

**NONLINEAR SIGNAL PROCESSING FOR
DIGITAL COMMUNICATIONS USING
SUPPORT VECTOR MACHINES AND A NEW
FORM OF ADAPTIVE DECISION FEEDBACK
EQUALIZER**

By

Daniel J. Sebald

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY
(ELECTRICAL ENGINEERING)

at the

UNIVERSITY OF WISCONSIN – MADISON

2000

Dedicated to the memory of

Richard A. Sebald

&

Rocky

Contents

List of Tables	iv
List of Figures	v
Acronym Glossary	viii
Acknowledgments	ix
1 Introduction	1
2 Support Vector Machine Techniques for Nonlinear Equalization	4
2.1 Introduction	4
2.2 Detection viewed as pattern recognition	6
2.3 SVM	7
2.3.1 Dual optimization problem	8
2.3.2 QP Formulation	9
2.3.3 Calculating the affine offset	10
2.3.4 Kernel options	11
2.4 SVM model of equalization	12
2.5 Simulations and results	15
2.6 Combating Temporal Dependence	21
2.7 More simulations and results	27
2.8 Conclusions	29
3 Support Vector Machines and the Multiple Hypothesis Test Problem	30
3.1 Introduction	30
3.2 Multicategory Classifier Overview	33
3.2.1 Parametric strategies	33
3.2.2 Binary classifier based strategies	34
3.2.3 Multiclass optimization strategies	36
3.3 Proposed M -ary classification	38
3.3.1 M -ary classification	38
3.3.2 Illustrative example	40
3.3.3 Comments on Gray codes	42
3.4 SVM with equality constraints	44
3.4.1 Example	48
3.5 Conclusions	49

4	A Binary Adaptive Decision-Selection Equalizer for Channels with Non-linear Intersymbol Interference	52
4.1	Introduction	52
4.2	Channel Model	57
4.3	Adaptive Decision-Selection Equalization	58
4.3.1	Pattern Spaces and the ADFE Limitation for Nonlinear Channels .	58
4.3.2	The Optimum Subclass Detector	62
4.3.3	ADSE Concept	64
4.3.4	Discussion about LMS Algorithm	66
4.3.5	More Robust LMS-Based Solution	68
4.4	Monte Carlo Simulations	69
4.5	Conclusions	81
A	Importance Sampling Transformations	82
	Bibliography	84

List of Tables

1	Simulation statistics for the SVM $\hat{x}(n) = \tilde{x}(n) - 0.9\tilde{x}^3(n)$, $\tilde{x}(n) = u(n) + 0.5u(n-1)$, and Gaussian white noise of power $\sigma_e^2 = 0.2$; and equalizer dimension $M = 2$, polynomial kernel order $d = 3$, and constraint $C = 5$. The number of training points is 500, the number of test points is 5000.	18
2	Version of the LMS-based ADSE algorithm where hyperplane orientation and distance from origin are allowed to vary.	65
3	Version of the LMS-based ADSE algorithm where only hyperplane distance from origin is allowed to vary.	69

List of Figures

1	Nonlinear system \mathcal{N} is equalized by an SVM. The training data for the SVM is obtained from appropriately arranging the nonlinear system output $x(n)$ of the training sequence $u(n)$ and the desired output y_n , which is simply a delayed version of the training sequence.	12
2	Model of the nonlinear transmission system.	13
3	Example of typical classification regions of an SVM and associated training points with channel $\hat{x}(n) = \tilde{x}(n) - 0.9 \tilde{x}^3(n)$, $\tilde{x}(n) = u(n) + 0.5 u(n-1)$, and Gaussian white noise of power $\sigma_e^2 = 0.2$; and equalizer dimension $M = 2$, polynomial kernel order $d = 3$, constraint $C = 5$, and lag $D = 0$	16
4	Example of typical classification regions of an SVM and associated training points with the channel described in Fig. 3, except the equalizer lag is now $D = 1$	16
5	Example of typical classification regions of an SVM and associated training points with the channel described in Fig. 3, except the equalizer lag is now $D = 2$	17
6	Example of a type of SVM error in which the classification boundary does not match the optimum in the region where empirical data is unlikely. The system is that described in Fig. 4.	19
7	Example of a type of SVM error in which an obvious region of misclassification appears in the outer regions of the classification space. The system is that described in Fig. 4.	19
8	Example of typical classification regions of an SVM and associated training points with channel $\hat{x}(n) = \tilde{x}(n) + 0.1 \tilde{x}^2(n) + 0.05 \tilde{x}^3(n)$, $\tilde{x}(n) = 0.5 u(n) + u(n-1)$, and Gaussian colored noise of power $\sigma_e^2 = 0.2$, correlation $\rho = 0.48$ (i.e., $\xi = 0.75$); and equalizer dimension $M = 2$, lag $D = 0$, polynomial kernel order $d = 3$ and constraint $C = 5$	21
9	Average BER for the SVM detector versus the Bayesian detector for the channel $\hat{x}(n) = \tilde{x}(n) + 0.2 \tilde{x}^2(n)$, $\tilde{x}(n) = 0.3482 u(n) + 0.8704 u(n-1) + 0.3482 u(n-2)$, and Gaussian colored noise correlation $\rho = 0.48$ (i.e., $\xi = 0.75$); and equalizer dimension $M = 3$, lag $D = 1$, polynomial kernel order $d = 3$ and constraint $C = 0.1$	22
10	A DFE in which past decisions are fed back through filter $A(f)$ to cancel ISI.	23
11	The constellation diagram for linear channel (9) for delay $D = 1$, assuming $u(n-2)$ and $u(n-3)$ are unknown.	24
12	The various constellations for delay $D = 1$ in different states of $u(n-2)$ and $u(n-3)$ for the linear channel (9).	25
13	A bank of SVMs with selection based upon the state of previous decisions. Each SVM constructs a different decision boundary based upon the constellation for a given state.	26

14	Average BER for the SVM, SVMB, DFSVM, and PDFSVM detectors compared against the optimum binary, non-ISI detector for the linear channel (9) with added Gaussian noise; and equalizer dimension $M = 2$, lag $D = 1$, polynomial kernel order $d = 3$, constraint $C = 0.5$, and DFSVM/PDFSVM feedback dimension $M' = 1$	28
15	M -ary nonlinear classification using $K' = \lceil \log_2 K \rceil$ SVMs. Classes are declared according to binary representation $y^N y^{N-1} \dots y^1$	39
16	Signal constellations for 8-PSK signaling using an (a) natural code and (b) Gray code.	40
17	Result for Gray code.	42
18	Result for natural code. The more complex regions result in the data not being separable when $d = 3$. Consequently, the results are poor.	43
19	For the natural code, increasing d to 5 results in the feature space data being separable. Boundaries are now reasonable, but with areas of ambiguity near the boundaries.	44
20	Pattern space for a noise-free, two user DS-CDMA system with correlation $\rho = 0.5$. Boundaries show the optimum decision regions under the assumption of additive white Gaussian noise.	49
21	Example decision boundaries of the M -ary SVM trained with the displayed data for the DS-CDMA space of Fig. 20.	50
22	Example decision boundaries of an M -ary SVM with an equality constraint forcing the border to pass through the origin. The data is the same as that for Fig. 21.	50
23	Block diagram of an ADFE. When the switch is in position A, the ADFE is in training mode. Position B corresponds to decision-directed mode.	53
24	The conditional constellations for a two input ($N_w = 2$), two decision state ($N_b = 2$), one delay ($D = 1$) equalizer where the channel is linear ISI with $\hat{x}(n) = 0.4084 u(n) + 0.8164 u(n - 1) + 0.4084 u(n - 2)$, conditioned on the correct past decision variables.	59
25	The conditional constellations for a two input, two decision, one delay equalizer where the channel is nonlinear ISI with $\tilde{x}(n) = 0.4084 u(n) + 0.8164 u(n - 1) + 0.4084 u(n - 2)$ and $\hat{x}(n) = \tilde{x}(n) + 0.2 \tilde{x}^2(n) + 0.1 \tilde{x}^3(n)$	61
26	The conditional constellations for a two input, two decision, one delay equalizer where the channel is nonlinear ISI with $\tilde{x}(n) = u(n) + 0.2 u^2(n) + 0.1 u^3(n)$ and $\hat{x}(n) = 0.4084 \tilde{x}(n) + 0.8164 \tilde{x}(n - 1) + 0.4084 \tilde{x}(n - 2)$	62
27	Conditional constellations ($j = 2$) subject to 17 dB AWGN for the system used to generate Fig. 25.	67
28	Block diagram of the ADSE (Algorithm II), where $\delta(\cdot, \cdot)$ is the Kronecker delta function.	70
29	Simulation results for ADSE Algorithm I with $N_b = 2$, $N_w = 6$, $D = 5$ and $\mu = 0.001$ on the channels of Section 4.3.	72
30	Simulation results for ADSE Algorithm II with $N_b = 2$, $N_w = 6$, $D = 5$ and $\mu = 0.001$ on the channels of Section 4.3.	74

31	Simulation results for ADSE Algorithm II with $N_b = 2$, $N_w = 6$, $D = 5$ and $\mu = 0.001$ on the channels of Section 4.3, but with no scaling by N_b in the ISI subtraction constant.	77
32	Simulation results for ADSE Algorithm II with $N_b = 2$, $N_w = 6$, $D = 5$ and $\mu = 0.001$ on the channels of Section 4.3, but with all adaptive coefficients initialized to zero.	79

Acronym Glossary

<i>k</i> -RLP	<i>k</i> -class RLP, 38
8-PSK	octaphase-shift-keying, 2
ADFE	adaptive decision feedback equalizer, 2
ADSE	adaptive decision-selection equalizer, 2
AWGN	additive white Gaussian noise, 52
BER	bit-error-rate, 3
BRG	binary reflected Gray, 41
CDMA	code division, multiple access, 48
DFE	decision feedback equalizer, 2
DFSVM	decision feedback SVM, 23
ECSVM	equality constrained SVM, 31
FIR	finite impulse response, 13
ISI	intersymbol interference, 1
LMS	least-mean-squared, 3
LP	linear program, 37
M-SVM	multicategory discrimination SVM, 38
ML	maximum likelihood, 62
MMSE	minimum mean-squared error, 55
PAM	pulse amplitude modulation, 13
PDFSVM	perfect DFSVM, 23
QP	quadratic programming, 1
RBF	radial basis function, 23
RLP	robust linear program, 37
SNR	signal-to-noise ratio, 3
SVMB	SVM-bank, 26
SVM	support vector machine, 1
<i>M</i> -ary SVM	Multicategory SVM with a power of two categories, 2
i.i.d.	independent, identically distributed, 52
p.d.f.	probability density function, 82
p.m.f.	probability mass function, 82

Acknowledgments

First and foremost, I would like to thank my adviser and all-around cool egg Jim Bucklew, without whom I surely would not have graduated. He gave me one last opportunity, and his thoughtful consideration and friendship throughout were greatly appreciated.

I thank the professors who served on the dissertation committee—John Gubner, Yu-Hen Hu, Barry Van Veen and Grace Wahba. They provided good guidance during the preliminary and final oral exams and took bona fide interest in Jim’s and my work. I owe a debt of gratitude to the readers, Prof. Van Veen and Prof. Hu. Reading a dissertation still may be the least glamorous of academic jobs, but now at least it is not a thankless one. I am honored that Prof. Wahba, from the Statistics Department, was part of the committee.

I must thank long time friend and bird of a feather, Tim Bawden. We always seemed to be on the same wavelength even though he spent most of his time atop Science Hall on the liberal end of campus, while I spent most of my time in the basement of Engineering Hall on the techie end of campus. A view of Memorial Union and the Capitol versus a view of asphalt core samples and bags of pea gravel—life isn’t fair. Along with Tim, I would also like to mention Wendy Bawden (and aviation expert Max), and Dave and Gloria Compton (hi Colin and Nate!).

I appreciate the support of my family—Rich and Roxann Sebald (Steven and Amanda), Linda and Brian Langum (hi Genna, Leigh and Eric!), David and Patti Haese, Brian and Amy Haese (Taylor, McKenzie and Delaney), Jodi and Troy Wagner, and especially my Mom Esther Sebald.

A warm thank you goes to Donna Lewis of the ECE Department for moral support and financial support in the form of a teaching assistanship when the chips were down. Also in the moral support category, I would like to thank the nicest professor on campus, Sandy

Shaw Courter of the EPD Department.

From the Computer Science Department, Prof. Olvi Mangasarian and Prof. Michael Ferris played an important role by kindly supplying us with a robust quadratic programming algorithm. And along the lines of computer resources, I must thank guru Bruce Orchard for taking care of so many Unix problems. I lost track a long time ago of how often Bruce helped in times of computer crisis. Also helpful at CAE were Beth Anderson and the straitlaced but likable Mary Baldwin.

I would like to mention, aside from those on my dissertation committee, ECE Department professors from whom I've learned a thing or two—Paul Milenkovic, Bob Barmish, Truong Nguyen, squash champ Rajeev Agrawal, and friend Bill Sethares.

From the many courses I've taken in the Math Department, I was lucky enough to learn from some gifted teachers—Prof. Jim Kuelbs, Steve Hutt, Paul Rabinowitz, Frank Forelli, Tom Kurtz, cantankerous Louis Rall, Patrick Ahearn, and especially Dan Shea. Prof. David Lindberg from the History of Science Department also falls under the above category.

There are a lot of people I've met in the last several years in regard to school and work, and some I've known from before UW, who have been supportive in some way. I'm sure I've overlooked someone I shouldn't have, but I'd like to thank (or remember) Rob and Heidi Nowak, Mike Nowak, Keith Burgess, Francis Neven, Eko Onggosanusi, Ron Leder, Ted Hea, Sharky the shark dog, Jerry Lopez, Tom Hundt, debate champion Gil Raz, debate champion understudy Saeed Asgari, Constantino Lagoa, the unique Surekha Palreddy, Valtino Afonso, Mark Mayotte, Mohammad Asgarian, Geoff “Ditch of Doom” Weinberg, Tamer Kadous, Ashwin Ganesan, Jianzhong Zhang, Anita Balasubramanian, Rajesh Sharma, Thomas Andre, Alberto Rodriguez, Yihchih Shern, amigo Yuris Fuentes for sending good karma in the direction of Madison during my defense, Laura Fuentes with various band incarnations (Sotevento y Paraguas) for local music, Jacob Eapen, Glen Luckjiff, Ian Wallace, German

King (and all those who played soccer across from the Waisman Center), Jason Bold, Chancellor Archie (no, not that chancellor), professor turned working stiff Mike Mulligan for good advice along the way, Prof. Jim Heinen, Prof. Russ Niederjohn, Rich Kelnhofer, Mike Chan, Jim Whitney, Tim Ruchti, Dave Steckl, long lost and newly found Kevin Melhorn, Barry Grabow, Tom Beckman, Prof. Bob Kern, Prof. Carol Diggelman for encouragement all those years ago, Steve Dimino, Mark Boutot, Joel Wichgers, Tim Dejewski, Wan-Hsin Tang, Mike McGrath, Marion McGavok, Murat Fahrioglu, Nikolai Krasnoperov, Daniel Mastropietro, Micki McCartney (Minnesota is her real name, but don't tell any one), John Pienkos, Neda Juričić, neighbor Jim Nelson, Bob Conway, Teri Sippel Schmitt, Ana Maria Tomé, Bob Rice, Dan Mahn (the Harry Potter of engineers), Brian Donahoe (the Mozart of engineers), Sharon Grinaway, Linda Vanasupa, Ronnie Mattheis, Karl Marshke, Eric Olmanson, bike-a-holic Robbie Weber, Pam Barrett, avenger of urban sprawl Mike Barrett, Tom Simmons, republican amid democrats Robert Summerfelt, Bob Gosse, Ron Gosse, Brian Michael, Bob Dylan fanatic Tim Eigenberger, Toby Funk, Tim Gohl, Alan Tefft, Terry Nimmer, Bill Harrison, Tony Fontana, Gary Schneider, Ray Heasty, gentleman farmer Rob Whisler, Steve Krueger, bike-and-ski-a-holic Greg Ladwig, Jon Milbrandt, Chris Barron, Jay Roberts, Jeff Schaetz, Bryan Piernot, Dennis Bahr for keeping the engineering flame burning, Mike Ihm, Dr. Alan Kahn, Jim Reuss, Mike Larsen, Gerhard Von der Ruhr, Oliver Wieben, Moncef Elaoud, Reza Ghodssi, Jun Kikuchi, Manuel Solorzano, Masoud Kasraian, Icelander Ragnar Palsson, Carol Nin, Chris Carlson-Dakes, Jodi Reeves, Eric Benedict, Michelle Rogers, Jordan Lee, Tris Utschig, Mike Baker, Prof. Mike Corradini, Picnic Point, The Terrace, Jean Brown, Debbie Diewald for organization, and Carol Chase for keeping Jim Bucklew in line (well, at least trying anyway).

There are a few good friends I met bicycling with Bombay and Wednesday Night Riders. It was always nice to get away from the daily grind and share a story with Jerry Greenfield, Jim and Lee Kleeman, Connie Daering, Al Abramson, David Pfaff, Dave Tenenbaum, Kate

and Chic Gladding, and especially Bob McGrath. Thanks to Trek for transportation and to all the riders who pulled up front.

Last, I would like to thank fickle yet flexible L^AT_EX for treating me so well while most other word processors gave me nothing but grief... I love you, L^AT_EX!