Cognitive Models for Learning to Control Dynamic Systems

Air Force STTR
Phase I Final Technical Report

Applied SR Technologies Inc.
1953, 68th Street
Brooklyn, NY 11204, USA
Tel: (718) 232 9243
FAX: (718) 232 9243
E-mail: zjang@control.poly.edu

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
26092008	final report	15 Dec 07 through 14 Sep 08

4. TITLE AND SUBTITLE
Cognitive Models for Learning to Control Dynamic Systems

5a. CONTRACT NUMBER
FA9550-08-C-00044

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Jiang, Zhong-Ping

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Polytechnic Institute of New York University (formerly Polytechnic University)
6 Metrotech Center, Brooklyn, NY 11201.

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
USAF, AFRL DUNS 143574726
AF Office of Scientific Research
875 N. Randolph ST. Room 3112
Arlington VA 22203
Timothy R. Harr (703) 696-5906

10. SPONSOR/MONITOR'S ACRONYM(S)
AFOSR

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
AFRL-SR-AR-TR-08-0537

12. DISTRIBUTION/AVAILABILITY STATEMENT
Distribution A: Approved for Public Release

13. SUPPLEMENTARY NOTES

14. ABSTRACT
Wald's sequential probability ratio test (SPRT) model and the equivalent discrete drift diffusion model have been widely used to explain human and animal decision making in psychophysical tasks. These models assume that observers gradually accumulate evidence from noisy inputs and make a decision when the evidence reaches a threshold. It is discovered that stochastic-resonance (SR) like behavior arises in the SPRT model when the actual input signal is significantly weaker than anticipated by the model.
Analytical expressions and conditions for the SR-like behavior are found. Therefore appropriate amount of noise can improve the decision making process when the input signal is significantly weaker than anticipated. Research in adaptive SPRT demonstrates that there is an optimal nonzero and finite weight to achieve the best accuracy when the real distributions are wider than the prior distribution.

15. SUBJECT TERMS
Cognitive models, decision making, SPRT, stochastic resonance, dynamic control system

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
			UU	32	Zhong-Ping Jiang
					19b. TELEPHONE NUMBER *(Include area code)* (718) 2603646

Contents

1 Technical Abstract

In this final technical report for the phase I Air Force STTR contract FA9550-08-C-0044, the technical objectives, work accomplished, results, and technical feasibility are summarized. The first and primary objective of this research is to systematically study the role of noise in human decision making. We achieve this objective by showing that stochastic resonance (SR) like behavior arises in the Wald's sequential probability ratio test (SPRT) model when the actual input signal is significantly weaker than anticipated by the model. We derive expressions for calculating the fraction of correct responses, the mean decision time, and the reward rate for the SPRT/DDD (discrete drift diffusion) model of decision making. We then demonstrate that both the fraction of correct responses and the reward rate have a peak as a function of noise strength while the mean decision time is a monotonically increasing function of the noise strength. We also examine the dependence of our results on the initial condition and the form of the input probability distributions. Finally, to gain analytical insights, we consider the continuous time limit of the SPRT/DDD model. We show that the closed-form expressions from the resulting continuous drift diffusion (CDD) model help us understand the SR-type behavior in the SPRT/DDD model but there are also important differences between the discrete and continuous time cases. Therefore, appropriate amount of noise can improve the decision making process when the input signal is weak.

The second objective of the project is to study adaptive SPRT. That is, we study the role of prior distribution in an adaptive SPRT algorithm. One simple adaptive SPRT algorithm without prior distribution is first presented. The simple algorithm is based on the estimation of the real distributions. The obtained result shows that the algorithm makes many quick and wrong decisions and sacrifices accuracy. Then, we use a wide prior distribution to improve the robustness of the algorithm. The new algorithm combines the prior distribution and the information from the input data to estimate the real distributions. The simulation result shows improvement in terms of accuracy while the response time remains moderate. We also plot accuracy and reward rate against the weight of the prior probability. It is observed that there exists an optimal non-zero and finite weight to achieve the best accuracy when the real distributions are wider than the prior distribution. Finally, we examine the evolving performance of this adaptive SPRT algorithm with prior distribution when feedback is provided.

A third but secondary objective of the project is to exploit the role of dynamic systems and control theories in decision making. We propose an extension of Busemeyer's decision field theory (DFT) by means of the contraction mapping principle. It is believed that such an extended DFT will assist operators in making better decision in supervisory control situations where manual control and automation coexist.

2 Introduction: Team Members and Phase I Publication Summary

2.1 Project Members

Principal Investigator: Zhong-Ping Jiang, Professor and IEEE Fellow
Department of Electrical and Computer Engineering, Polytechnic Institute of New York University, Brooklyn.

Dr. Jiang is in charge of all aspects of this project. His expertise is in nonlinear control theory, nonlinear dynamics, optimization and their applications.

Project Manager: Xiaoming Zhuang, President
Applied SR Technologies Inc., Brooklyn, NY

President Zhuang is in charge of the coordination of this project.

Academic Researcher: Xingxing Wu, Ph.D and part-time researcher
Department of Electrical and Computer Engineering, Polytechnic Institute of NYU, Brooklyn, NY

Under the guidance of Professor Jiang, Xingxing plays an active role in the research efforts on the theory of stochastic resonance and applications in signal and image processing. He also helps graduate students on the SR-related issues.

Visiting Scholar: Jerome Busemeyer, Professor
Psychology Department, Indiana University

Dr. Busemeyer provides general comments on the research related to his DFT and how dynamic system (and control) theory is useful in developing computational and mathematical models for human decision making.

Consultant: Ning Qian, Associate Professor
Department of Neuroscience, Columbia University, 1051 Riverside Dr., New York.

Dr. Qian plays a role in the research using his expertise in neuroscience and detection theory, in particular his prior work on the development of models for stochastic-resonance-type behavior in sensory perception.

Research Students: Xiao Han, Feng Ma, Shiyun Xu, Juan Zhang

These graduate students help us with various tasks such as computer simulations using Matlab and C++, extensive literature review, preparation of technical reports.

2.2 Publications

Our Phase I research has gained significant progress. This has lead to the following publications:

- S. Xu, Z. P. Jiang, L. Huang and Daniel W. Repperger, "Control-oriented approaches in dynamic decision making," Proceedings of the 8th WSEAS Int. Conference on Robotics, Control and Manufacturing Technology, April 2008, Hangzhou, pp. 138-146.

- X. Han, Z.P. Jiang, D. Repperger and N. Qian, Stochastic-resonance-like behavior in Wald's sequential probability ratio test model for decision making, to be submitted to *Neural Computation*.

- X. Han, Z.P. Jiang and N. Qian, The role of prior distribution in an adaptive SPRT algorithm. Under preparation.

3 Technical Objectives

The primary phase I research and development objectives are as follows:

1. Investigation of stochastic resonance (SR) behavior in the sequential probability ratio test (SPRT) model for decision making. Stochastic resonance can enhance signal detection by improving certain performance measures such as signal-to-noise ratio, reward rate and fraction of correct responses. One of the project tasks is to demonstrate the stochastic resonance phenomenon in the SPRT model by taking advantage of noise. As is well-known, the SPRT model and the equivalent discrete drift diffusion (DDD) model have been widely used to explain human and animal decision making in psychophysical tasks. These models assume that observers gradually accumulate evidence from noisy inputs and make a decision when the evidence reaches a threshold. In this report, we show that stochastic-resonance type behavior arises in the SPRT model when the actual input signal is significantly weaker than anticipated by the model. Specifically, we derive expressions for calculating the fraction of correct responses, the mean decision time, and the reward rate for the SPRT model of decision making. We then demonstrate that both the fraction of correct responses and the reward rate have a peak as a function of noise strength while the mean decision time is a monotonically increasing function of noise strength.

2. Differences between the predictions of the CDD model and the DDD model in dynamic decision making. Choice of which model for decision making depends on what type of questions we want to answer. We propose to develop mathematically rigorous analysis to illustrate the key differences between the predictions of continuous- and discrete-time diffusion models known as CDD and DDD. We have made some interesting observations which are consistent with MATLAB-based computer simulations. First, if the SPRT/DDD model selects the thresholds according to the equations from the CDD model, then CR (the mathematical expectation of accuracy) would be better than desired, but at the price of longer response time than the anticipation of the CDD model. Second, the response time always exhibits a peak which is not predicted by the CDD model. As a result, if the actual noise strength is to the right of the peak, then a moderate amount of extra noise could possibly reduce the response time without violating the constraints on the Type I and Type II errors, due to the first difference. Third, the tail of the reward rate curve drops significantly with stronger noise, while the CDD model predicts constant reward rate. Such a result may be important for psychological experiments.

3. Effects of non-zero initial condition and extensions to other forms of distributions. For the purpose of testing the robustness of the SR-like behavior, we look at the CR behavior when the initial diffusion position is biased (because of asymmetric prior probability) or asymmetric constraints on Type I and Type II errors. The simulation results show that the SR-like behavior gets more evident as initial bias moves toward the upper threshold and weaker if the initial bias is negative. Numerical studies are also run on other forms of distributions such as Gamma distribution. For the Gamma distribution with wrongly assumed shape parameter K, it is observed that SR-like behavior also occurs. Unlike the Gaussian case, the peak gets higher and narrower with larger Z but the position of the peak does not seem to move much with Z. Like the Gaussian case, the response time also exhibits a peak and gets longer almost linearly with increasing Z and the reward rate also exhibits a peak and gets smaller with increasing Z.

4. Adaptive SPRT. We study the role of prior distribution in dynamic decision making via an adaptive SPRT algorithm. The ultimate goal is to use a wide prior distribution to improve the robustness of the algorithm. The key strategy is to develop a new algorithm that combines the prior distribution and the information from the input data to estimate the real distributions. In addition, we examine the evolving performance of this new adaptive SPRT algorithm with prior distribution when feedback is provided.

5. Applications of control and dynamic system theories to the development of mathematical and computational models for human dynamic decision making. It is widely recognized that feedback and control play a crucial role in human decision making. We have obtained some control-theoretic results on dynamic decision making using the celebrated contraction mapping principle. It is shown that an extension of Busemeyer's Decision Field Theory (DFT), named as Generalized DFT, can be proposed to solve supervisory control problems involving manual control and automation. An application to the benchmark example of sugar factory task yields promising results. The obtained results have been presented at the 8th WEAS International Conference ROCOM (April 2008, pp. 138-146, ISBN 978-960-6766-51-0), and an expanded version has been selected for publication in a special issue of a Springer journal.

4 Technical Approach and Accomplished Work

4.1 Background

Stochastic resonance (SR) occurs when output signal-to-noise ratio in a nonlinear system is maximized by a moderate level of noise. The term is also used generally to describe any phenomena where noise plays a positive role. The SR was first observed in the physics literature (Benzi et al. [1]), and then widely studied in other fields of sciences and engineering. Gammaitoni, et al. have an in-depth review on SR [2]. The SR-like behavior has also been well documented in psychophysical experiments. For example, Collins and his co-workers [11, 12] found that human detection of weak tactile stimulus can be enhanced by adding a certain amount of noise. Gong et al. [13] have proposed a theory that explains Collins et al.'s data very well.

An integral component of any psychophysical experiment is decision making. In the fixed-time (or interrogation) paradigm used in many experiments, including those of Collins et al. [11, 12], a stimulus is presented for a fixed duration in each trial, and an observer has to decide, for example, whether the stimulus is noise or signal plus noise. However, a different approach, called the reaction-time (or free-response) paradigm, is often employed when one wants to study the dynamical decision making process more explicitly. In this paradigm, the stimulus presentation in each trial is terminated by the observer when he or she feels ready to make a decision. Typically, observers earn points for correct decisions and lose points for wrong ones. Since longer time is needed to make more accurate decisions, observers have to compromise between speed and accuracy to maximize the total points per unit time (reward rate).

Wald's SPRT model and the equivalent DDD model [3, 4] provide a natural framework for understanding decision making in the reaction-time paradigm. These models assume that observers gradually accumulate evidence from noisy inputs and make a decision when the evidence reaches a threshold. In most psychophysical experiments, observers have to decide between two alternatives (e.g., presence or absence of a signal, leftward or rightward motion). This maps naturally to the upper and lower thresholds in the models. The models choose one or the other alternative depending on whether the accumulated evidence reaches the upper or the lower threshold. The models have been applied to a wide range of psychophysical experiments; see [8] for a recent review.

While there are both experimental and theoretical studies of the SR-like behavior in the fixed-time paradigm [11, 12, 13], to our knowledge no SR studies have been done for the reaction-time paradigm and the associated models for decision making. The focus of this paper is to demonstrate SR-like behavior in the SPRT/DDD model, and in its continuous time limit, the CDD model. Previous applications of these models [7] often assume that the decision-making process has perfect knowledge of the probability distributions from which inputs are drawn. This is, however, unlikely to be true. Instead, the brain must rely on its prior experience over a long period of time to estimate the probability distributions of the inputs. Consequently, the estimated distributions must be different from the actual ones arbitrarily picked by experimenters for a given psychophysical experiment. In particular, when the input signals are very weak, the decision process must incorrectly assume that the signals are drawn from distributions containing much stronger signals. One of our research goals is to show that this is precisely the situation where the SR-like behavior emerges.

4.2 Stochastic Resonance Behavior in Wald's Sequential Probability Ratio Test (SPRT)

4.2.1 SPRT and Its Drift Diffusion Models

Take as an example the two-alternative forced choice (2AFC) task in the reaction time paradigm. The Wald's SPRT model for decision making is based on the probability distributions of the input x under hypotheses H_0 and H_1:

$$p(x|H_n), \quad n = 0, 1$$

For each new sampled input x_k, the model accumulates evidence by updating the running product of likelihood ratios:

$$L(k_0) = \prod_{k=1}^{k_0} \frac{p(x_k|H_1)}{p(x_k|H_0)}$$

and compares the product with a lower threshold $\frac{1-p(H_1)}{p(H_1)} \cdot \frac{1-P}{1-\alpha}$ and an upper threshold $\frac{1-p(H_1)}{p(H_1)} \cdot \frac{P}{\alpha}$, where $p(H_1)$ is the prior probability of H_1, P is the power of the test and α is the significance level of the test. The thresholds [4] are determined so that the probability of Type I error (H_1 is correct but the model chooses H_0) is $1 - P$ and the probability of Type II error (H_0 is correct but the model chooses H_1) is α. If the product is smaller than the lower threshold, the model decides that H_0 is correct (This decision is called D_0). If the product is larger than the upper threshold, then the model decides that H_1 is correct (This decision is called D_1). If neither condition is satisfied, the model continues to sample the input and update the product until one of the thresholds is reached. The SPRT model is optimal (Wald [3, 4]) in the sense that for fixed constraints on type I and type II errors, no other test allows a shorter decision time on average.

By taking logarithm, we can convert the product of likelihood ratios into a sum of log likelihood ratios

$$Y(k_0) = log\left(L(k_0)\right) = \sum_{k=1}^{k_0} \log\left(\frac{p(x_k|H_1)}{p(x_k|H_0)}\right) = \sum_{k=1}^{k_0} s_k \tag{1}$$

which is referred to as the discrete drift diffusion (DDD) model, with the new diffusion boundaries Z_0 and Z_1:

$$Z_0 = \log\left(\frac{1-p(H_1)}{p(H_1)} \cdot \frac{1-P}{1-\alpha}\right), \quad Z_1 = \log\left(\frac{1-p(H_1)}{p(H_1)} \cdot \frac{P}{\alpha}\right) \ . \tag{2}$$

Note that s_k is the diffusion steps while $Y(k_0)$ is the diffusion position after k_0 steps.

On the other hand, when the diffusion steps are infinitesimally small, the SPRT/DDD model converges to the CDD model which is described by the following stochastic equation

$$dy = Adt + cdW; \quad y(0) = y_0. \tag{3}$$

where A is the drift per unit time, W is standard Weiner process with expected value 0 and variance 1, c is the standard deviation of the Weiner process, and $y(0)$ is the initial bias. We will first consider the case of zero initial bias and then discuss the case of nonzero initial conditions.

The CDD parameters A and c correspond to the mathematical expectation and standard deviation of a single diffusion step s in the DDD model. To be specific, and use A_n to denote the drift rate under hypothes H_n,

$$A_n = \frac{(2\mu_n - \hat{\mu}_0 - \hat{\mu}_1)(\hat{\mu}_1 - \hat{\mu}_0)}{2\hat{\sigma}^2}, \quad c = \frac{\sigma(\hat{\mu}_1 - \hat{\mu}_0)}{\hat{\sigma}^2}, \quad n = 0, 1 \tag{4}$$

4.2.2 Some Technical Assumptions

DC signal and signal-independent Gaussian noise: We assume that the signal is DC and the noise is Gaussian and signal-independent. This is justified when the two hypotheses are symmetric such as leftward motion vs. rightward motion. It can also be justified for asymmetric hypotheses such as "noise vs. signal plus noise" as long as the signal is very weak.
To be specific, the actual distributions from which the inputs are drawn under H_0 and H_1 are given by:

$$\begin{array}{rcl} p(x|H_0) & \sim & N(\mu_0, \sigma^2) \\ p(x|H_1) & \sim & N(\mu_1, \sigma^2) \end{array}$$

where $N(\mu, \sigma^2)$ denotes Gaussian distribution with mean μ and variance σ^2.
SPRT parameters inaccurate and non-adaptive: Since the brain does not have direct access to these distributions and has to rely on prior experience to estimate these distributions, we assume that the distributions used in the model for calculating the likelihood ratios are different:

$$\begin{array}{rcl} \hat{p}(x|H_0) & \sim & N(\hat{\mu}_0, \hat{\sigma}^2) \\ \hat{p}(x|H_1) & \sim & N(\hat{\mu}_1, \hat{\sigma}^2) \end{array}$$

Throughout this paper, we will analyze the non-adaptive SPRT model. The SPRT parameters are assumed by the SPRT model prior to the decision making process and does not adapt to the input data.
Symmetry: We further assume that the two hypotheses are equally probable, that is,

$$p(H_0) = p(H_1) = 0.5$$

and the constraints on Type I and Type II errors are symmetric, that is,

$$P + \alpha = 1$$

Following these assumptions, the boundaries (2) for the diffusion process are symmetric with respect to zero:

$$Z_0 = log\left(\frac{1-P}{P}\right), \quad Z_1 = log\left(\frac{P}{1-P}\right) \tag{5}$$

We define the distance from the boundaries to zero to be Z, i.e. $Z = Z_1 = -Z_0$. As a result, we have

$$P = \frac{e^Z}{1+e^Z} \quad \text{and} \quad \alpha = \frac{1}{1+e^Z}, \quad \text{where} \quad Z > 0. \tag{6}$$

Under these assumptions, a single diffusion step in the DDD model is:

$$\begin{aligned} s_k &= \log\left(\frac{\hat{p}(x_k|H_1)}{\hat{p}(x_k|H_0)}\right) = \log\left(\frac{e^{-(x_k-\hat{\mu}_1)^2/2\hat{\sigma}^2}}{\sqrt{2\pi\hat{\sigma}^2}} \Bigg/ \frac{e^{-(x_k-\hat{\mu}_0)^2/2\hat{\sigma}^2}}{\sqrt{2\pi\hat{\sigma}^2}}\right) \\ &= \frac{2x_k(\hat{\mu}_1-\hat{\mu}_0)+\hat{\mu}_0^2-\hat{\mu}_1^2}{2\hat{\sigma}^2} . \end{aligned}$$

Since the input x_k is Gaussian distributed, and s_k is linearly related to x_k, s_k is also Gaussian distributed. The probability distributions of s_k under H_0 and H_1 are given by:

$$\begin{aligned} p(s_k|H_0) &\sim N\left(\tfrac{(2\mu_0-\hat{\mu}_0-\hat{\mu}_1)(\hat{\mu}_1-\hat{\mu}_0)}{2\hat{\sigma}^2}, (\tfrac{\sigma(\hat{\mu}_1-\hat{\mu}_0)}{\hat{\sigma}^2})^2\right), \\ p(s_k|H_1) &\sim N\left(\tfrac{(2\mu_1-\hat{\mu}_0-\hat{\mu}_1)(\hat{\mu}_1-\hat{\mu}_0)}{2\hat{\sigma}^2}, (\tfrac{\sigma(\hat{\mu}_1-\hat{\mu}_0)}{\hat{\sigma}^2})^2\right) . \end{aligned} \tag{7}$$

4.2.3 Expressions for the performance measures of the SPRT/DDD model

To demonstrate the SR-like behavior in the SPRT/DDD model, we need to derive the expressions for the fraction of correct responses, the mean decicion time, and the reward rate of the SPRT/DDD model:

$$CR = p(H_0)\cdot p(D_0|H_0) + p(H_1)\cdot p(D_1|H_1) , \tag{8}$$

$$\langle DT\rangle = p(H_0)\cdot \langle DT_{H_0}\rangle + p(H_1)\cdot \langle DT_{H_1}\rangle , \tag{9}$$

$$RR = \frac{CR - q\cdot ER}{\langle DT\rangle + T_0} . \tag{10}$$

According to equations (8), (9), (10), we only need to derive the expressions for the first passage probability $p(D_1|H_n)$ and the mean decision time $\langle DT_{H_n}\rangle$.

By decomposition, we have

$$p(D_1|H_n) = \sum_{k_0=1}^{\infty} p(D_1, DT{=}k_0|H_n), \quad n = 0, 1$$

$$\langle DT_{H_n}\rangle = \sum_{k_0=1}^{\infty} k_0 \cdot p(DT{=}k_0|H_n), \quad n = 0, 1$$

and

$$p(DT{=}k_0|H_n) = p(D_1, DT{=}k_0|H_n) + p(D_0, DT{=}k_0|H_n)$$

Thus, to determine all the performance measures, we only need to calculate

$$p(D_1, DT{=}k_0|H_n) \quad \text{and} \quad p(D_0, DT{=}k_0|H_n), \quad n = 0, 1$$

i.e. the probability of passing the upper boundary at time k_0 and the probability of passing the lower boundary at time k_0.

We will use Figure 1 to illustrate the derivation. We will also use $g_{H_n}(y)$ to denote the PDF of a single diffusion step, where $n = 0, 1$.

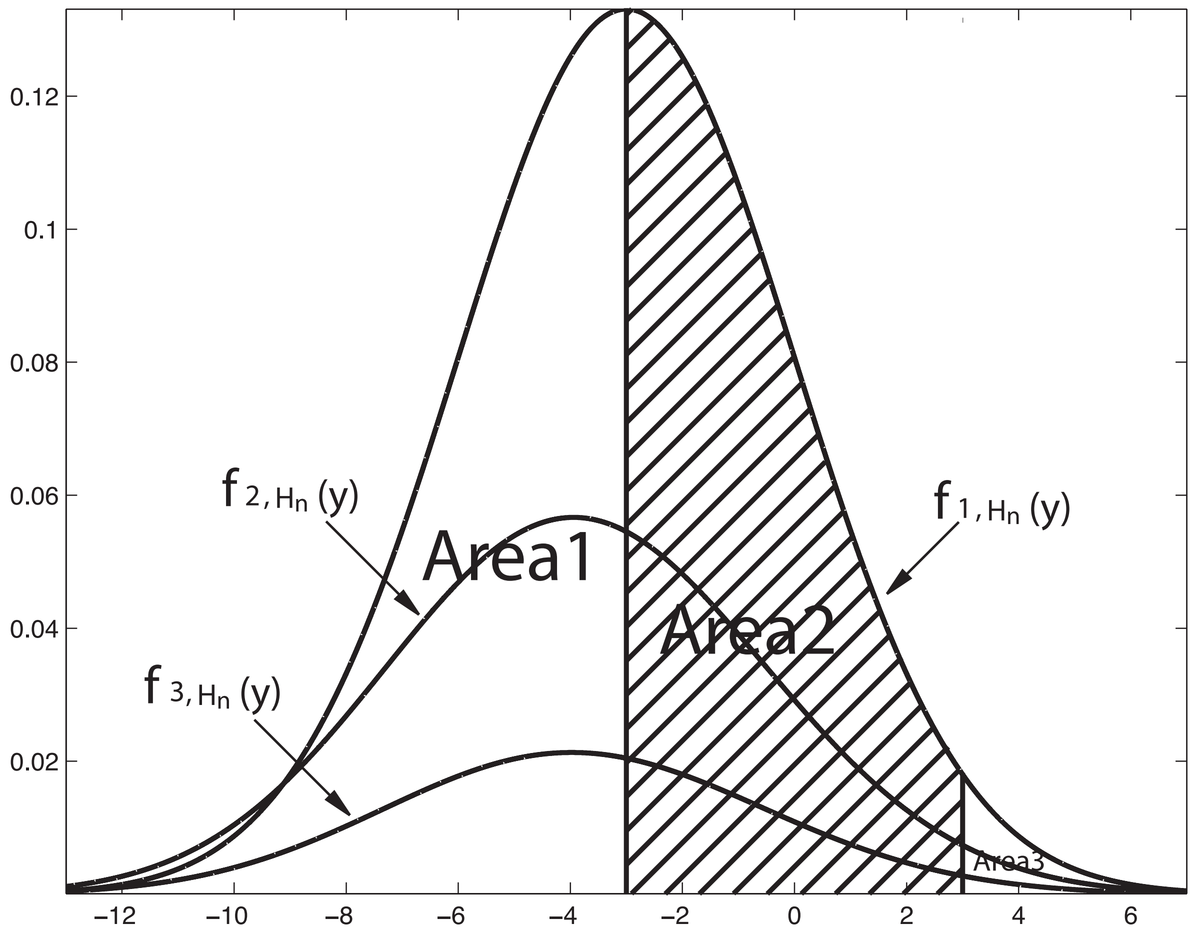

Figure 1: Functions $f_{k_0,H_n}(y)$, with $Z = 3$.

Apparently, the PDF of $Y(1)$ is simply the PDF of a single diffusion step, $g_{H_n}(y)$. We define $f_{1,H_n}(y)$ as $g_{H_n}(y)$. As a result, the probability of passing the upper boundary at time $k_0 = 1$ is the area3 in $f_{1,H_n}(y)$.

If the first step of diffusion reached neither boundaries ($-Z$ and Z), then with this further information, the conditional PDF of $Y(1)$ (We will denote this condition as $DT > 1$) is the part2 of $f_{1,H_n}(y)$ multiplied by a normalization factor.

We will denote the part2 of $f_{1,H_n}(y)$ as $T(f_{1,H_n}(y))$ where $T(\cdot)$ is a truncation operator which takes out the part of a function between $-Z$ and Z.

By definition, $Y(k_0) = Y(k_0 - 1) + s_{k_0}$, where $Y(k_0 - 1)$ and s_{k_0} are independent random variables, thus the PDF of $Y(2)$ is the convolution of $T(f_{1,H_n}(y))$ and $g_{H_n}(y)$ multiplied by a normalization factor.

If we define $f_{2,H_n}(y)$ as $T(f_{1,H_n}(y)) * g_{H_n}(y)$ where $*$ denotes convolution, then the PDF of $Y(2)$ is $f_{2,H_n}(y)/\int_{-\infty}^{\infty} f_{2,H_n}(y)\,dy$.

Generally, we will define

$$f_{k_0+1,H_n}(y) = T(f_{k_0,H_n}(y)) * g_{H_n}(y) \quad k_0 \geq 1 \tag{11}$$

Proposition 4.1: $p(DT \geq k_0|H_n) = \displaystyle\int_{-\infty}^{\infty} f_{k_0,H_n}(y)$.

Proof: Apparently, we have $p(DT \geq 1|H_n) = 1 = \displaystyle\int_{-\infty}^{\infty} f_{1,H_n}(y)$. Generally, if $p(DT \geq k_0|H_n) =$

$\int_{-\infty}^{\infty} f_{k_0,H_n}(y)$ for a certain k_0, then $p(DT \geq k_0+1|H_n)$ is simply the area2 of $f_{k_0,H_n}(y)$. To be explicit,

$$p(DT \geq k_0+1|H_n) = \int_{-Z}^{Z} f_{k_0,H_n}(y) = \int_{-\infty}^{\infty} f_{k_0+1,H_n}(y)$$

Proposition 4.2: The conditonal PDF of $Y(k_0)$ (with condition $DT \geq k_0$) is $f_{k_0,H_n}(y)/\int_{-\infty}^{\infty} f_{k_0,H_n}(y)\,dy$.

Proof: The conditional PDF of $Y(1)$ (with condition $DT \geq 1$) is apparently

$$f_{1,H_n}(y) \Big/ \int_{-\infty}^{\infty} f_{1,H_n}(y)\,dy = g_{H_n}(y)$$

Generally, if the conditional PDF of $Y(k_0)$ is $f_{k_0,H_n}(y)/\int_{-\infty}^{\infty} f_{k_0,H_n}(y)\,dy$, and the k_0th step of diffusion did not reach any boundaries ($-Z$ and Z), neither, thus the condition changes from $DT \geq k_0$ to $DT \geq k_0+1$, and the conditional PDF of $Y(k_0+1)$ is

$$\begin{aligned} &C_{k_0} \cdot T\left(f_{k_0,H_n}(y) \Big/ \int_{-\infty}^{\infty} f_{k_0,H_n}(y)\,dy\right) * g_{H_n}(y) \\ &\qquad = \left(C_{k_0} \Big/ \int_{-\infty}^{\infty} f_{k_0,H_n}(y)\,dy\right) \cdot f_{k_0+1,H_n}(y) = N_{k_0+1} \cdot f_{k_0+1,H_n}(y) \end{aligned} \tag{12}$$

where C_{k_0} and N_{k_0+1} are normalization factors. The normalization factor N_{k_0+1} is apparently $1/\int_{-\infty}^{\infty} f_{k_0+1,H_n}(y)\,dy$ to ensure the total area of a PDF is 1. Thus the conditional PDF of $Y(k_0+1)$ is $f_{k_0+1,H_n}(y)/\int_{-\infty}^{\infty} f_{k_0+1,H_n}(y)\,dy$.

Proposition 4.1.3: The probability of passing the upper boundary at time k_0 is $\displaystyle\int_{Z}^{\infty} f_{k_0,H_n}(y)$.

Proof: In other words, this is the probability that the first k_0-1 steps reached neither boundaries and the k_0th step crosses the upper boundary, to be specific, "$DT \geq k_0$ and $Y(k_0) \geq Z$".

$$\begin{aligned} &p(D_1, DT{=}k_0|H_n) = p(Y(k_0) \geq Z|DT \geq k_0, H_n) \cdot p(DT \geq k_0|H_n) \\ &\qquad = \frac{\int_{Z}^{\infty} f_{k_0,H_n}(y)}{\int_{-\infty}^{\infty} f_{k_0,H_n}(y)} \cdot \int_{-\infty}^{\infty} f_{k_0,H_n}(y) = \int_{Z}^{\infty} f_{k_0,H_n}(y) \end{aligned} \tag{13}$$

4.2.4 Condition for the SR-like behavior in the CDD model

The CDD model predicts that, under the above-mentioned assumptions,

$$p(D_0|H_0) = 1 - \frac{1}{1+e^{2A_0Z/c^2}}, \quad p(D_1|H_1) = \frac{1}{1+e^{2A_1Z/c^2}} \tag{14}$$

$$\langle DT_{H_0}\rangle = \frac{Z}{A_0} tanh\left(\frac{A_0Z}{c^2}\right), \quad \langle DT_{H_1}\rangle = \frac{Z}{A_1} tanh\left(\frac{A_1Z}{c^2}\right)$$

As a result, (according to (8) and (9)),

$$CR = \frac{1}{2} - \frac{1}{2}\left(\frac{1}{1+e^{2A_0Z/c^2}} - \frac{1}{1+e^{2A_1Z/c^2}}\right) \tag{15}$$

$$\langle DT\rangle = \frac{Z}{2A_0} tanh\left(\frac{A_0Z}{c^2}\right) + \frac{Z}{2A_1} tanh\left(\frac{A_1Z}{c^2}\right) \tag{16}$$

Under the assumptions in Subsection 4.2.2, we propose that the sufficient and necessary condition for the SR-like behavior is

$$A_0 \cdot A_1 > 0 \quad \text{and} \quad A_0 < A_1 \tag{17}$$

Specifically, CR, as a function of σ, has a global maximum at a non-zero and finite σ.

The above condition is equivalent to (according to (4)),

$$\mu_0 < \mu_1 < \frac{\hat{\mu}_0+\hat{\mu}_1}{2}, \quad \text{or} \quad \frac{\hat{\mu}_0+\hat{\mu}_1}{2} < \mu_0 < \mu_1 \tag{18}$$

i.e. the signal under H_1 is much weaker than anticipated or the signal under H_0 is much stronger than anticipated.

4.3 Adaptive SPRT

The presence of SR-like behavior in Wald's SPRT model can be viewed as an indication of the mismatch between the actual distribution and the assumed input distribution that initiates the adaptive process. Indeed, it is possible that the decision-making process in the brain gradually adapts its internally assumed input distributions to match the actual distributions if feedback is provided. There has been a previous study on adaptive SPRT [19]. However, that study does not assume any prior probability distributions and attempts to estimate the distributions from the available input data alone. Therefore, the method may not be robust when the input data are few.

In order to study the role of prior distribution in adaptive SPRT, we proceed with two cases: adaptive SPRT without prior distribution and SPRT with prior distribution. First, let us assume that the input distributions under H_0 and H_1 are $\mathcal{N}(\mu_0, \sigma_0^2)$ and $\mathcal{N}(\mu_1, \sigma_1^2)$. We want the SPRT algorithm to decide between H_0 and H_1, with the objective of minimizing the response time while achieving a desired accuracy described by power P and significant level α.

Adaptive SPRT algorithm without prior distribution: First pass. In this case, the algorithm finds the unbiased MMSE estimation of the parameters μ_0, σ_0^2, μ_1, σ_1^2 using the following formulas:

$$\hat{\mu} = \frac{1}{N}\sum_{i=1}^{N} X_i, \quad \hat{\sigma}^2 = \frac{1}{N-1}\sum_{i=1}^{N}(X_i - \bar{X})^2$$

It turns out that

$$\hat{\mu} \sim \mathcal{N}(\mu, \frac{\sigma^2}{N}), \quad \hat{\sigma}^2 \sim \Gamma\left(\frac{N-1}{2}, \frac{2\sigma^2}{N-1}\right)$$

with $E(\hat{\mu}) = \mu$, $Var(\hat{\mu}) = \dfrac{\sigma^2}{N}$, $E(\hat{\sigma}^2) = \sigma^2$, $Var(\hat{\sigma}^2) = \dfrac{2\sigma^4}{N-1}$.

For each step of diffusion, the nth step, for example, the model estimates the model parameters based on all the input samples except the current sample, i.e. $x_1, x_2, \ldots, x_{n-1}$. The diffusion size is then computed using the new model parameters $\hat{\mu}_0$, $\hat{\sigma}_0^2$, $\hat{\mu}_1$, $\hat{\sigma}_1^2$.

Note that if we only make one decision, then we do not know whether the input samples are drawn from $p(x|H_0)$ or $p(x|H_1)$. As a result, we can only adapt the common parameters of $p(x|H_0)$ and $p(x|H_1)$. For example, if $\sigma_0 = \sigma_1 = \sigma$, then we can estimate σ using the input samples without knowing whether they are drawn from $p(x|H_0)$ or $p(x|H_1)$.

Next, we will study a simple case when $\mu_0 = 0$ and $\mu_1 = \mu$ is known and we estimate only $\sigma = \sigma_0 = \sigma_1$. In this case, the unbiased MMSE estimation becomes:

$$\hat{\sigma}^2 = \frac{1}{N}\sum_{i=1}^{N}(X_i - \mu)^2$$

where $\hat{\sigma}^2$ follows Chi-square distribution and the variance is $\frac{2}{N} \cdot \sigma^2$.

The performance of this simple case is plotted in Figure 2 and Figure 3. The desired accuracy is 0.95. This simple adaptive SPRT is more robust to noise variance but achieves only about 0.83 accuracy. As anticipated, the response time appears quadratically dependent on σ.

Adaptive SPRT algorithm with prior distribution: Second pass Suppose that we want to estimate a random variable α using observations from sensor 1 and sensor 2. The two observations α_1 and α_2 have variances σ_1 and σ_2, respectively. It turns out that the MMSE estimation of α is

$$\alpha = \frac{C_1 \cdot \alpha_1 + C_2 \cdot \alpha_2}{C_1 + C_2}$$

where $C_1 = 1/\sigma_1^2$ and $C_2 = 1/\sigma_2^2$.

We will use $\mathcal{N}(\mu_{0p}, \sigma_{0p}^2)$ and $\mathcal{N}(\mu_{1p}, \sigma_{1p}^2)$ to denote the prior distributions. The estimation from the input data are $\mathcal{N}(\mu_{0d}, \sigma_{0d}^2)$ and $\mathcal{N}(\mu_{1d}, \sigma_{1d}^2)$.

To combine $\mathcal{N}(\mu_p, \sigma_p^2)$ and $\mathcal{N}(\mu_d, \sigma_d^2)$ into $\mathcal{N}(\mu, \sigma^2)$, we model the prior distribution $\mathcal{N}(\mu_p, \sigma_p^2)$ as observation from sensor 1 and $\mathcal{N}(\mu_d, \sigma_d^2)$ as observation from sensor 2.

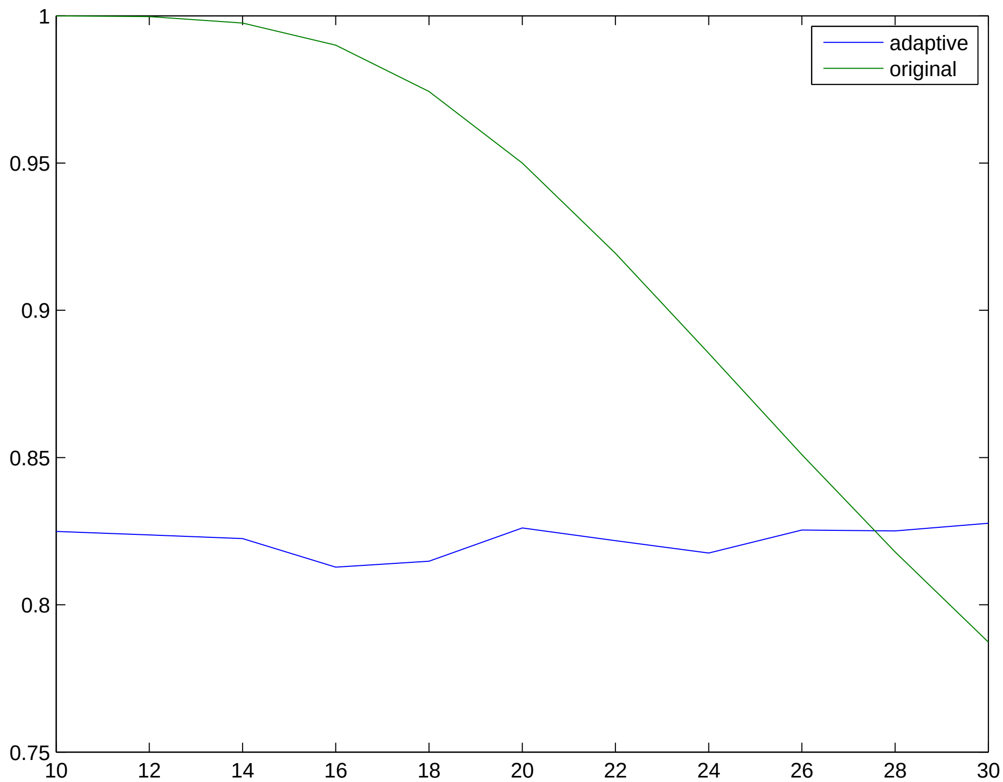

Figure 2: CR versus sigma, No prior distribution, power=0.95, significance=0.05. $\mu = 6$, $\sigma_p = 20$, $\sigma \in \{10, 12, \cdots, 30\}$. Averaged over 10000 trials.

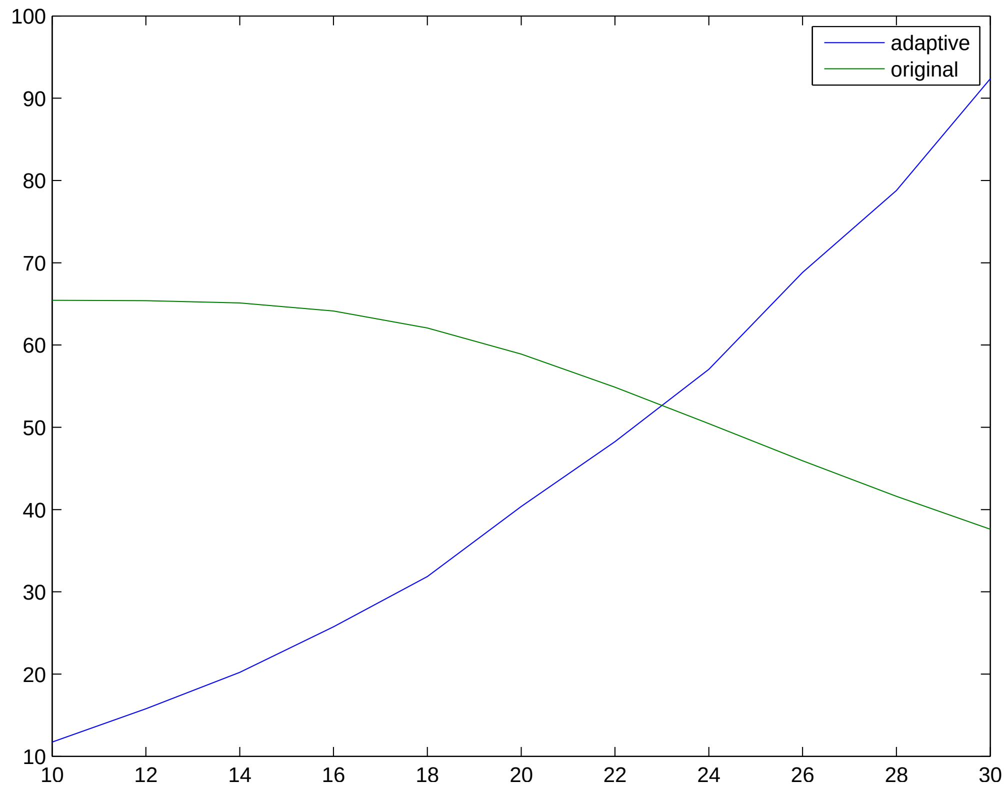

Figure 3: CR versus sigma, No prior distribution, power=0.95, significance=0.05. $\mu = 6$, $\sigma_p = 20$, $\sigma \in \{10, 12, \cdots, 30\}$. Averaged over 10000 trials.

We assume that the prior distribution is estimated from M "virtual inputs". The virtual inputs are drawn from $\mathcal{N}(\mu, \sigma^2)$ but "unfortunately" have mean μ_p and variance σ_p^2.

When μ_0 and μ_1 are unknown, using the unbiased MMSE estimation, the formulas for combination are

$$\mu = \frac{M \cdot \mu_p + N \cdot \mu_d}{M + N}, \quad \hat{\sigma}^2 = \frac{(M-1)\sigma_p^2 + (N-1)\sigma_d^2}{M+N-2} \tag{19}$$

where M is the number of "virtual inputs" and N is the number of "real inputs".

When μ_0 and μ_1 are known, using the unbiased MMSE estimation, the formula for combination is:

$$\hat{\sigma}^2 = \frac{(M)\sigma_p^2 + (N)\sigma_d^2}{M+N} \tag{20}$$

Intuitively, prior distribution dominates if $M > N$, estimation from real input data dominates if $M < N$.

The performance of this adaptive SPRT algorithm (also adapts only σ) with prior distribution is plotted in Figure 4 and Figure 5. The desired accuracy is 0.95. This adaptive SPRT with prior distribution is less robust to noise variance but achieves about 0.95 accuracy. The response time remains moderate compared with the simple adaptive SPRT algorithm without prior distribution.

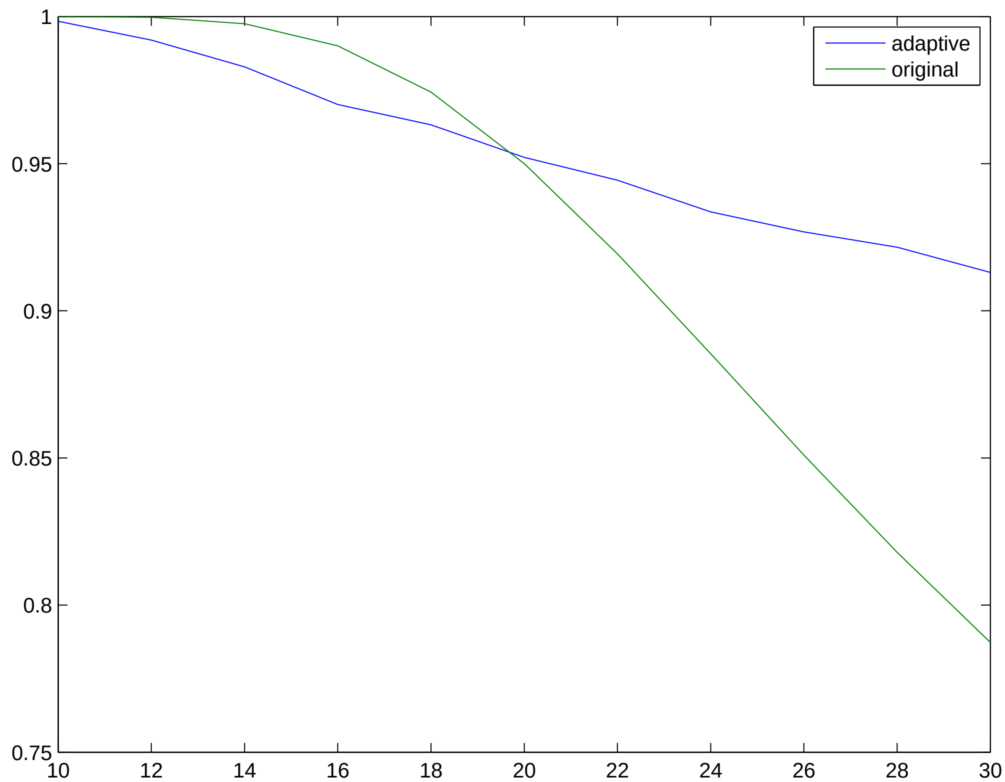

Figure 4: CR versus sigma, with prior distribution, M=10, power=0.95, significance=0.05. $\mu = 6$, $\sigma_p = 20$, $\sigma \in \{10, 12, \cdots, 30\}$. Averaged over 10000 trials.

4.4 Control-Theoretic Approach to Dynamic Decision Making

Dynamic decision making tasks include important activities such as stock trading, air traffic control, and managing continuous production processes. In these tasks, decision makers make multiple recurring decisions

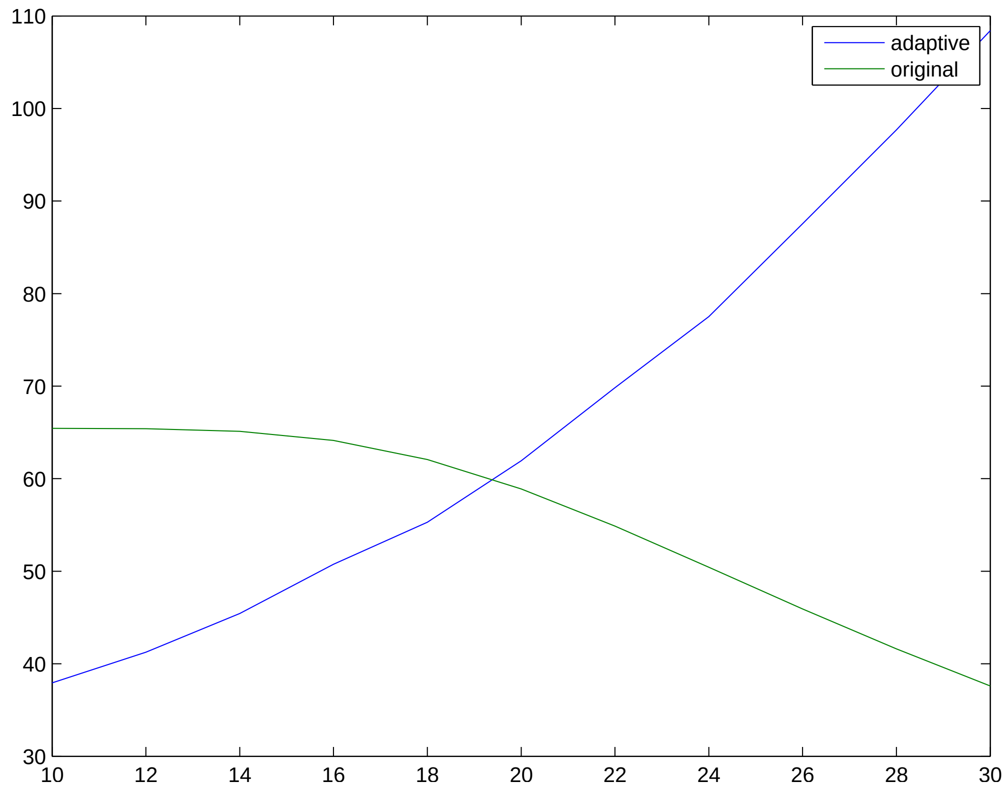

Figure 5: CR versus sigma, with prior distribution, M=10, power=0.95, significance=0.05. $\mu = 6$, $\sigma_p = 20$, $\sigma \in \{10, 12, \cdots, 30\}$. Averaged over 10000 trials.

to reach a target, and they receive feedback on the outcome of their efforts along the way (see [45, 52, 54] and the references therein).

A transfer of insights from other related domains makes it possible to develop a formulation of learning building on the application of control theory to the study of human performance in dynamic decision making [53]. Brehmer uses control theory [37, 38] as a framework to analyze the goal-directed behavior in dynamic decision-making environments. People who use less sophisticated environment models are able to learn to improve their performance only when feedback is timely and continuous [37, 38]. Jordan and Rumelhart [54, 56] address similar issues in the area of motor learning. A key idea of their approach to dynamic decision making is to divide the learning problem into two interdependent subproblems. A broad set of topics including feedback control, feedforward control, delay and learning algorithms are then introduced into this area [55]. Gibson [51] inherits Jordan's connectionist network and applies online learning in parallel distributed processing, or a neural network control model to illustrate the Sugar Factory (SF) Task [50].

The SF model is a simple dynamic decision-making task in which decision makers are expected to learn from experience [35, 50, 44]. It is of interest to computational organization theorists, and there have been various kinds of tests conducted on it. A typical phenomenon arising from these experiments is that while participants progressively improve their capacity to control the system, they remain unable to describe how the system works or how does it reach the target value, leading to large amounts of repetitive work and low efficiency. Upon such backgrounds, an automatic design is required and presented as a reference.

Automation can improve the efficiency and safety of complex and dangerous operating environments by reducing the physical or mental burden on human operators [64]. Despite this fact, it is always a critical distinction whether or not automation is engaged, and the operator's role has to be changed from controllers directly involved with the system to supervisory controllers [58]. In such supervisory control systems, operators monitor the performance of automation during normal operations, and intervene to take manual control when necessary.

Studies have shown that operator's use of automation reflects automation reliability, and inappropriate reliance associated with misuse and disuse partly depends on how well trust matches the true capabilities of the automation [68]. In order to guide design, evaluation and training to enhance human-automation partnerships as well as high specificity of trust are required, and through which misuse and disuse of automation can be mitigated [59]. Consequently, a better operator knowledge of how automation works and the automation design philosophy are both required for more appropriate use of automation [63].

The operator's choice plays such an important role in the automated system performance that the allocation of functions is becoming a critical decision making process, and to optimize this process will be of great importance [60]. A dynamic approach capitalized on the power of the DFT (Decision Field Theory) has been developed to characterize operators' reliance on automation in a supervisory control system by describing a quantitative model of trust in automation, and an EDFT (Extended DFT) model is proposed [49]. As trust and self-confidence are closely associated with the capacity of automation and manual control separately [57, 69], it behooves us to improve the existed model in order to help the operator gain a better understanding of capacities.

Our research is targeted at developing a framework to modify the EDFT method based on the celebrated contraction mapping principle. The result is supported via computation simulations on the benchmark example of Sugar Factory supervisory control scenario.

Due to the complexity and variability of automation performance, the operator's choice between automatic and manual control in supervisory control situations can be considered both a preferential choice problem and a decision-making process described by Decision Field Theory (DFT) [41, 42]. The standard elementary DFT model used to investigate decision making under risk or uncertainty could be described through a straightforward example in supervisory control. Suppose one is facing the problem of choosing whether to rely on automation (A) or to intervene with manual control (M), as shown in the following chart. In Figure 5, S_1 and S_2 are two interdependent uncertain events, one of which may occur at a certain time point. S_1 denotes the occurrence of an automation fault and S_2 represents the incidence of a fault that compromises manual control. During the course of decision making, the valence of an action V_i (i = A or M) is defined as the subjective expected payoff for each action also fluctuates from sample to sample, which is relevant to the subjective probability weight $W(S_j)$ and the utility of the payoff [40]. The preference state

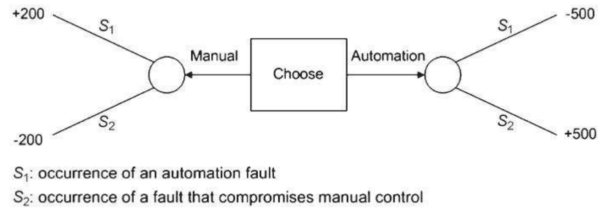

Figure 6: DFT choosing model in a supervisory control situation [49]

at sample n is derived based on the accumulated valence difference:

$$\begin{aligned} P(n) &= (1-s) \times P(n-1) + [V_A(n) - V_M(n)] \\ &= (1-s) \times P(n-1) + [d + \varepsilon(n)]. \end{aligned} \tag{21}$$

Let C represent the true capability of the automation (C_A) or manual control (C_M). The former symbol describes the reliability of the automation in terms of fault occurrence and general ability to accomplish the task under normal conditions, while the latter one describes how well the operator can manually control the system in various situations. B_C denotes the belief or estimation of the capability of automation (B_{CA}) or the operator's manual capability (B_{CM}). In the EDFT model, sequential decision processes are linked by dynamically updating beliefs regarding the capability of automation or manual control based on previous decisions in order to guide the next decision as follows [49]:

$$B_C(n) = B_C(n-1) + \frac{1}{b_1}(C(n-1) - B_C(n-1)). \tag{22}$$

The value b_1 ($b_1 \geq 1$) represents the level of transparency of the system interface, describing how well information is conveyed to the operator when capability information is available. $b_1 = 1$ means the information is perfectly conveyed to the operator. The larger b_1 is, the more poorly information is conveyed to operator.

There is a formulation depicted in [48] that beliefs represent the information base that determines attitudes and then attitudes determine intentions and consequently behaviors. Under the circumstance of supervisory control, trust and self-confidence are both attitudes that depend on beliefs, while at the same time, they determine preference and reliance. Take T and SC as the denotation of trust and self-confidence, which are updated by B_{CA} and B_{CM} as the new input respectively. Preference of A over M is defined as the difference between trust and self-confidence at time step n in the EDFT model, denoted by $P(n)$ [49]:

$$\begin{aligned} P(n) = T(n) - SC(n) = {} & [(1-s) \times T(n-1) \\ & + s \times B_{CA}(n) + \varepsilon(n)] - [(1-s) \times SC(n-1) \\ & + s \times B_{CM}(n) + \varepsilon(n)] = (1-s) \times P(n-1) \\ & + s \times [B_{CA}(n) - B_{CM}(n)] + \varepsilon_P(n). \end{aligned} \tag{23}$$

Here the difference between C_A and C_M corresponds to d, and $P(n)$ combined with other factors such as time constraints will determine whether to actually rely on automation or not.

In a supervisory control system, operators are sensitive to the ability of predicting the capacity of automation or manual control, and previous findings suggest that operator's trust is closely linked with the capacity of automation [61]. More specifically, people's trust on automation may vary according to the change of discrepancy between the operators' expectation and the true behavior (the capacity) of automation. Consequently, though it is useful to get to know the influence of capacity C on trust, it is necessary to examine whether the expectation of capacity is close to the practical situation if we are to develop a predictive model of trust in automation and intervention behavior. Improving the accuracy of operators' perception to the system capacities will also greatly enhance the appropriateness of their trust in automation. Based on this, it is necessary to develop a modified model that can better reflect appropriate trust.

One of the effective ways to modify the EDFT model is to consider the discrepancy between the capacity of two sequential time steps. Accordingly, belief is expressed as:

$$\begin{aligned} B_C(n) = B_C(n-1) + \frac{1}{b_1}(C(n-1) - B_C(n-1)) \\ + (1 - \frac{1}{b_1})(C(n-1) - C(n-2)). \end{aligned} \tag{24}$$

By transposing (24), we will get:

$$X(n-1) = (1 - \frac{1}{b_1})X(n-2), \tag{25}$$

where $X(n-1) = B_C(n-1) - C(n-2)$. Equation (25) constitutes a contraction mapping, from whose definition we know that $X(n-1)$ converges to 0 for a enough large n. Consequently, $B_C(n-1)$ will eventually converge to $C(n-2)$ as time step n increases.

Modifying the generation of belief in pattern of (24) enables operators to generate their belief much closer to the true capacity, and it provides a better understanding of how automation works. As a result, the operators' trust in automation will grow, and thus lead to more appropriate reliance on automation. The effectiveness is supported by simulation results.

5 Results

5.1 Analytical Expression and Conditions for the SR-like Behavior in the SPRT Model

Based on the suggestion of Dr. Jun Zhang, the AFOSR Program Manager, we launched and completed a rather exhaustive review of the past literature on SPRT and SR-related work. Particularly, we focused our attention on the papers published in the mainstream journal "Annals of Mathematical Statistics" from 1945 to 2007, where the pioneering work of Wald and his co-workers (1945, 1948) was published. We also reviewed past literature on SR-like behaviors in psychological experiments, as well as some textbooks or edited books on SPRT, such as those entitled Detection Theory, Handbook of Sequential Analysis, Response Time, and Optimum Stopping Rules. Finally, we also searched for previous work on first passage probability and recursive alternating truncation and convolution. After all these serious efforts, it is our conclusion that the expression we derived for the SPRT/DDD model, as reported in Subsection 4.2.3, is new and the calculations involve alternating multiplication of functions in time domain and frequency domain and in our view cannot be further simplified.

On the other hand, we have used the continuous-time drift diffusion (for short, CDD) model to find the necessary and sufficient condition for the SR-like behavior in CR. It is shown that the following condition is "necessary and sufficient" for the SR-behavior in terms of metric CR:

$$\frac{(2\mu_0 - \hat{\mu}_0 - \hat{\mu}_1)(\hat{\mu}_1 - \hat{\mu}_0)}{2\hat{\sigma}^2} \times \frac{(2\mu_1 - \hat{\mu}_0 - \hat{\mu}_1)(\hat{\mu}_1 - \hat{\mu}_0)}{2\hat{\sigma}^2} > 0 \tag{26}$$

It is easy to see that this condition (26) reduces to the following simplified form when $\mu_0 = \hat{\mu}_0 = 0$, $\hat{\mu}_1 > 0$:

$$\mu_1 < \frac{1}{2}\hat{\mu}_1$$

In addition, simulation results have confirmed our theoretical findings. In the following MATLAB simulation, Figure 7, the SR-like behavior clearly occurs in the fraction of correct response when the actual signal strength is much weaker than anticipated. The vertical axis is the fraction of correct response for various threshold Z while the horizontal axis is σ, the standard deviation of the noise.

From the CDD model, it is also proven that the height of the peak is independent of Z, and the position of the peak (i.e. the optimal noise strength) scales with $\sqrt{Z}$ and $\hat{\sigma}$. As a result, for large Z or $\hat{\sigma}$, the position of the peak corresponds to a strong noise, and expectedly extra noise is helpful under the proposed condition. We have also shown that to guarantee a fixed CR, the response time must increase linearly with the variance of the noise.

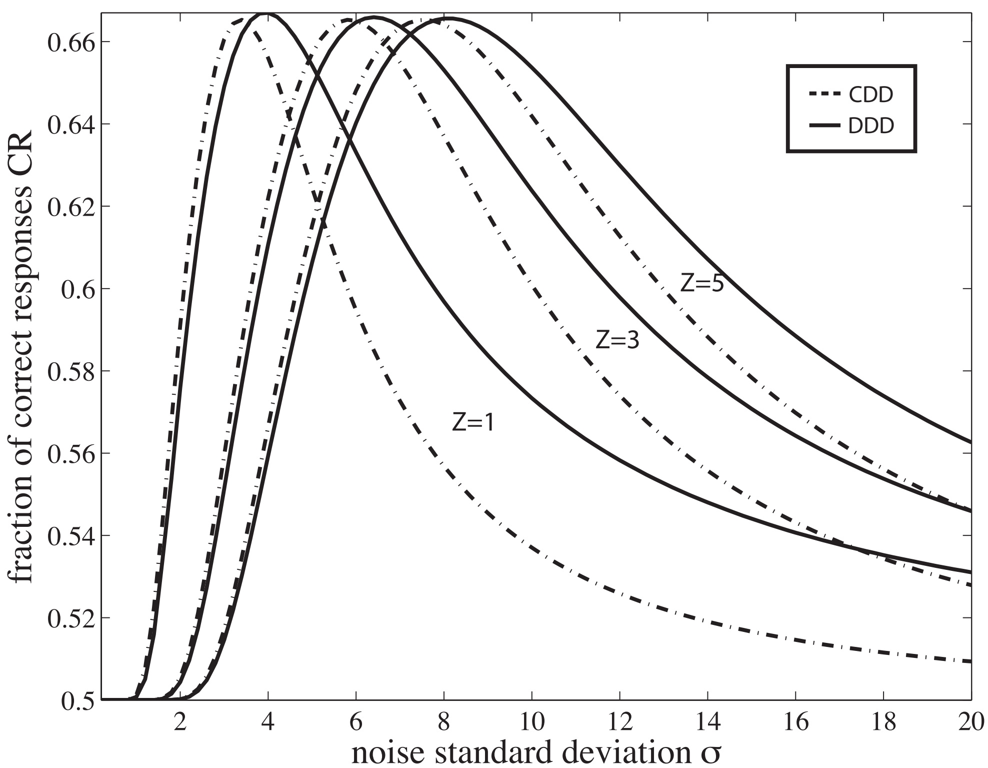

Figure 7: Fraction of Correct Responses vs. Noise. (SR case)

5.2 Differences Between the Predictions of the CDD and DDD Models

Through a comparative study supported by mathematical analysis and computer simulations (see Figures 8–10), three major differences between the predictions of the CDD model and the DDD model are observed as follows.

- First, if the SPRT/DDD model selects the thresholds according to the equations from the CDD model, then CR (the mathematical expectation of accuracy) would be better than desired, but at the price of longer response time than the anticipation of the CDD model.
- Second, the response time always exhibits a peak which is not predicted by the CDD model. As a result, if the actual noise strength is to the right of the peak, then a moderate amount of extra noise could possibly reduce the response time without violating the constraints on the Type I and Type II errors, due to the first difference.
- Third, the tail of the reward rate curve drops significantly with stronger noise, while the CDD model predicts constant reward rate. Such a result may be important for psychological experiments.

5.3 Nonzero Initial Conditions and Extensions to Other Forms of Distributions

For the purpose of testing the robustness of the SR-like behavior, we have plotted the CR when the initial diffusion position is biased (because of asymmetric prior probability) or asymmetric constraints on Type I and Type II errors. The simulation results (see Figure 11 for example) show that the SR-like behavior gets more evident as initial bias moves toward the upper threshold and weaker if the initial bias is negative. Numerical studies are also run on other forms of distributions such as Gamma distribution. See Figure 12. For the Gamma distribution with wrongly assumed shape parameter K, it is observed that SR-like behavior

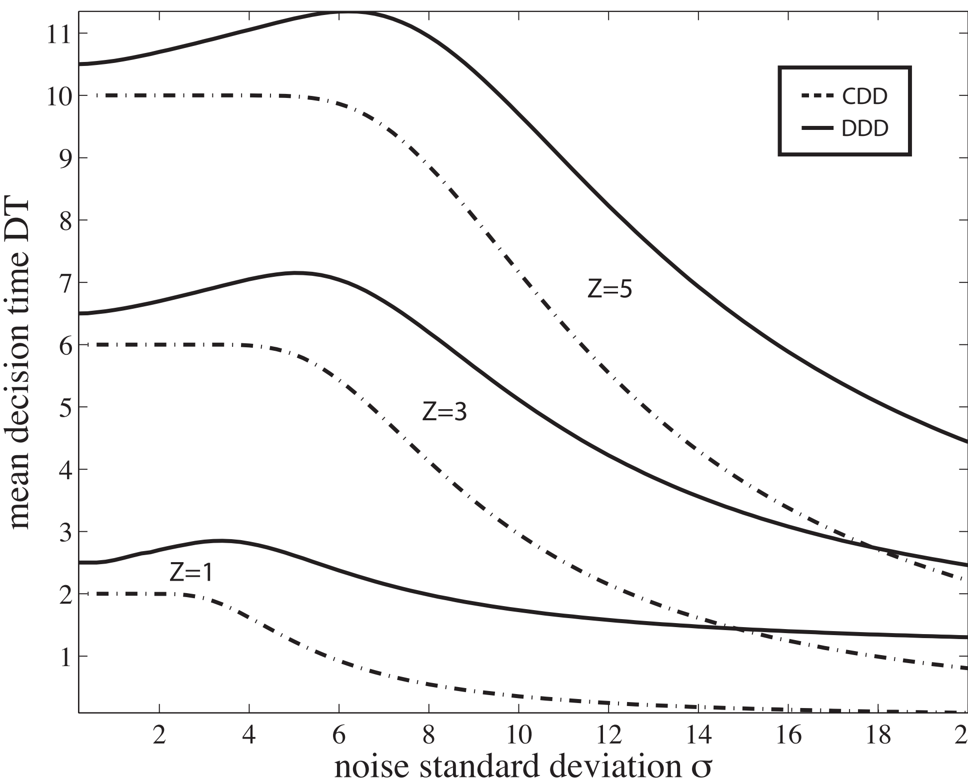

Figure 8: Response Time vs. Noise. (Optimal case)

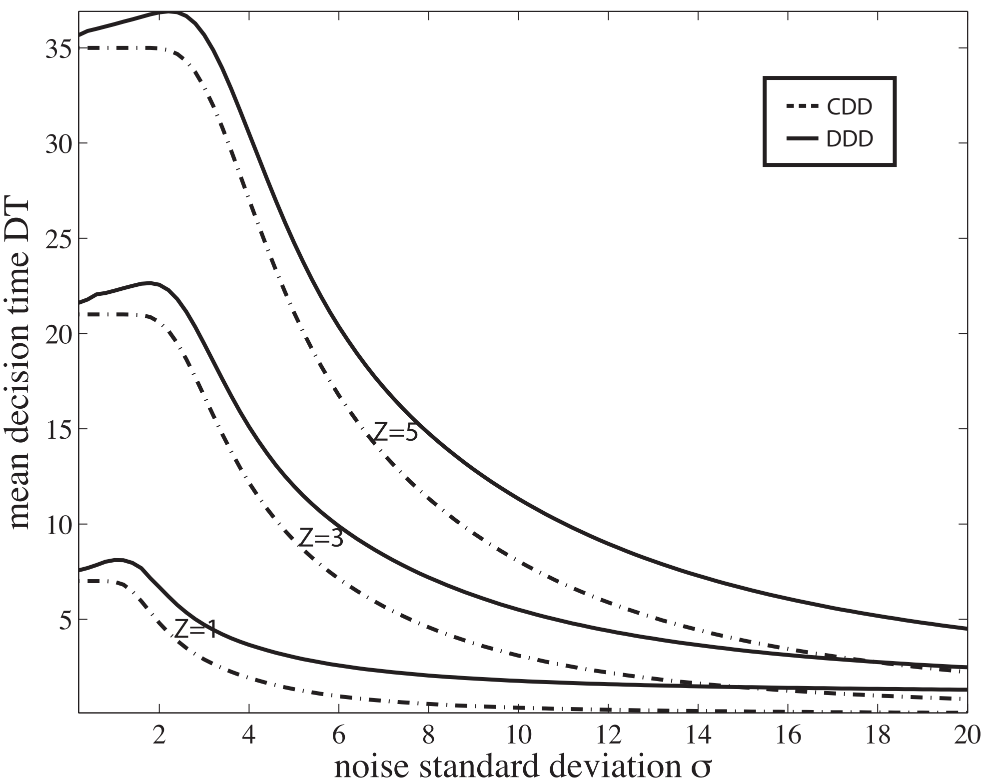

Figure 9: Response Time vs. Noise. (SR case)

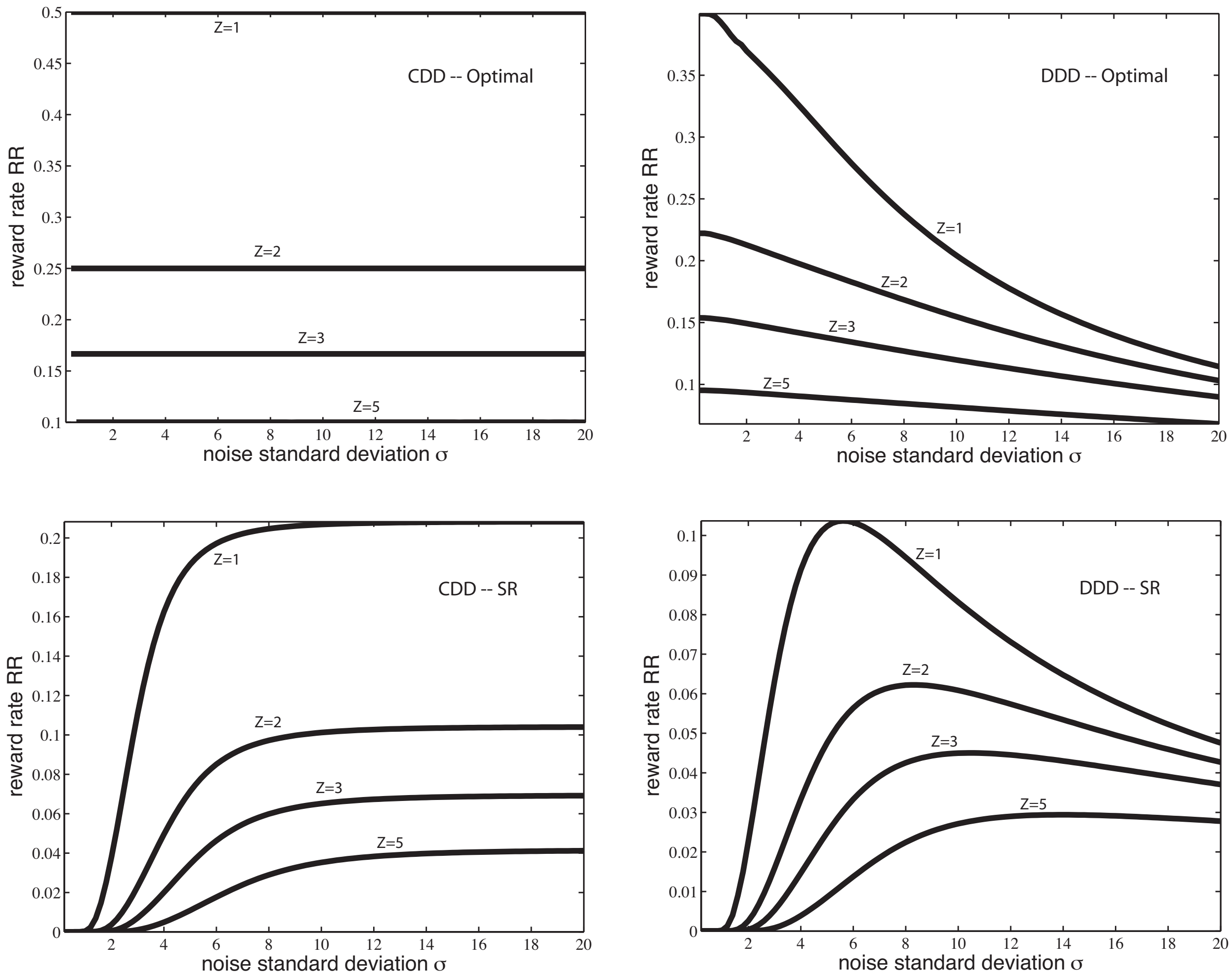

Figure 10: Reward Rate vs. Noise. (Optimal case and SR case)

also occurs. Unlike the Gaussian case, the peak gets higher and narrower with larger Z but the position of the peak does not seem to move much with Z. Like the Gaussian case, the response time also exhibits a peak and gets longer almost linearly with increasing Z and the reward rate also exhibits a peak and gets smaller with increasing Z.

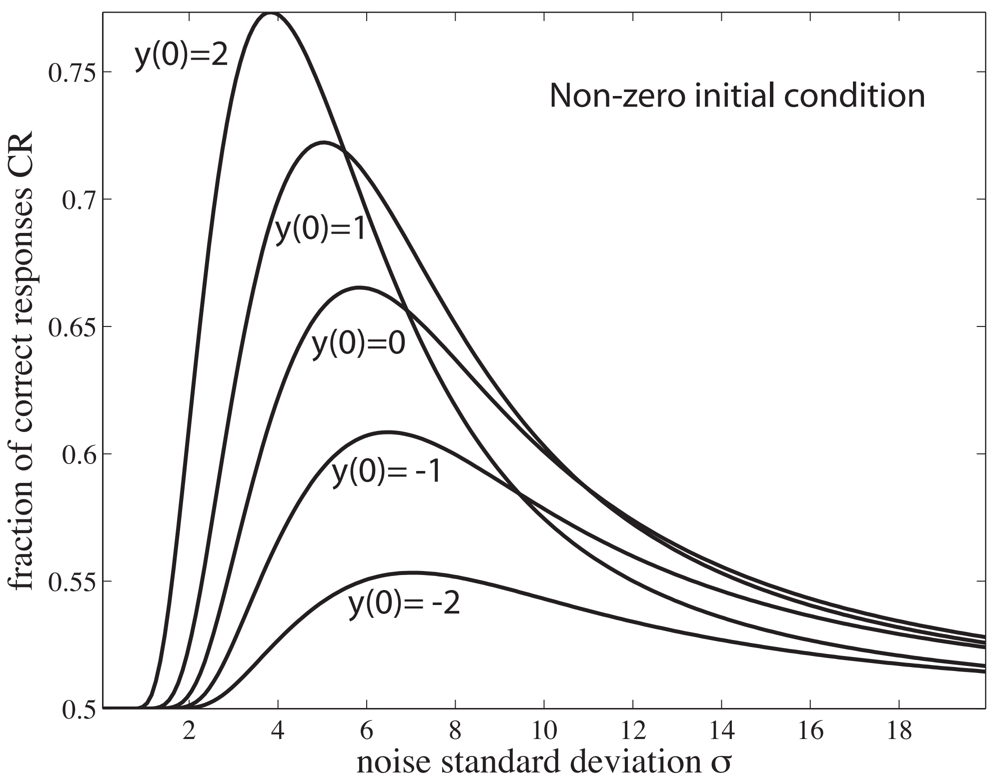

Figure 11: Fraction of Correct Responses vs. Noise with Non-zero Initial Condition. (SR case)

5.4 Role of Prior Distribution in Adaptive SPRT

Our research demonstrates performance improvement in terms of accuracy (while the response time remains moderate) when using the adaptive SPRT algorithm with prior distribution. See Figures 13 and 14.

For these simulations, we assume that the distributions of the input data remain constant during the trials. As a result, we will use all the input sequences that were drawn from H_0 to adapt $p(x|H_0)$. We also apply equal weights to all the input sequences. This assumption can be readily relaxed. For example, if we assume a changing environment, we can apply exponentially decreasing weight for the sequences.

To be more specific, the first time of decision making is based solely on the prior distributions because we do not know which distribution should be adapted. After we receive the feedback, we know that the input sequence in the first trial was drawn from H_0, then we adapt the parameters of $p(x|H_0)$ using that input sequence (also combines with the prior distribution) and apply the new parameters in the next trial. Then suppose we received feedback of the second decision and knew it was drawn from H_1, then we will use this input sequence to adapt the parameters of $p(x|H_1)$. Suppose after 5 decisions, we knew trials $1, 3, 5$ were drawn from H_0, then we estimate $p(x|H_0)$ using input sequences $1, 3, 5$ (also combines with the prior distribution) and estimate $p(x|H_1)$ using input sequences $2, 4$ (also combines with the prior distribution).

For comparison, we still assume that $\sigma_0 = \sigma_1 = \sigma$ and let the adaptive process update σ, i.e. update $p(x|H_0)$ and $p(x|H_1)$ simultaneously. The formula for the adapting process is equation (20). The parameters

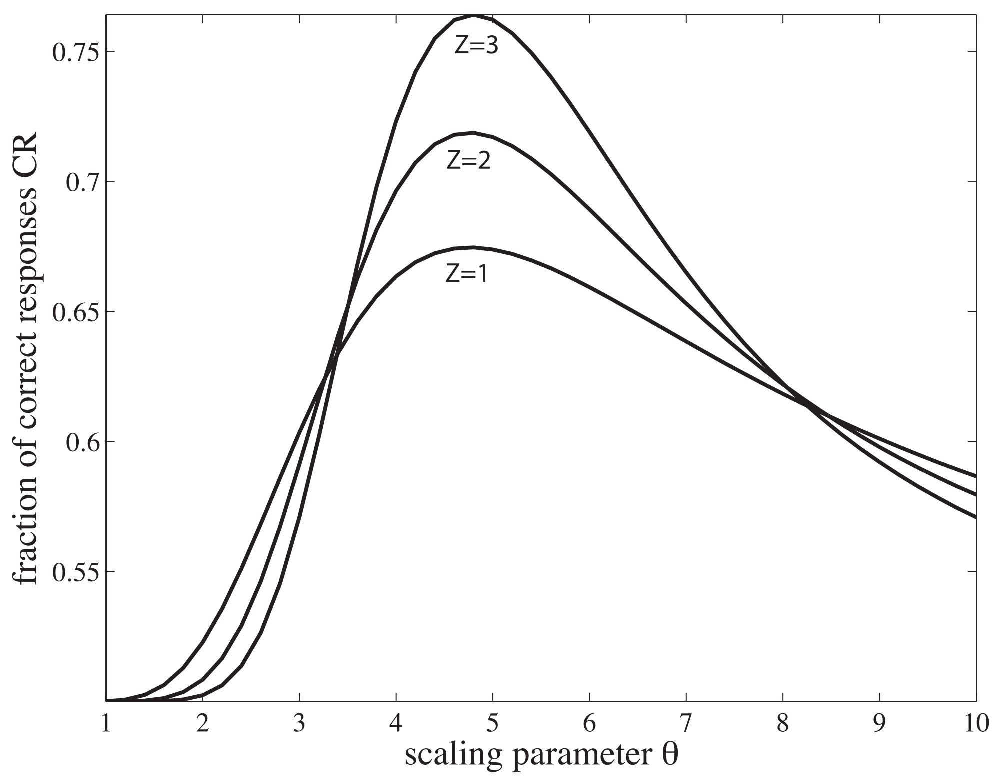

Figure 12: Fraction of Correct Responses vs. scaling parameter θ. Gamma distribution with parameters: $\hat{\theta} = 3$, $\hat{K}_0 = 2$, $\hat{K}_1 = 6$, $K_0 = 2$, $K_1 = 3$. (SR case)

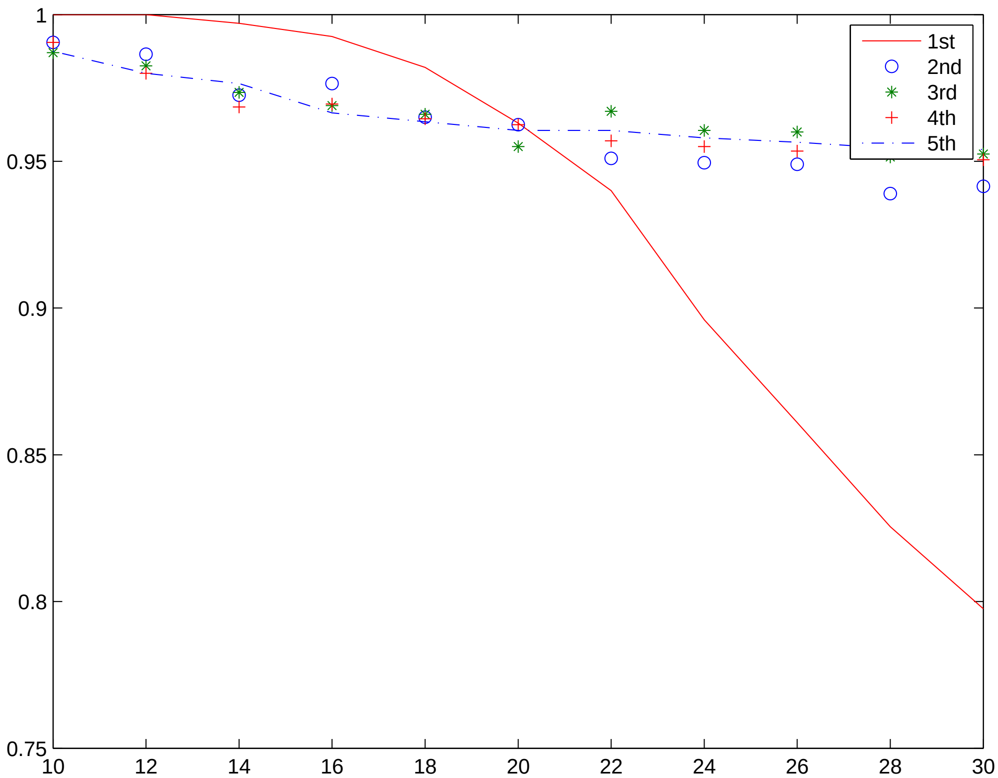

Figure 13: CR of the first 5 steps versus sigma, M=10, power=0.95, significance=0.05. $\mu = 6$, $\sigma_p = 20$, $\sigma \in \{10, 12, \cdots, 30\}$. Averaged over 2000 trials.

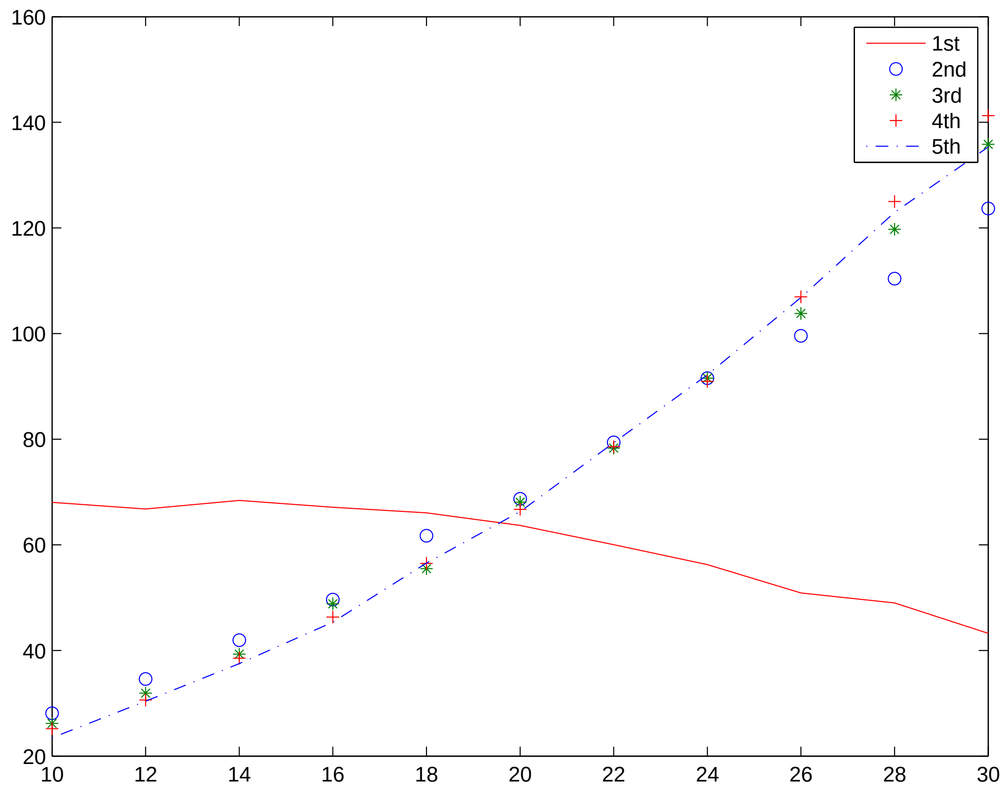

Figure 14: DT of the first 5 steps versus sigma, M=10, power=0.95, significance=0.05. $\mu = 6$, $\sigma_p = 20$, $\sigma \in \{10, 12, \cdots, 30\}$. Averaged over 2000 trials.

are also chosen to be the same as in the previous section.

In Figures 13 and 14, we plot the CR and DT under each trials to see how fast the algorithm converges to the real distributions and reaches the desired optimality. The results are also plotted with respect to the mismatched paramters such as μ_1 and σ to check its robustness. The result shows that CR converges to the desired accuracy while DT remains reasonable and converges to the performance of the optimal SPRT. The first trial does not adapt at all. It only benefits the following trials. As a result, its performance is the same as the original non-adaptive SPRT. The performance of the adaptive SPRT boosts from the second trial, because the first step of SPRT collects a lot of information about the real distributions.

5.5 Extended Decision Field Theory for Dynamic Decision Making

Our research introduces a control-theoretic approach to learning in dynamic decision making tasks to the study of Sugar Factory task. By constructing a control model, it presents a fairly good estimation of automation control capability to participants. Also, the model provides an accurate approximation and a reliable reference to participants through the demonstration of simulation. Aiming at enhancing appropriate trust in automation in a supervisory control system, a modified approach to the previous EDFT model is proposed to provide a more accurate approximation of trust. Feasibility is demonstrated by both theoretic analysis and simulation through a Sugar Factory supervisory control system. The model becomes robust to disturbance irrespective of the fluctuations after modification, and the effectiveness is demonstrated. See Figures 15–17 on the implementation of our generalized DFT on the benchmark example of SF model.

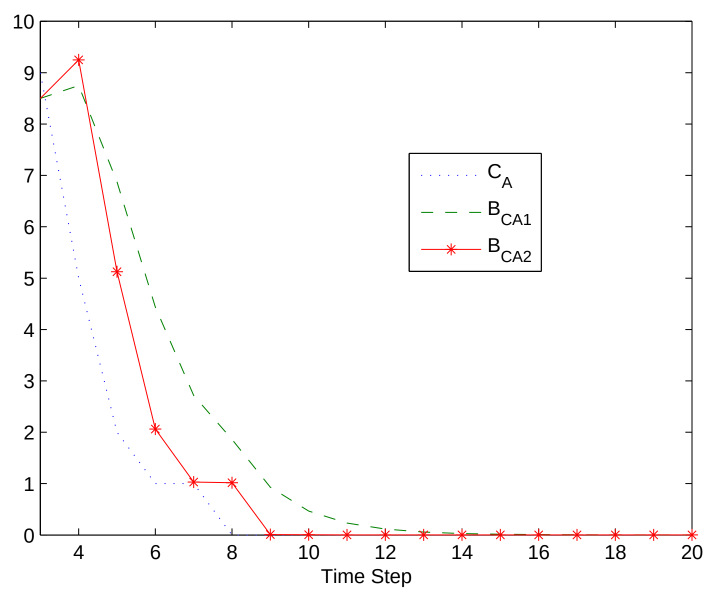

Figure 15: Comparison of B_{CA1} and B_{CA2} when no disturbance exists ($b = 2$)

6 Estimates of Technical Feasibility

Our phase I research reveals theoretically that stochastic resonance like behavior arises in the SPRT model when the actual input signal is significantly weaker than anticipated by the model. Theoretical analysis and computational results demonstrate that both the fraction of correct responses and the reward rate have a peak as a function of noise strength under the SR condition while the mean decision time is a generally decreasing function of the noise strength with the exception of an insignificant peak. That appropriate

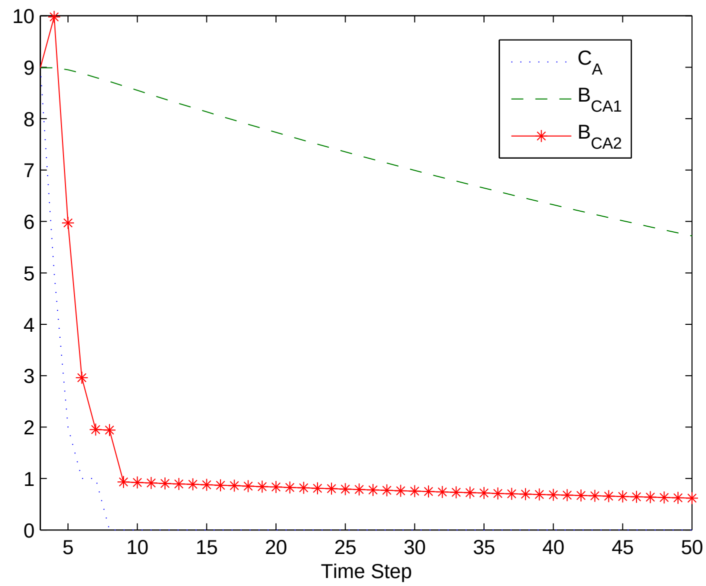

Figure 16: Comparison of B_{CA1} and B_{CA2} when no disturbance exists ($b = 100$)

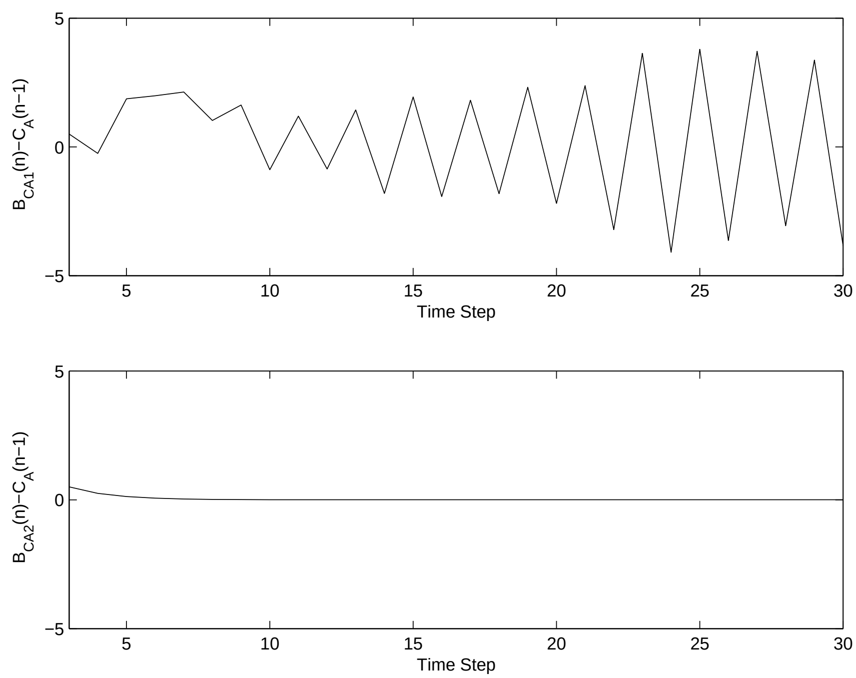

Figure 17: Difference between $B_{CAi}(i = 1, 2)$ and C_A in the presence of disturbance ($b = 2$)

amount of noise can improve the decision making process (when the input signal is significantly weaker than anticipated) is a useful conclusion which may explain human and animal decision making in psychophysical tasks. The novel expressions and conditions for the SR-like behavior in the SPRT model lay a solid foundation for experimental validation of our prediction.

We have proposed an extension of Busemeyer's Decision Field Theory to handle the supervisory control situations where manual control and automation coexist. By means of the celebrated contraction mapping principle, we have developed an improved model describing reliance on automation (trust versus self-confidence) and hopefully will help the operator make appropriate use of automation and manual control.

7 Summary and Future Work

We studied the SPRT/DDD model for decision making under the condition that the input signal is much weaker than anticipated by the model. We derived expressions for calculating the fraction of correct responses, the mean decision time, and the reward rate for the model. Under the assumption that the noise and signal-plus-noise inputs follow Gaussian distributions of equal variance, we found that the fraction of correct responses and the reward rate show SR-like behavior when the actual mean signal level is less than the average of the noise mean and signal-plus-noise mean used in the calculation of the likelihood ratio. Such discrepancy between the actual input distributions and those used by the decision-making process is likely to occur because the brain does not have direct access to the actual input distributions in a psychophysical experiment or in a real-world situation but have to rely on prior experience over a long period of time to estimate the distributions. The SR condition can be understood analytically by using the closed-form expressions from the CDD model. There are also important differences between the SPRT/DDD model and its continuous counterpart, CDD model. For example, if the thresholds are chosen according to the CDD equations, then the accuracy is always better than expected while the response time is always longer. The response time also has a peak which can be useful to reduce the response time without violating the constraints on Type I and Type II errors. The reward rate also appears lower than expected and gets even lower with stronger noise. Finally, to establish the robustness of our findings, we showed that a similar pattern of results is obtained when the initial position of the diffusion process is varied and when the inputs are drawn from gamma, instead of Gaussian, probability distributions. Based on the study, we conclude that a moderate amount of noise can improve decision making when the input signal is weak. This prediction can be tested psychophysically with the reaction-time paradigm. We also studied the role of prior distribution in adaptive SPRT, and analyzed its effects in accuracy and reward rate.

Topics for future research include the systematic study of the signal-dependent noise case and generalizing our preliminary results on adaptive SPRT to other forms of distributions.

Acknowledgement. We would like to thank our sponsor for funding this project.

References

[1] R. Benzi, A. Sutera and A. Vulpiani, "The mechanism of stochastic resonance," *J. Phys. A*, vol. 14, no. 11, L453, 1981.

[2] L. Gammaitoni, P. Hanggi, P. Jung, F. Marchesoni, Stochastic Resonance, *Reviews of Modern Physics*, Vol. 70, No. 1, pp. 223-287, January 1998.

[3] A. Wald, Sequential Tests of Statistical Hypotheses, *Ann. Math. Stat.* Vol. 16, pp 117-186, 1945.

[4] A. Wald, Sequential Analysis, *John Wiley and Sons*, New York, 1947.

[5] J. Wolfowitz, The Efficiency of Sequential Estimates and Wald's Equation for Sequential Processes, *Ann. Math. Stat.* Vol 18, pp 215-230, 1947.

[6] A. Dvoretzky, J. Kiefer, J. Wolfowitz, Sequential Decision Problems for Processes with Continuous time Parameter Hypotheses, *Ann. Math. Stat.* Vol. 24, No. 2, (Jun., 1953), pp. 254-264.

[7] J. Moehlis, E. Brown, R. Bogacz, P. Holmes, Jonathan D. Cohen, Optimizing reward rate in two alternative choice tasks: Mathematical formalism, *Princeton University*, Technical report No. 04-01.

[8] R. Bogacz, E. Brown, J. Mochlis, P. Holmes, and J. D. Cohen, The Physics of Optimal Decison Making: A Formal Analysis of Models of Performance in Two-Alternative Forced-Choice Tasks, *Psychological Review*, Vol. 113, No. 4, pp. 700-765, 2006.

[9] M. Zacksenhouse, P. Holmes and R. Bogacz, Robust versus optimal strategies for two-alternative forced choice tasks, Preprint, 2007.

[10] B. K. Gosh and P. K. Sen, Handbook of Sequential Analysis, *Marcel Dekker*, New York, ISBN 0-8247-8404-1, 1991.

[11] J. J. Collins, C. C. Chow and T. T. Imhoff, Aperiodic stochastic resonance in excitable systems, *Physical Review E*, vol. 52, no. 4, pp. 3321-3324, 1995.

[12] J. J. Collins, T. T. Imhoff and P. Grigg, Noise-enhanced tactile sensation, *Nature*, 383 (6603) : 770, 1996.

[13] Y. Gong, N. Matthews, and N. Qian, Model for stochastic-resonance-type behavior in sensory perception, *Physical Review E*, vol. 65, 031904, 2002.

[14] X. Wu, Z. P. Jiang, D.W. Repperger and Y. Guo, Enhancement of stochastic resonance using optimization theory, *Communications in Information and Systems*, Vol. 6, No. 1, pp. 1-18, 2006.

[15] B. Xu, F. Duan and F. Chapeau-Blondeau, Comparison of aperiodic stochastic resonance in a bistable system realized by adding noise and by tuning system parameters, *Physical Review E*, Vol. 69, 061110, 2004.

[16] R. Cloux and D. J. McConalogue, A Numerical Algorithm for Recursively-Defined Convolution Integrals Involving Distribution Functions, *Management Science*, Vol. 22, No. 10 (Jun., 1976), pp. 1138-1146

[17] C. M. Harris, D. M. Wolpert, Signal-dependent noise determines motor planning, *Nature*, 1998 Aug 20, Vol 394(6695):780-4.

[18] H. Tanaka, J. W. Krakauer, and N. Qian, An Optimization Principle for Determining Movement Duration, *J Neurophysiol*, June 1, 2006, Vol. 95(6):3875-3886.

[19] A. G. Tartakovsky, An Efficient Adaptive Sequential Procedure for Detecting Targets, *Proceedings of the IEEE Aerospace Conference*, Big Sky, Montana, 9-16 March 2002.

[20] T.W. Anderson, A modification of the Sequential Probability Ratio Test to Reduce the Sample Size, *Ann. Math. Stat.* Vol. 31, No. 1, (Mar., 1960), pp. 165-197.

[21] P. J. Huber, A robust version of the probability ratio test, *Ann. Math. Stat.* Vol. 36, pp. 1753-1758, 1965.

[22] A. G. Tartakovsky, An efficient adaptive sequential procedure for detecting targets, IEEEAC paper #10, 2002.

[23] V.P. Dragalin and A.A. Novikov, "Adaptive Sequential Tests for Composite Hypotheses", Technical Report 94.4, Institute for Applications of Mathematics and Informatics, Milan, 1994.

[24] I.V. Pavlov, "Sequential Procedure of Testing Composite Hypotheses With Applications to the Kiefer-Weiss Problem", Theory Prob. Appl., 35, 280-292, 1990.

[25] H. Robbins and D. Siegmund, "A Class of Stopping Rules for Testing Parametric Hypotheses", Proceedings of the Sixth Berkeley Symposium on Theory of Prob. and Math. Stat., 4, 37-41, 1973.

[26] H. Robbins and D. Siegmund, "The Expected Sample Size of Some Tests of Power One", THe Annals of Statistics, 2, 415-436, 1974.

[27] Gilles R. Ducharme and Teresa Ledwina, Efficient and adaptive nonparametric test for the two-sample problem, Ann. Statist. Volume 31, Number 6 (2003), 2036-2058.

[28] Andres Azuero, A Wald's SPRT-Based Group Sequential Testing Procedure with Flexible Multiplicity Correction for Genetic Association Studies and Genome-Wide Association Scans, First International Workshop in Sequential Methodologies (IWSM 2007), July 22-25, 2007

[29] CHANG Yuan-Chin Ivan, Application of sequential interval estimation to adaptive mastery testing, 2005, vol. 70, no4, pp. 685-713.

[30] A. Tartakovsky, "Asymptotic Properties of M-SPRT and Adaptive Sequential Test in Multihypothesis Problems with an Indifference Zone." Joint Statistical Meetings. Dallas, Texas: IMS Bulletin, 1998, vol. 27, No. 3, p. 167.

[31] P. Ioannou and J. Sun, Robust adaptive control, Prentice Hall, 1995.

[32] R. S. Michalski, J. G. Carbonell, and T. M. Mitchell, Machine learning, Tioga Publishing, 1983.

[33] T. Miller, R. S. Sutton, and P. J. Werbos, Neural networks for control, MIT press, 1991.

[34] R. J. Sutton, and A. C. Barto, Reinforcement learning, MIT press, 1999.

[35] D. C. Berry and D. E. Broadbent, On the relationship between task performance and associated verbalizable knowledge, *Quarterly Journal of Experimental Psychology* 36A, 1984, pp. 209–231.

[36] D. C. Berry and D. E. Broadbent, The combination of explicit and implicit learning processes in task control, *Psychological Research* 49, 1987, pp. 7–15.

[37] B. Brehmer, Strategies in real-time, dynamic decision making. In R. Hogarth (Ed.), *Insights from decision making*, Chicago: IL: University of Chicago Press, 1990, pp. 262–279.

[38] B. Brehmer, Dynamic decision making: Human control of complex systems, *Acta Psychologica* 81, 1992, pp. 211–241.

[39] A. Buchner, J. Funke and D. C. Berry, Negative correlations between control performance and verbalizeable knowledge: Indicators for implicit learning in process control tasks, *Quarterly Journal of Experimental Psychology* 48A(1), 1995, pp. 166–187.

[40] J. R. Busemeyer, Decision making under uncertainty: A comparison of simple scalability, fixed-sample, and sequential-sampling models, *Journal of Experimental Psychology* 11, 1985, pp. 538–564.

[41] J. R. Busemeyer and J. T. Townsend, Decision field theory: A dynamic cognitive approach to decision making in an uncertain environment, *Psychol. Rev.* 100(3), 1993, pp. 432–459.

[42] J. R. Busemeyer and A. Diederich, Survey of decision field theory, *Math. Soc. Sci.* 43(3), 2002, pp. 345–370.

[43] D. Fum and A. Stocco, Instance vs. rule based learning in controlling a dynamic system, *Proceedings of the international conference on cognitive modelling*, Universitats-Verlag Bamberg, Bamberg, Germany, 2003, pp. 105–110.

[44] Z. Dienes and R. Fahey, Role of specific instances in controlling a dynamic system, *Journal of Experimental Psychology: Learning, Memory, and Cognition* 21, 1995, pp. 848–862.

[45] S. Farrell and S. Lewandowsky, A connectionist model of complacency and adaptive recovery under automation, *Journal of Experimental Psychology: Learning, Memory, and Cognition* 26(2), 2000, pp. 395–410.

[46] G. F. Franklin and J. D. Powell and M. L. Workman, *Digital Control of Dynamic Systems*, Addison-Wesley Publishing Company, Inc. 1992

[47] G. F. Franklin and J. D. Powell, *Feedback Control of Dynamic Systems*, New Jersey: Pearson Prentice Hall 2006

[48] M. Fishbein and I. Ajzen, *Belief, Attitude, Intention, and Behavior*, Reading, MA: Addison-Wesley 1975

[49] J. Gao and J. D. Lee, Extending the Decision Field Theory to Model Operators' Reliance on Automation in Supervisory Control Situations, *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans* 36(5), 2006, pp. 943–959.

[50] F. P. Gibson and D. C. Plaut, A connectionist formulation of learning in dynamic decision-making tasks,*Proceedings of the 17th Annual Conference of the Cognitive Science Society* , 1995, pp. 512–517.

[51] F. P. Gibson, The Sugar Production Factory–A Dynamic Decision Task, *Computational and Mathematical Organization Theory 2:1*, 1996, pp. 49–60.

[52] F. P. Gibson, M. Fichman and P. Plaut, Learning in dynamic decision tasks: computational model and empirical evidence, *Organizational Behavior and Human Decision Processes* 71(1), July, 1997, pp. 1–35.

[53] R. M. Hogarth, Generalization in decision research: The role of formal models, *IEEE Transactions on Systems, Man, and Cybernetics* 16, 1986, pp. 439–449.

[54] M. I. Jordan, Constrained supervised learning, *Journal of Mathematical Psychology* 36, 1992, pp. 396–425.

[55] M. I. Jordan, Computational aspects of motor control and motor learning, In H. Heuer, and S. Keele (Eds.), *Handbook of perception and action: Motor skills*, New York: Academic Press 1997

[56] M. I. Jordan and D. E. Rumelhart, Forward models: Supervised learning with a distal teacher, *Cognitive Science* 16(3), 1992, pp. 307–354.

[57] J. D. Lee and N. Moray, Trust, control strategies, and allocation of function in human-machine systems, *Ergonomics* 35(10), 1992, pp. 1243–1270.

[58] J. D. Lee and N. Moray, Trust, self-confidence, and operators' adaptation to automation, *Int. J. Human-Comput. Stud.* 40(1), 1994, pp. 153–184.

[59] J. D. Lee and K. A. See, Trust in Automation: Designing for Appropriate Reliance, *Human Factors* 46(1), Spring 2004, pp. 50–80.

[60] B. M. Muir, Trust in automation 1: Theoretical issues in the study of trust and human intervention in automated systems, *Ergonomics* 37(11), 1994, pp. 1905–1922.

[61] B. M. Muir and N. Moray, Trust in automation 2: Experimental studies of trust and human intervention in a process control simulation, *Ergonomics* 39(3), 1996, pp. 429–460.

[62] K. Ogata, *Discrete-time Control Systems*, New Jersey: Prentice-Hall 1987

[63] R. Parasuraman and V. Riley, Humans and automation: Use, misuse, disuse, abuse, *Hum. Factors* 39(2), 1997, pp. 230–253.

[64] D. W. Repperger and C. A. Phillips, The Human Role in Automation, *Handbook of Automation*, Springer –Verlag, April 2008

[65] R. Roe, J. R. Busemeyer and J. T. Townsend, Multialternative decision field theory: A dynamic connectionist model of decision-making, *Psychol. Rev.* 108(2), 2001, pp. 370–392.

[66] W. B. Stanley, R. C. Mathews, R.R. Buss and S. Kotler-Cope, Insight without awareness: On the interaction of verbalization, instruction, and practice in a simulated process control task, *Quarterly Journal of Experimental Psychology* 41A(3), 1989, pp. 553–577.

[67] N. A. Taatgen and D. Wallach, Whether skill acquisition is rule or instance based is determined by the structure of the task, *Cognitive Science Quarterly* 2, 2002, pp. 163–204.

[68] A. C. Wick, S. L. Berman and T. M. Jones, The structure of optimal trust: Moral and strategic, *Academy of Management Review* 24, 1999, pp. 99–116.

[69] R. E. Wood and A. Bandura, Impact of conceptions of ability on self-regulatory mechanisms and complex decision making, *Journal of Personality and Social Psychology* 56, 1989, pp. 407–415.