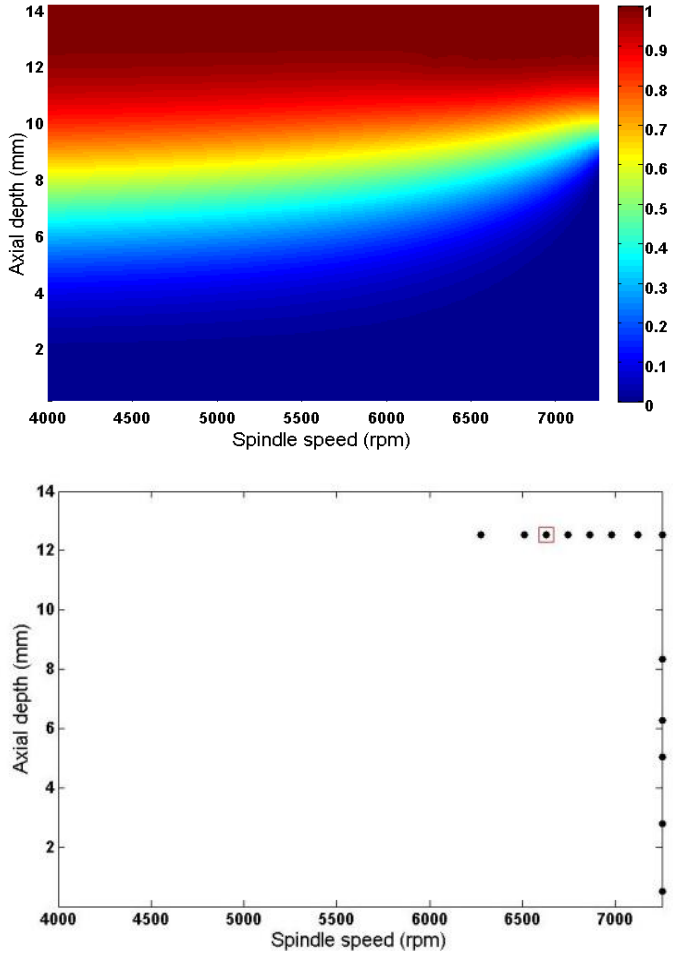

Figure 3.14: CDF representation of probability of instability (top) and the points tested (bottom). The final test point is indicated by a square.

Because the true location of the stability limit is unknown, it is impossible to determine how the cost obtained using the inference model compares with the true minimal cost. Based on the final prediction, it is probable that there is a peak at a lower spindle speed which the model did not locate. If this is the case, then there could be stable operating parameters which require only two (axial) tool passes rather than the three passes provided by the inference model. Overall, the insights from the experimental testing scenario are essentially the same as from the simulated

testing scenario: the performance of the model is acceptable, but could be improved by an improved prior and a non-heuristic optimization scheme.

3.4 CONCLUSION

The goal of this work was to apply a Bayesian inference model to stability prediction for high-speed machining. The motivation for implementing a Bayesian inference model was: 1) a Bayesian inference model enables a prediction which considers both theory and experimental results; and 2) when using a Bayesian inference model, experiments can be chosen such that the expected value added by performing the experiment is maximized. In other words, Bayesian inference models remove the guesswork that is normally involved when designing experiments.

For the study presented here, a worst case situation was assumed where no prior knowledge of machining dynamics was available. Only stability testing was considered and the optimal experimentation policy was generated heuristically. The results obtained provide a lower bound on what can be achieved using Bayesian inference models. Even so, the model performed fairly well. Although the performance was somewhat dependent on the input parameters, in the majority of simulated (numerical) testing scenarios the difference between the cost of machining the pocket at the parameters determined by the inference model and the true minimal cost of machining the pocket was negligible. This is most likely true for the experimental testing scenario as well judging from the cost function and the unstable experimental results.

REFERENCES

[3.1] Tobias, S.A., 1965, Machine-Tool Vibration, Blackie and Sons Ltd., Glasgow, Scotland.

[3.2] Arnold, R.N., 1946, The Mechanism of Tool Vibration in the Cutting of Steel, Proceedings of the Institution of Mechanical Engineers, 154/4: 261-284.

[3.3] Tobias, S.A. and Fishwick, W., 1958, The Chatter of Lathe Tools under Orthogonal Cutting Conditions, Transactions of the ASME, 80: 1079.

[3.4] Tobias, S.A. and Fishwick, W., 1958, Theory of Regenerative Machine Tool Chatter, The Engineer, 205.

[3.5] Tlusty, J. and Polocek, M., 1963, The Stability of the Machine-Tool against Self-Excited Vibration in Machining, Proceedings of the International Research in Production Engineering Conference Pittsburgh, PA, 465.

[3.6] Koenisberger, F. and Tlusty, J., 1967, Machine Tool Structures-Vol. I: Stability Against Chatter, Pergamon Press.

[3.7] Merrit, H., 1965, Theory of Self-Excited Machine Tool Chatter, Journal of Engineering for Industry, 87/4: 447-454.

[3.8] Tlusty, J., Zaton, W., Ismail, F., 1983, Stability Lobes in Milling, Annals of the CIRP, 32/1: 309-313.

[3.9] Altintas, Y. and Budak, E., 1995, Analytical Prediction of Stability Lobes in Milling, Annals of the CIRP, 44/1: 357-362.

[3.10] Budak, E. and Altintas, Y., 1998, Analytical Prediction of Chatter Stability Conditions for Multi-Degree of Freedom Systems in Milling. Part I: Modeling, Part II:

Applications, Journal of Dynamic Systems, Measurement and Control, Transactions of the ASME, 120: 22-36.

[3.11] Abbas, A., Yang, L., Zapata, R., and Schmitz, T., 2009, Application of Decision Analysis to Milling Profit Maximisation: An Introduction, International Journal of Materials and Product Technology, 35(1-2): 64-88

[3.12] Zapata, R., Schmitz, T., Traverso, M., and Abbas, A., 2009, Value of Information and Experimentation in Milling Profit Optimization, International Journal of Mechatronics and Manufacturing Systems, 2/5-6: 580-599

[3.13] Altintas, Y., 2000, Manufacturing Automation, Cambridge University Press, Cambridge, UK.

[3.14] Schmitz, T. and Smith, K.S., 2009, Machining Dynamics: Frequency Response to Improved Productivity, Springer, New York, NY.

[3.15] Taira, K., 1988, Diffusion Processes and Partial Differential Equations, Academic Press Inc., Boston, MA.

[3.16] Nagasawa, M., 1993, Schrödinger Equations and Diffusion Theory, Basel, Birkhäuser Verlag.

[3.17] Gibbs, J.W., 1898, Fourier Series, Nature, 59: 200 [and 606, 1899].

CHAPTER 4

DEMAND CURVE PREDICTION VIA PROBABILITY ASSIGNMENT OVER A FUNCTIONAL SPACE

In this chapter, the methods derived in the first chapter will be applied to the problem of demand curve prediction. The motivation for doing this is at least two-fold. First, it illustrates that these methods are versatile and applicable to more than one field. In addition, the problem of demand curve prediction is particularly interesting in that demand curves are generally assumed to satisfy certain regularity conditions and this must be reflected in the prediction. In particular, it will be assumed that the demand curve satisfies 1.) demand for a good decreases monotonically with price, 2.) demand for a good goes to zero as price goes to infinity, and 3.) demand for a good goes to infinity as price goes to zero. Thus, when making a prediction for a demand curve, one must choose a probability distribution which assigns a nonzero probability to any function which satisfies these conditions and a probability of zero to any function which does not. To accomplish this, a coordinate transformation is first applied which greatly simplifies the constraints. The remaining constraint is then enforced through the choice of transition probabilities.

4.1 INTRODUCTION

Currently, the majority of demand curve predictions are based on curve fits. While these may be adequate when making simple decisions, there are a number of issues that make them somewhat impractical in real world decision making. For one, they only output a single curve prediction. As a result, they are not compatible with rigorous decision making in the presence of uncertainty. In addition, with curve fits it is impossible to construct a predictive model for demand curves that is both globally regular (meets the monotonicity constraints of demand curves) and locally

flexible [4.1] (given enough observations of an arbitrary "true" demand curve meeting the global constraints, the prediction will converge to the true curve). In this work, we present an alternative method of demand curve prediction which is fundamentally different from curve fitting. In particular, a probability distribution is assigned over the space of all functions that satisfy the regularity conditions. First, a hyperbolic coordinate transformation is used to impose the regularity conditions is introduced. Then, the problem of assigning a probability distribution over the space of functions and its update with information is discussed.

4.2 THE HYPERBOLIC COORDINATE TRANSFORMATION

For a price P>0 and demand Q>0, consider the hyperbolic coordinate transformation given by:

$$u = \frac{1}{2}(\log P - \log Q).$$

$$v = \frac{1}{2}(\log P + \log Q).$$

The inverse transformation is given by:

$$P = e^{v+u}.$$

$$Q = e^{v-u}.$$

This transformation maps the first quadrant to the entire R^2 plane. In particular, the point $\{P = 0, Q = \infty\}$ is mapped to the line $u = -\infty$, the point $\{P = \infty, Q = 0\}$ is mapped to the line $u = \infty$, the P axis and Q axis are mapped to the line $v = -\infty$, and the lines $P = \infty$ and $Q = \infty$ are mapped to the line $v = \infty$ (Figure 4.1).

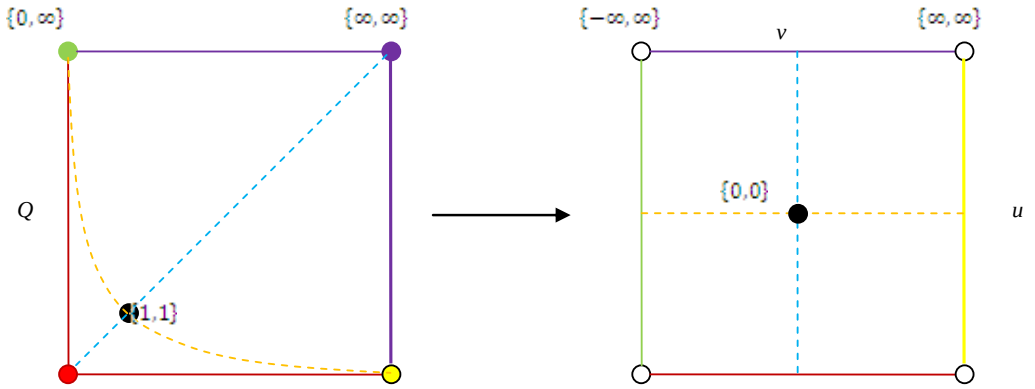

Figure 4.1: Mapping of the hyperbolic coordinate transformation

The advantage of using hyperbolic coordinates is that the regularity conditions imposed on the demand system can be expressed nicely in u-v coordinates [4.2]. Any function $Q(P)$ that is monotonically decreasing, and has the positive real values as both a domain and range can be expressed as a function $v(u)$ which satisfies

$$\operatorname{dom} v(u) = \mathbb{R}.$$

$$\left|\frac{dv}{du}\right| < 1. \qquad (4.1)$$

In addition, u-v coordinates have nice interpretations in terms of economic variables. For one, the total revenue R generated (the product $P * Q$ in P-Q coordinates) is given by

$$R = e^{2v}.$$

As a result, maximizing the revenue for a given demand curve is equivalent to finding the maximum v of that curve. Furthermore, the elasticity E, given in P-Q space by

$$E = -\frac{P}{Q}\frac{dQ}{dP} = -\frac{d\ln Q}{d\ln P},$$

is given in u-v space by

$$E = \frac{\left(1 - \frac{dv}{du}\right)}{\left(1 + \frac{dv}{du}\right)}. \qquad (4.2)$$

Therefore, lines of constant slope in the u-v plane are equivalent to lines of constant elasticity and horizontal lines are equivalent to lines of unit elasticity.

4.3 CONSTRUCTING A PROBABILITY DISTRIBUTION ON THE HYPERBOLIC PLANE

In order to construct a probability distribution over the space of functions, a discrete random walk model can be used. In particular, the following two assumptions are made:

Assumption 4.1 *When predicting the behavior of the demand curve at any price P_0, only information regarding the curve in an infinitesimal neighborhood of P_o needs to be considered.*

Assumption 4.2 *Knowledge of the second or higher derivatives is irrelevant when predicting the demand curve.*

Now, suppose that it is known that the demand curve goes through the point $\{P_1, Q_1\}$ and one is interested in calculating the probability that the demand curves travels through the point $\{P_1 + \Delta P, Q_1 - \Delta Q\}$ for some $0 < \Delta P$ and $0 < \Delta Q < Q_1$. Because of assumptions 1 and 2, this conditional probability can be calculated using a lattice walk model. In P-Q space, these conditional probabilities are difficult to evaluate since the transition probabilities MUST depend explicitly on P and Q, and these dependencies are necessarily singular near the axes (Figure 4.2).

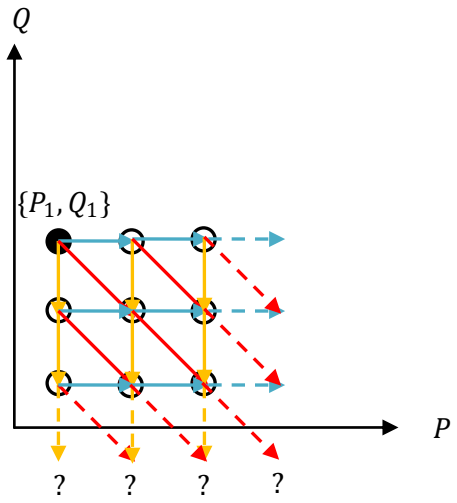

Figure 4.2: Evaluating a random lattice walk in $\boldsymbol{P}$-$\boldsymbol{Q}$ space is difficult due to the boundaries on the axes. In $\boldsymbol{u}$-$\boldsymbol{v}$ space, this is not an issue since the $\boldsymbol{P}$-$\boldsymbol{Q}$ axes are mapped out of the plane (to the line v=-∞).

However, in u-v space these dependencies are already accounted for. Because of this and assumptions 4.1 and 4.2, the transition probabilities are independent and identically distributed. If equation (4.1) (resulting from the monotonicity in P-Q space) was not imposed, these transition probabilities would converge to a Wiener process and the resultant conditional probabilities would be normally distributed [4.3]. With the requirement of equation (4.1), however, the conditional probabilities must be of compact support and therefore cannot be normally distributed. Instead, the conditional probabilities are given by a Lorentz invariant distribution [4.4]. The lattice walk and the resultant conditional probability distribution are shown in Figure 4.3. After transforming back into P-Q space, by using these conditional probabilities and assumptions 1 and 2, the probability that the demand curve goes through any point $\{P, Q\}$ can be calculated.

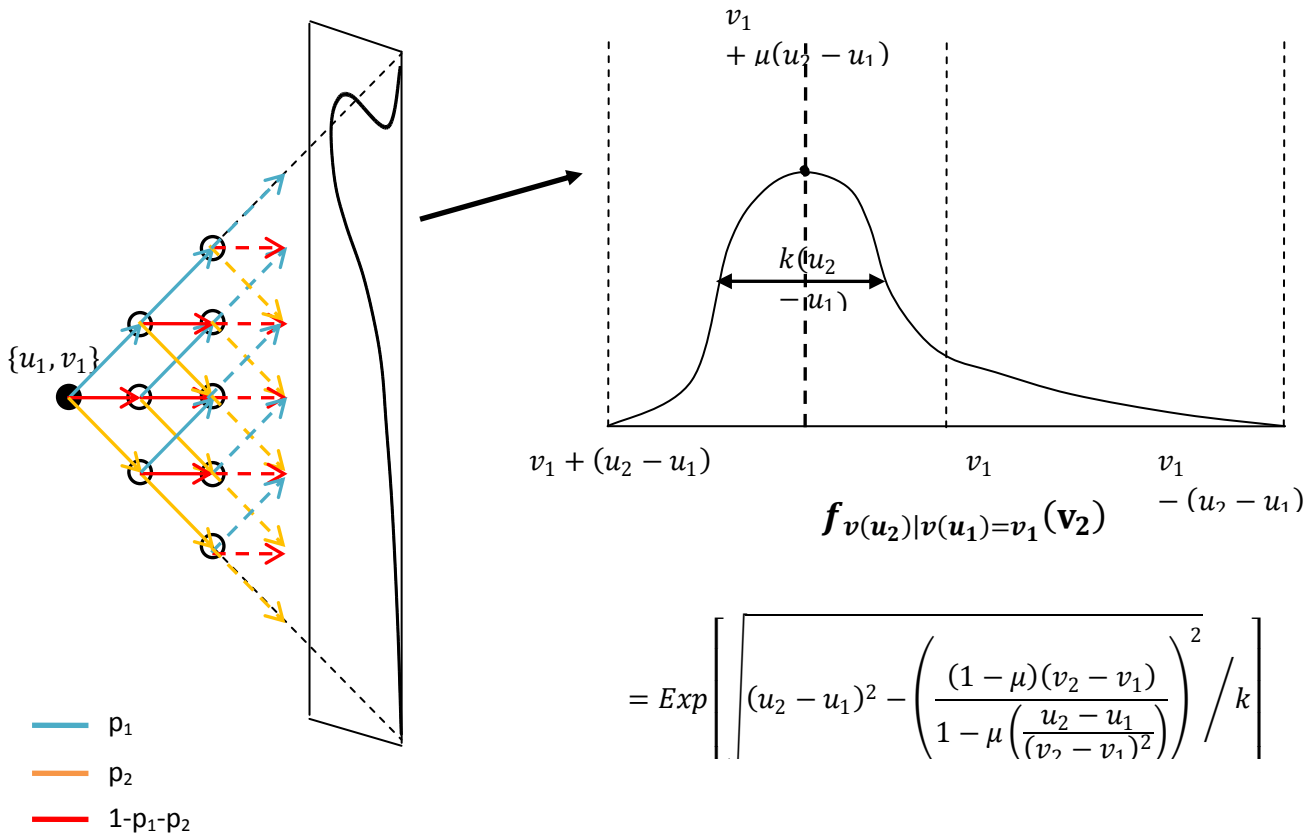

Figure 4.3: On the left, the random lattice walk is shown. As the number of steps increases to infinity, the distribution converges to a two parameter Lorentz-invariant distribution (shown on the right) where the two parameters (mean **μ** and scale / spread parameter ***k***) are determined from the transition probabilities p_1 and p_2.

4.4 USING THE PROBABILITY MODEL

To make an initial prediction of the demand curve, one needs to define which points of the demand curve are already known, as well as assigning values to the μ and k parameters. In the general case, they can take different values at each $\{u, v\}$ coordinate, so initial parameter *fields*, $\mu_0(u, v)$and $k_0(u, v)$ need to be chosen. In terms of the P-Q coordinate system, assigning these parameter fields is equivalent to answering the following question for every $\{\mathrm{P}_0, \mathrm{Q}_0\}$:

Given that the demand curve travels through $\{\mathrm{P}_0, \mathrm{Q}_0\}$, *what is the expectation and variance of the slope of the demand curve at this point?*

Or using equation (4.2)

Given that the demand curve travels through $\{P_0, Q_0\}$*, what is the expectation and variance of the elasticity of the demand curve at this point?*

For example, suppose someone is interested in predicting the demand curve for some type of product X. Using historical sales data, they believe that the demand curve travels through the points $\{P, Q\}$={\$29, 3960 units}, {\$60, 2040 units} , {\$200, 810 units} and {\$400, 240 units}. In addition, they define the μ_0 and k_0 fields as constant between any two known points (Table 4.1). In particular, they choose the μ_0 field as the average elasticity between two known points, and a k_0 field that decreases near the axes (less uncertainty/spread near axes). Using these assumptions, the resultant prediction for the demand curve of product X is as shown in Figure 4.4. The price which maximizes expected revenue generated by selling product X can then be calculated from this prediction, which in this case is \$200. In addition, these predictions can be used to place a value on information gathering activities. For example, suppose that the demand Q_{test} for product X at price P_{test} can be estimated in a small test market. Since the expected revenue generated prior to the test, the probability of each possible outcome of the test (the demand curve prediction at P_{test}) and the expected revenue generated with each outcome can all be calculated prior to performing the test, the expected value gained from a market test at any P_{test} can be calculated. P_{test} can then be chosen to maximize this value.

Table 4.1: Parameter fields for predicting the demand curve of product X.

Region:	**Drift Field $\mu_0(u,v)$:**	**Diffusion Field $k_0(u,v)$:**
\$0< P <\$29	0.310*	0.03
\$29< P <\$60	0.0460	0.05
\$60< P <\$200	0.132	0.1
\$200< P <\$400	-0.274	0.05
P >\$400	-0.263*	0.03

*-Values chosen arbitrarily such that $P * Q$ goes to 0 as P or Q goes to 0.

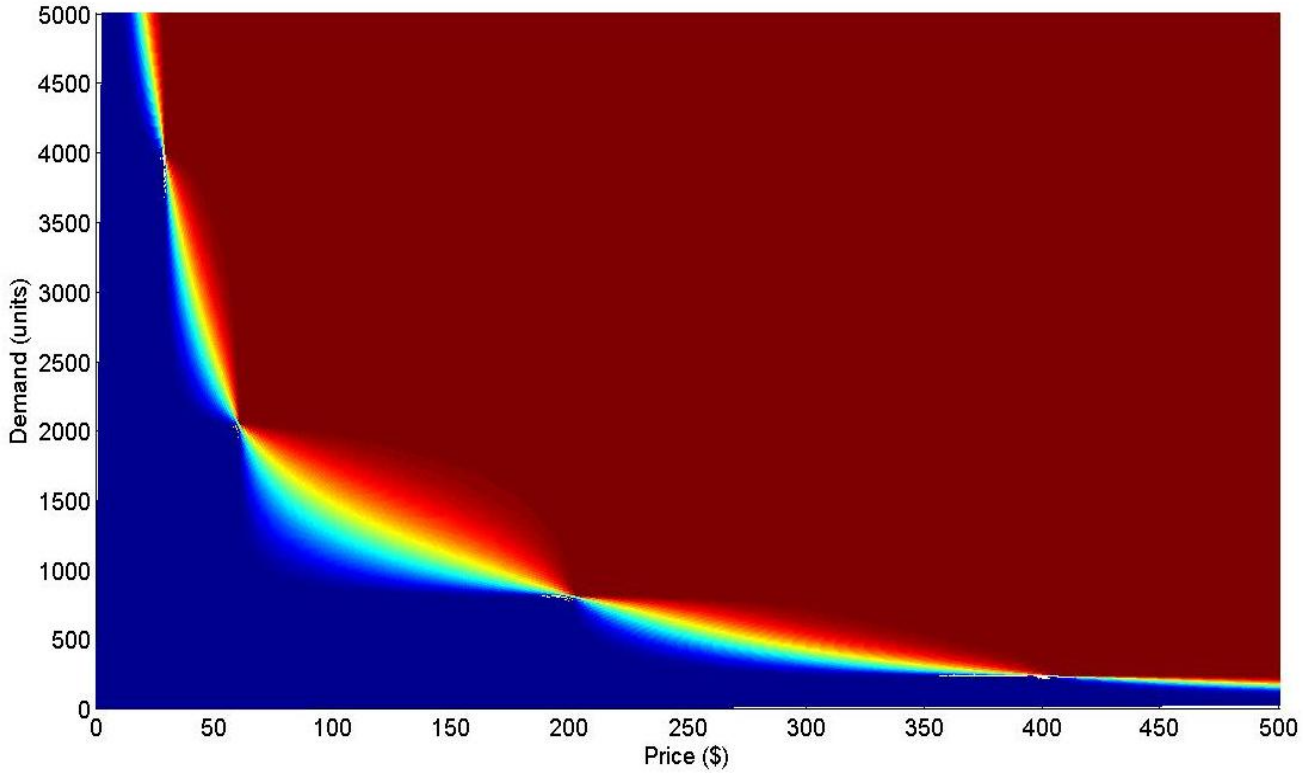

Figure 4.4: Demand curve prediction for product X.

4.5 CONCLUSION

In conclusion, the predictive model presented here has a number of advantages over traditional demand curve predictions which utilize statistical curve fits. Fundamentally, it differs from curve fitting since a probability is assigned over a full space of functions instead of assuming a strict functional form for the demand curve. As a result, these methods result in predictions that are both globally regular and locally flexible, which is impossible with traditional curve fitting. In addition, since this predictive model is compatible with rational decision making, it can be used to determine not only optimal pricing strategies, but also optimal information gathering strategies. Overall, although these methods are somewhat more complex mathematically than curve fits, they also provide stronger functionality which is often preferred when making real world decisions involving real dollars.

REFERENCES

[4.1.] Barnett WA., Yue P., 1998, Semiparametric estimation of the Asymptotically Ideal Model: the AIM demand system, *In Nonparametric and Robust Inference, Advances in Econometrics* 7, JAI Press, Greenwich, CT; 229-52.

[4.2.] Dunkel, J. and Hänggi, P., 2009, Relativistic Brownian Motion, Physics Reports, Vol. 471, p. 1-73.

[4.3.] Nagasawa, M., 1993, Schrödinger Equations and Diffusion Theory, Basel, Birkhäuser Verlag.

[4.4.] Ikeda, N. and Matsumoto, H., 1999, Brownian Motion on the Hyperbolic Plane and the Selberg Trace Formula, *Journal of Functional Analysis,* Vol 163, p. 63-110.

APPENDIX A

PARAMETERS FOR NUMERICAL TEST CASES IN CHAPTER 2

Table A.1: Parameters used to generate the reference stability limit.

Description	Value	Units
Tangential cutting coefficient	2500	N/mm^2
Radial cutting coefficient	750	N/mm^2
Tool stiffness	$5x10^6$	N/m
Damping ratio	0.05	-
Tool natural frequency	2.4	kHz

Table A.2: Parameters defining the Brownian distribution prior to experimentation.

Description	Parameter name	Value
Upper bound on stability limit	g_{max}	1.2 mm
Diffusion field	$k(x,g)$	0.00005
Drift field	$\mu(x,g)$	0
Creation and killing field	$c(x,g)$	$c_{up}(x,g) = \begin{cases} \infty \; for \; g < 0, g > g_{max} \\ 0 \; otherwise \end{cases}$

Table A.3: Parameters for profit calculations.

Description	Value	Units
Width of work piece	500	Mm
Radial depth	5	Mm
Tool diameter	10	Mm
Tool change time	4	Sec
Number of teeth	4	-
Machining cost per unit time	1	$/min
Cost of tool	114	$

Table A4: Inputs for FRF based stability limit prediction.

Parameter	**Value**	**Units**
Stiffness	$1.6x10^7$	N/m
Damping ratio	0.05	-
Natural frequency	2400	Hz
Number of teeth	4	-
Radial immersion	50%	-
Tangential cutting coefficient	3000±1000	N/mm^2
Radial cutting coefficient	900±300	N/mm^2