


Low Complexity Near Optimal Hybrid Detectors for Large-Scale MIMO Uplink Systems Based on Complex Support Vector Regression

Tianpei Chen

 Department of Electrical & Computer Engineering
McGill University
Montreal, Quebec, Canada

November 2015

A thesis submitted to McGill University in partial fulfillment of the requirements for the degree of Master of Engineering.

© 2015 Tianpei Chen

Abstract

This report describes the use of \LaTeX to format a thesis. A number of topics are covered: content and organization of the thesis, \LaTeX macros for controlling the thesis layout, formatting mathematical expressions, generating bibliographic references, importing figures and graphs, generating graphs in MATLAB, and formatting tables. The \LaTeX macros used to format a thesis (and this document) are described.

Acknowledgments

Thesis regulations require that contributions by others in the collection of materials and data, the design and construction of apparatus, the performance of experiments, the analysis of data, and the preparation of the thesis be acknowledged.

Contents

1	Introduction	1
1.1	Large MIMO system	1
1.2	Large MIMO detections	1
1.3	Support Vector Regression	3
1.4	Thesis Contribution	5
1.5	Thesis Outline	5
2	Theoretical Analysis of Channel Hardening Phenomenon	6
2.1	System Model	6
3	Complex Support Vector Preliminary Detector (CSVPD) for Large-Scale MIMO Uplink Systems	8
3.1	Brief Introduction to ϵ -Support Vector Regression (ϵ -SVR)	9
3.2	Combined Single Direction Searching Strategy	11
3.3	Complex Support Vector Preliminary Detector	12
3.4	Channel Hardening Approximation	12
3.5	CSVPD-LAS versus MMSE-LAS	12
3.6	CSVPD-OPIC versus MMSE-OPIC	12
4	Mathematical Layout Styles	13
5	Tables	14
5.1	Tables in L ^A T _E X	14
A	L^AT_EX Macros	16
A.1	Thesis Preamble	16

Contents	iv
References	17

List of Figures

2.1	Large MIMO uplink system	7
3.1	Regression Model of ϵ -SVR	10

List of Tables

5.1	Filter specifications	15
-----	---------------------------------	----

List of Acronyms

16-QAM	16-point Quadrature Amplitude Modulation
3GPP	Third Generation Partnership Project
3GPP2	Third Generation Partnership Project 2
64-QAM	64-point Quadrature Amplitude Modulation
ADSL	Asymmetric Digital Subscriber Line
ARQ	Automatic Repeat Request
WPAN	Wireless Personal Area Network

Chapter 1

Introduction

1.1 Large MIMO system

One of the biggest challenges the researchers and industry practitioners are facing in wireless communication area is how to bridge the sharp gap between increasing demand of high speed communication of rich multimedia information with high level Quality of Service (QoS) requirements and the limited radio frequency spectrum over a complex space-time varying environment. The promising technology for solving this problem, Multiple Input Multiple Output (MIMO) technology has been of immense research interest over the last several tens of years is incorporated into the emerging wireless broadband standard like 802.11ac [1] and long-term evolution (LTE) [2]. The core idea of MIMO system is to use multiple antennas at both transmitting and receiving end, so that multiplexing gain (multiple parallel spatial data pipelines that can improve spectrum efficiency) and diversity gain (better reliability of communication link) are obtained by exploiting the spatial domain. Large MIMO (also called Massive MIMO) is an upgraded version of conventional MIMO technology employing hundreds of low power low price antennas at base station (BS), that serves several tens of terminals simultaneously. This technology can achieve full potential of conventional MIMO system while providing additional power efficiency as well as system robustness to both unintended man-made interference and intentional jamming. [3] [4].

1.2 Large MIMO detections

The price paid for large MIMO system is the increased complexities for signal processing at both transmitting and receiving ends. The uplink detector is one of the key components

in a large MIMO system. With orders magnitude more antennas at the BS, benefits and challenges coexist in designing of detection algorithms for the uplink communication of large MIMO systems. On one hand, a large number of receive antennas provide potential of large diversity gains, on the other hand, complexities of the algorithms become crucial to make the system practical.

Vertical Bell Laboratories Layered Space-Time (V-BLAST) architecture for MIMO system can achieve high spectrum efficiency by spatial multiplexing (SM), that is, each transmit antenna transmits independent symbol streams. However the optimal maximum likelihood detector (MLD) for V-BLAST systems that perform exhaustive search over the transmit symbol vector space has a complexity that increases exponentially with the number of transmitted antennas, which is prohibitive for practical applications.

In order to alleviate this problem, linear detectors (LD) such as zero-forcing (ZF) and minimum mean square error (MMSE) aided by successive interference cancellation with optimal ordering (ZF-OSIC, MMSE-OSIC) are exploited in V-BLAST architecture [5] [6] [7], although ZF-OSIC and MMSE-OSIC can provide significant improvement comparing to their LD counterparts, a common drawback of SIC aided LD is the error propagation effect, which can not be eliminated by ordering. That results in inferior performances comparing to MLD [8] [9] [10].

Sphere Decoder (SD) [11] is the most prominent algorithm that utilizes the lattice structure of MIMO systems, which can achieve optimal performance with relatively much lower complexity comparing to MLD. However, SD has two major shortages that make it problematic to be integrated into a practical systems. The first shortage is SD has various complexities under different signal to noise ratios (SNR), while a constant processing data rate is required for hardware. The second shortage is the complexity of SD still has a lower bound that increases exponentially with the number of transmit antennas and the order of modulation scheme [12]. The fixed complexity sphere decoder (FCSD) [13] makes it possible to achieve near optimal performance with a fixed complexity under different values of SNR. The FCSD inherits the principle of list based searching algorithms, which first generate a list of candidate symbol vectors and then the best candidate is chosen as the solution. The other sub optimal detectors belong to this class include Generalized Parallel Interference Cancellation (GPIC) [14] and Selection based MMSE-OSIC (sel-MMSE-OSIC) [15]. However, all these list based searching algorithms have the same shortage - their complexities increase exponentially with the number of transmit antennas and the order of modulation scheme [15]. Therefore, such algorithms are prohibitive when it comes to a

large number of antennas or a high order modulation scheme, for example in IEEE 802.11ac standard [1], the modulation scheme is 256QAM.

Besides the above detection algorithms designed for conventional MIMO systems, in the last several years, a variety of metaheuristic based local search algorithms invoked from machine learning field [16] have been proposed for large MIMO systems. These algorithms have complexities that are comparable or slightly higher comparing to MMSE detector and near-optimal performance. Such algorithms include likelihood ascend searching (LAS) algorithm and variants [17] [18] [19] and Reactive Tabu search (RTS) algorithms and variants (e.g. Layered Tabu search (LTS) [20], Random Restart Reactive Tabu search (R3TS) [21]). Additionally, there are other algorithms proposed for large MIMO systems including Message passing technique based algorithms (e.g. Belief propagation (BP) detectors based on graphic model and Gaussian Approximation (GA) [22] [23] [24] [25]), Probabilistic Data Association (PDA) based algorithms [26], Monte Carlo sampling technique based algorithms (e.g. Multiple Restart- Mixed Gibbs Sampling (MR-MGS) algorithm [27]) and Element based Lattice Reduction (LR) aided algorithms [28].

Considering MIMO detection from a Combinatorial Optimization (CO) problem viewpoint, as powerful tool for solving CO problems, metaheuristic algorithms [29] are good choices for designing large MIMO detectors, driven by demand of achieving acceptable performance with significantly lower computational complexity. Besides the metaheuristic algorithms that based on local search strategies which use trajectory methods based on single solution, another class of metaheuristic algorithms is defined as population based. The population based metaheuristic algorithms deal with a population of candidate solutions. Intrinsically the population based metaheuristic algorithms can provide wider and more efficient exploration of search space. The major population based metaheuristic algorithms includes Evolutionary Computation (EC) and Ant Colony Optimization (ACO).

Genetic algorithm (GA) is one kind of EC algorithms, which are designed for CO problems. GA mimics the natural evolution process of a population and is powerful tool in searching a solution that is close enough to global optimum [30].

1.3 Support Vector Regression

Firmly grounded in framework of statistical learning theory or VC (Vapnik–Chervonenkis) theory, the Support Vector (SV) technique has become a powerful tool to solve real world supervised learning problems such as classification, regression and prediction. The SV

method is a nonlinear generalization of Generalized Portrait algorithm developed by Vapnik in 1960s [31] [32], which can provide good generalization performance [33].

Interest in SV algorithms boosted since 1990s, promoted by the works of Vapnik and co-workers at AT& T Bell laboratory [34] [35] [36] [37] [38] [39]. Moreover, the kernel based methods [33] were proposed in order to extending the SV algorithms to nonlinear learning cases. The input data samples are mapped into some high dimensional feature space (also called Reproducing Kernel Hilbert Space (RKHS)) and then linear tools are applied to the feature mappings of the input data samples. This is equivalent to transforming the nonlinear learning tasks in the original space into the linear learning tasks in high dimensional feature space. The mathematical notion underlying kernel based methods is that of RKHS [33], in which the inner products of the feature mappings can be simply replaced by the computationally economical kernel functions. Because SV algorithms only deal with the inner products of the the feature mappings. Therefore, by kernel based methods, it is sufficient to use the specific kernel functions based on the RKHS discarding the actual structure of the feature space.

Based on the principle of structural risk minimization [40], the ϵ -SVR [36] [41] solves an original constraint optimization problem (primal objective problem) by transforming it into a Lagrange dual form (dual objective problem), which is a Quadratic Programming (QP) problem. Efficient methods for training SV algorithms which are based on large scale data sets were proposed, which is called decomposition. Decomposition process is performed by decomposing a large QP problem into a sequence of sub QP problems and solve them in an iterative manner [42] [43], Sequential Minimal Optimization (SMO) algorithm [44] is one of well known representatives of decomposition methods.

Complex valued signal arises in signal processing and digital communication areas etc. Therefore, developing signal processing algorithms which are suitable to be directly applied to complex valued systems is typically natural and concise choice. Furthermore, for MIMO systems, although one can transform the complex valued system model into an equivalent real valued system model, the detection algorithms built based on the complex valued model is more preferable due to the flexibility for signal constellation choice and hardware implementations [45]. In recent years, a mathematical framework for pure complex valued SV algorithms was developed which can deal with complex valued tasks in signal processing, digital communication and related areas in an elegant and computational efficient manner [46].

Chapter 2

Theoretical Analysis of Channel Hardening Phenomenon

2.1 System Model

Consider a uncoded complex large MIMO uplink spatial multiplexing (SM) system with N_t users, where each is equipped with transmit antenna. The number of receive antennas at the Base Station (BS) is N_r , $N_r \geq N_t$. Typically large MIMO systems have hundreds of antennas at the BS, as shown in Fig 2.1.

Bit sequences, which are modulated to complex symbols, are transmitted by the users over a flat fading channel. The discrete time model of the system is given by:

$$\mathbf{y} = \mathbf{H}\mathbf{s} + \mathbf{n}, \quad (2.1)$$

where $\mathbf{y} \in \mathbb{C}^{N_r \times 1}$ is the received symbol vector, $\mathbf{s} \in \mathbb{C}^{N_t}$ is the transmitted symbol vector, with components that are mutually independent and taken from a finite signal constellation alphabet \mathbb{O} (e.g. BPSK, 4-QAM, 16-QAM, 64-QAM), $|\mathbb{O}| = M$. The transmitted symbol vectors $\mathbf{s} \in \mathbb{C}^{N_t}$, satisfy $\mathbb{E}[\mathbf{s}\mathbf{s}^H] = \mathbf{I}_{N_t}E_s$, where E_s denotes the symbol average energy, $\mathbb{E}[\cdot]$ denotes the expectation operation, \mathbf{I}_{N_t} denotes identity matrix of size $N_r \times N_t$. Furthermore $\mathbf{H} \in \mathbb{C}^{N_r \times N_t}$ denotes the Rayleigh fading channel propagation matrix, each component is independent identically distributed (i.i.d) circularly symmetric complex Gaussian random variable with zero mean and unit variance. Finally, $\mathbf{n} \in \mathbb{C}^{N_r}$ is the additive white Gaussian noise (AWGN) vector with zero mean components and $\mathbb{E}[\mathbf{n}\mathbf{n}^H] = \mathbf{I}_{N_r}N_0$, where N_0 denotes the noise power spectrum density, and hence $\frac{E_s}{N_0}$ is the signal to noise ratio (SNR).

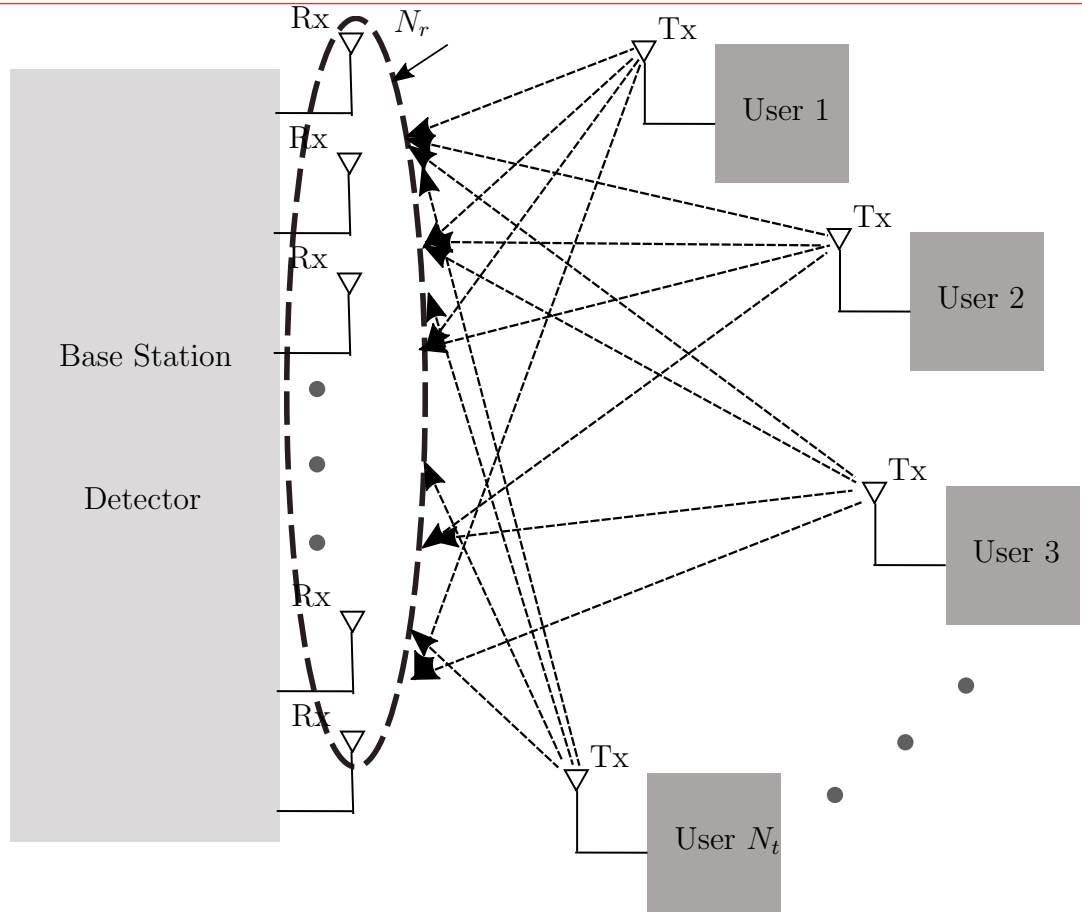


Fig. 2.1 Large MIMO uplink system

The task of a MIMO detector is to estimate the transmit symbol vector \mathbf{s} , based on the knowledge of receive symbol vector \mathbf{y} and channel matrix \mathbf{H} .

Chapter 3

Complex Support Vector Preliminary Detector (CSVPD) for Large-Scale MIMO Uplink Systems

In this chapter, we proposed a complex support vector preliminary detector (CSVPD), which has a complexity comparable with linear detectors but performs much better. The proposed CSVPD can work in the complex valued Large-Scale MIMO (LS-MIMO) systems in an elegant and efficient manner. We exploit the specific widely linear regression model without constant offset, which is suitable for LS-MIMO systems model. Furthermore, based on the analysis of channel hardening phenomenon from Chapter 2 and the recent advances in machine learning field concerning decomposition process for SVM [47], we propose a combined single direction searching strategy in CSVR training process which can approximately maximize the gain of the sub dual objective functions whose work set size are two. The combined single direction searching strategy proposed can achieve a much smaller searching times than the optimal double direction searching while make dual objective function converge using as few iterations as the latter one.

Furthermore, we show that channel hardening phenomenon can be further exploited in CSVPD to reduce the computational complexity and simplify the implementation of the algorithm for LS-MIMO systems. Finally, we compare the performance of CSVPD aided likelihood ascend search (LAS) detector and CSVPD aided parallel interference cancellation with ordering (OPIC) detector with MMSE-LAS and MMSE-OPIC. These results demonstrate the potential of CSVPD to be utilized as a preliminary detector that can improve the LS-MIMO detectors whose performances are influenced by the reliability of

3.1 Brief Introduction to ϵ -Support Vector Regression (ϵ -SVR)

Suppose we are given a training data $((\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \dots, (\mathbf{x}_L, y_L))$, L denotes the size of the training data set, $\mathbf{x}_i \in \mathbb{R}^V$ denotes input data vector, V is the number of features in \mathbf{x}_i . y_i denotes the output. Let \mathbf{w} denotes regression coefficient vector, $\Phi(\mathbf{x}_i)$ denotes the feature mapping of \mathbf{x}_i , $\mathbf{w}, \Phi(\mathbf{x}_i) \in \mathbb{R}^\Omega$, $\Omega \in \mathbb{R}$ denotes the dimension of mapped feature space (For linear model, $\mathbf{x}_i = \Phi(\mathbf{x}_i)$, $\Omega = V$). The regression estimate $g(\mathbf{x}_i)$ (either linear or non-linear) is given by

$$g(\mathbf{x}_i) = \mathbf{w}^T \Phi(\mathbf{x}_i) + b \quad i = 1, 2, \dots, L \quad (3.1)$$

In ϵ -SVR, as shown in Fig 3.1, the solid line represents the regression estimate in 3.1, ϵ controls the precision of the regression, the area between two dash lines (ϵ -bound) is called ϵ -tube. Only the regression estimates that have the deviation larger than ϵ (the data points located outside ϵ -tube in Fig.3.1) denoted by ξ_i and $\hat{\xi}_i$, contribute to the estimation errors. the goal of ϵ -SVR is to minimize the risk introduced by estimate errors while keeping \mathbf{w} small, which is also called regularized risk minimization principle, therefore, the primal constraint optimization problem is formulated as

$$\begin{aligned} \min_{\mathbf{w}, \xi, \hat{\xi}_i} f(\mathbf{w}, \xi_i, \hat{\xi}_i) &= \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^L (R(\xi_i) + R(\hat{\xi}_i)) \\ s.t. \quad &\begin{cases} y_i - \mathbf{w}^T \Phi(\mathbf{x}_i) - b \leq \epsilon + \xi_i, i = 1, 2 \dots, L \\ \mathbf{w}^T \Phi(\mathbf{x}_i) + b - y_i \leq \epsilon + \hat{\xi}_i, i = 1, 2 \dots, L \\ \epsilon, \xi_i, \hat{\xi}_i \geq 0, i = 1, 2 \dots, L \end{cases} \end{aligned} \quad (3.2)$$

In (3.2), $\frac{1}{2} \|\mathbf{w}\|^2$ is the regularization term in order to ensure the flatness of regression model. slack variables ξ_i and $\hat{\xi}_i$ are introduced based on the "soft margin" principle [37] that can cope with the infeasible constraints of the optimization problem and allows the existence of some additive noise to the observations. $R(u)$ denotes the cost function. The simplest cost function is $R(u) = u$, The choice of the cost function is determined by the statistical distribution of the additive noise [41]. For example if the noise is Gaussian distributed, then the optimal cost function is $R(u) = \frac{1}{2} u^2$. The term $C \sum_{i=1}^L (R(\xi_i) + R(\hat{\xi}_i))$ denotes

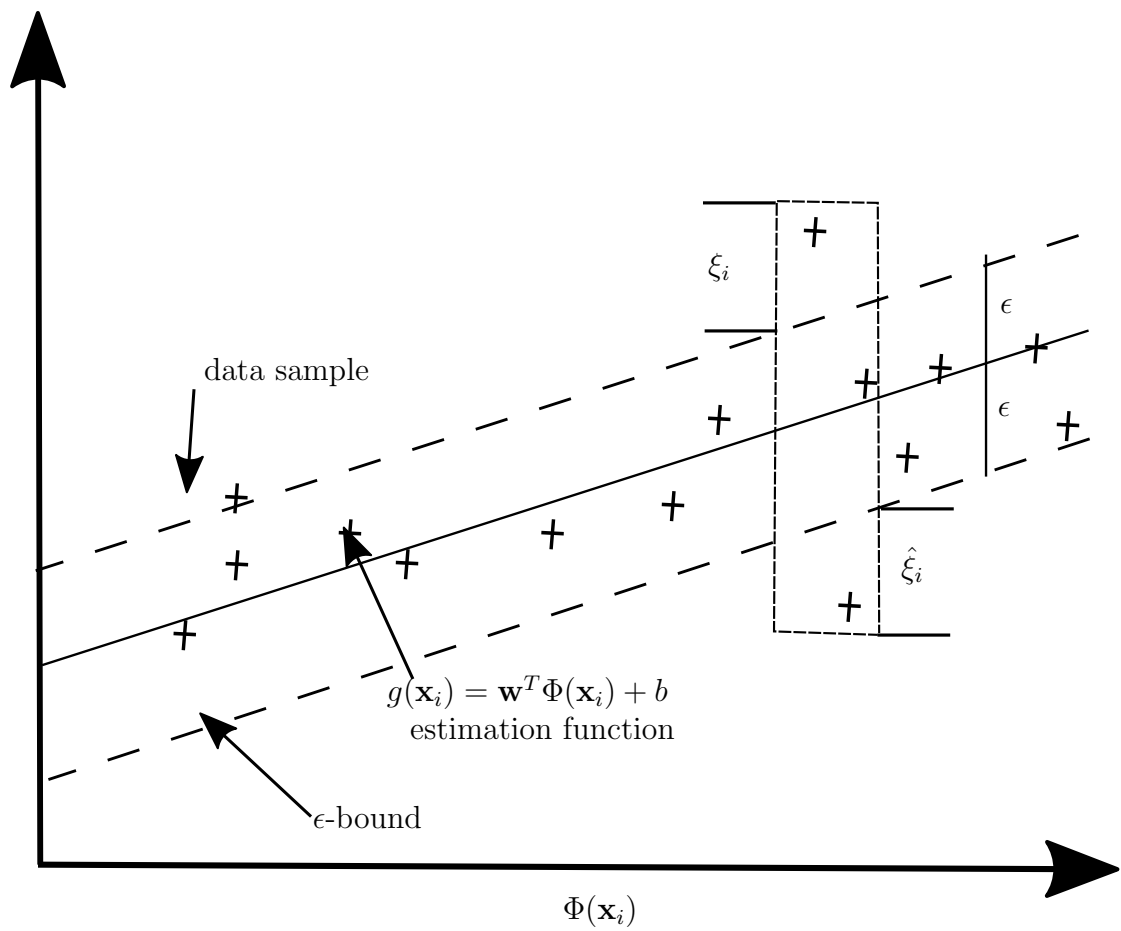


Fig. 3.1 Regression Model of ϵ -SVR

the penalty of the additive noise, $C \in \mathbb{R}$ and $C > 0$ that controls the trade off between regularization term and cost function term.

In ϵ -SVR, the objective to exploit slack variables ξ_i and $\hat{\xi}_i$ is to compensate the influences from the outliers that exceed the ϵ -tube which are caused by noise. Therefore in ϵ -SVR, ξ_i and $\hat{\xi}_i$ are defined as

$$\xi_i = \max(0, y_i - \mathbf{w}^T \Phi(\mathbf{x}_i) - b - \epsilon) \quad (3.3)$$

$$\hat{\xi}_i = \max(0, \mathbf{w}^T \Phi(\mathbf{x}_i) + b - y_i - \epsilon). \quad (3.4)$$

Because the deviation between the estimates $g(\mathbf{x}_i)$ and the observation y_i as defined in 3.1 can only exceeds the ϵ -tube in one direction, therefore there is at most one of ξ_i and $\hat{\xi}_i$ can be non zero. That is $\xi_i \hat{\xi}_i = 0$.

3.2 Combined Single Direction Searching Strategy

Decomposition methods were proposed to solve this QP problem by decomposing it into sub QP problems and solving them iteratively [44]. Therefore, the computational intensive numerical methods can be avoided. Decomposition is performed by sub set selection solver, which refers to a set of algorithms that separate the optimization variables (Lagrange multipliers) into two sets S and N, S is the work set and N contains the remaining optimization variables. In each iteration, only the optimization variables in the work set is updated while keeping other optimization variables fixed. The Sequential Minimal Optimization (SMO) algorithm [44] is an extreme case of the decomposition solvers. An important issue of the sub set selection solvers is the selection of the work set. One strategy is to choose Karush-Kuhn-Tucker (KKT) condition violators, ensuring the final converge [43]. The SMO algorithm restricts the size of the work set to two. A method to train SVM without offset was proposed In [47], with the comparable performance to the SVM with offset. A set of sequential single variable work set selection strategies, which require $O(n)$ searching time are proposed. The optimal double variable work set selection strategy, which performs exhaustively searching, however, requires $O(n^2)$ searching time. The authors demonstrate that with the combination of two different proposed single variable work set selection strategies, convergence can be achieved by a iteration time that is as few as optimal double variable work set selection strategy.

The mathematical foundation of kernel based methods is RKHS which is defined in

3 Complex Support Vector Preliminary Detector (CSVPD) for Large-Scale MIMO Uplink Systems

~~complex domain, however most of the practitioners are dealing with real data sets.~~ **12**
communication and signal processing area, the channel gains, signals, waveforms etc. are all represented in complex form. Recently, a pure complex SVR & SVM based on complex kernel was proposed in [46], which can deal with the complex data set purely in complex domain. The results in [46] demonstrate the better performance as well as reduced complexity comparing to simply split learning task into two real case by real kernels. Based on this work, we derive a complexity-performance controllable detector for large MIMO systems based on a dual channel complex SVR (CSVSR). The detector can work in two parallel real SVR channels which can be solved independently. Moreover, only the real part of kernel matrix is needed in both channels. This means a large amount of computation cost saving can be achieved. Based on the discrete time MIMO channel model, in our regression model, this CSVSR-detector is constructed without offset, Therefore, for each real SVR without offset, Two types of combined single optimization variable selection strategy are proposed based on the work in [47]. The proposed combined single optimization variable selection strategy can approximate optimal double optimization variable selection strategy. The former one can achieve as few as iteration time while enjoy significant speed gain in each iteration.

3.3 Complex Support Vector Preliminary Detector

3.4 Channel Hardening Approximation

3.5 CSVPD-LAS versus MMSE-LAS

3.6 CSVPD-OPIC versus MMSE-OPIC

Chapter 4

Mathematical Layout Styles

T_EX does a marvelous job of setting mathematical formulas, most often choosing pleasing spacing. However, on occasion one should intercede to improve the layout. This chapter defines a few such occasions. In addition, this chapter documents some features of the `amsmath` package which overcome difficulties in typesetting some mathematical forms. The `amsmath` package is documented in *The L^AT_EX Companion* [?].

The modified setup is typeset as

$$G(z) = \begin{cases} \frac{P(z)}{1+z^{-1}} & \text{for } p \text{ even,} \\ P(z) & \text{for } p \text{ odd.} \end{cases} \quad (4.1)$$

With the modified definitions, we get the following.

$$\begin{aligned} \boldsymbol{d}^{(i)} &= \hat{\boldsymbol{v}}^{(i)} - \hat{\tilde{\boldsymbol{v}}}^{(i)} \\ \boldsymbol{n}^{(i)} &= \boldsymbol{u}^{(i)} - \tilde{\boldsymbol{v}}^{(i)} \end{aligned} \quad (4.2)$$

Chapter 5

Tables

5.1 Tables in L^AT_EX

Tables of many different sorts can be made with L^AT_EX. This chapter gives suggestions on producing tables, along with a number of examples.

To illustrate these rules, here is a table and the L^AT_EX input which was used to generate it.

Table 5.1 Filter specifications

Taps (N)	Transition Band	Stopband Weighting (α)	Passband Ripple dB	Stop-band Rejection dB	Ultimate Stop Band dB
8	A	1	0.06	31	31
12			0.025	48	50
16			0.008	60	75
12	B	1	0.04	33	36
16			0.02	44	48
24			0.008	60	78
16	C	1	0.07	30	36
24		1	0.02	44	49
32		2	0.009	51	60
48		2	0.006	50	66
24	D	1	0.1	30	38
48		2	0.006	50	66
64		5	0.002	65	80
48	E	2	0.07	32	46
64		5	0.025	40	51

Transition Code Letter	Normalized Transition Band
A	0.14
B	0.10
C	0.0625
D	0.043
E	0.023

The normalized transition band is the width of the transition band normalized to 2π ; that is, $(\omega_s - \pi/2)/(2\pi)$.

Appendix A

L^AT_EX Macros

The L^AT_EX commands and macros used in formatting the title page for this document are shown in this appendix.

A.1 Thesis Preamble

The commands used to create the title page for a thesis are shown below. The McGill University crest is brought in via a macro `McGillCrest` which allows for setting the size and colour of an imported PostScript file which contains the actual crest. The title page also includes a red separator line.

References

- [1] “IEEE Standard for Information technology– Telecommunications and information exchange between systemslocal and metropolitan area networks– Specific requirements– Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications–Amendment 4: Enhancements for Very High Throughput for Operation in Bands below 6 GHz.” *IEEE Std 802.11ac-2013 (Amendment to IEEE Std 802.11-2012, as amended by IEEE Std 802.11ae-2012, IEEE Std 802.11aa-2012, and IEEE Std 802.11ad-2012)*, pp. 1–425, Dec 2013.
- [2] E. Dahlman, S. Parkvall, J. Skold, and P. Beming, *3G Evolution: HSPA and LTE for Mobile Broadband*. Academic press, 2010.
- [3] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, “Scaling up MIMO: Opportunities and challenges with very large arrays,” *Signal Processing Magazine, IEEE*, vol. 30, no. 1, pp. 40–60, 2013.
- [4] E. Larsson, O. Edfors, F. Tufvesson, and T. Marzetta, “Massive MIMO for next generation wireless systems,” *Communications Magazine, IEEE*, vol. 52, no. 2, pp. 186–195, 2014.
- [5] P. W. Wolniansky, G. J. Foschini, G. Golden, R. Valenzuela *et al.*, “V-BLAST: An architecture for realizing very high data rates over the rich-scattering wireless channel,” in *Signals, Systems, and Electronics, 1998. ISSSE 98. 1998 URSI International Symposium on*. IEEE, 1998, pp. 295–300.
- [6] G. J. Foschini, G. D. Golden, R. Valenzuela, P. W. Wolniansky *et al.*, “Simplified processing for high spectral efficiency wireless communication employing multi-element arrays,” *Selected Areas in Communications, IEEE Journal on*, vol. 17, no. 11, pp. 1841–1852, 1999.
- [7] J. Benesty, Y. Huang, and J. Chen, “A fast recursive algorithm for optimum sequential signal detection in a blast system,” *Signal Processing, IEEE Transactions on*, vol. 51, no. 7, pp. 1722–1730, 2003.
- [8] S. Loyka and F. Gagnon, “Performance analysis of the v-blast algorithm: an analytical approach,” *Wireless Communications, IEEE Transactions on*, vol. 3, no. 4, pp. 1326–1337, 2004.

- [9] N. Prasad and M. K. Varanasi, "Analysis of decision feedback detection for mimo rayleigh-fading channels and the optimization of power and rate allocations," *Information Theory, IEEE Transactions on*, vol. 50, no. 6, pp. 1009–1025, 2004.
- [10] Y. Jiang, X. Zheng, and J. Li, "Asymptotic performance analysis of v-blast," in *Global Telecommunications Conference, 2005. GLOBECOM'05. IEEE*, vol. 6. IEEE, 2005, pp. 5–pp.
- [11] M. O. Damen, H. El Gamal, and G. Caire, "On maximum-likelihood detection and the search for the closest lattice point," *Information Theory, IEEE Transactions on*, vol. 49, no. 10, pp. 2389–2402, 2003.
- [12] J. Jaldén and B. Otterste, "On the complexity of sphere decoding in digital communications," *Signal Processing, IEEE Transactions on*, vol. 53, no. 4, pp. 1474–1484, 2005.
- [13] L. G. Barbero and J. S. Thompson, "Fixing the complexity of the sphere decoder for MIMO detection," *Wireless Communications, IEEE Transactions on*, vol. 7, no. 6, pp. 2131–2142, 2008.
- [14] Z. Luo, M. Zhao, S. Liu, and Y. Liu, "Generalized parallel interference cancellation with near-optimal detection performance," *Signal Processing, IEEE Transactions on*, vol. 56, no. 1, pp. 304–312, 2008.
- [15] D. Radji and H. Leib, "Interference cancellation based detection for v-blast with diversity maximizing channel partition," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 3, no. 6, pp. 1000–1015, 2009.
- [16] A. Chockalingam, "Low-complexity algorithms for large-mimo detection," in *Communications, Control and Signal Processing (ISCCSP), 2010 4th International Symposium on*. IEEE, 2010, pp. 1–6.
- [17] K. V. Vardhan, S. K. Mohammed, A. Chockalingam, and B. S. Rajan, "A low-complexity detector for large MIMO systems and multicarrier CDMA systems," *Selected Areas in Communications, IEEE Journal on*, vol. 26, no. 3, pp. 473–485, 2008.
- [18] B. Cerato and E. Viterbo, "Hardware implementation of a low-complexity detector for large mimo," in *Circuits and Systems, 2009. ISCAS 2009. IEEE International Symposium on*. IEEE, 2009, pp. 593–596.
- [19] P. Li and R. D. Murch, "Multiple output selection-LAS algorithm in large MIMO systems," *Communications Letters, IEEE*, vol. 14, no. 5, pp. 399–401, 2010.
- [20] N. Srinidhi, T. Datta, A. Chockalingam, and B. S. Rajan, "Layered tabu search algorithm for large-MIMO detection and a lower bound on ML performance," *Communications, IEEE Transactions on*, vol. 59, no. 11, pp. 2955–2963, 2011.

- [21] T. Datta, N. Srinidhi, A. Chockalingam, and B. S. Rajan, "Random-restart reactive tabu search algorithm for detection in large-mimo systems," *Communications Letters, IEEE*, vol. 14, no. 12, pp. 1107–1109, 2010.
- [22] P. Som, T. Datta, N. Srinidhi, A. Chockalingam, and B. S. Rajan, "Low-complexity detection in large-dimension MIMO-ISI channels using graphical models," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 5, no. 8, pp. 1497–1511, 2011.
- [23] P. Som, T. Datta, A. Chockalingam, and B. S. Rajan, "Improved large-MIMO detection based on damped belief propagation," in *Information Theory Workshop (ITW), 2010 IEEE*. IEEE, 2010, pp. 1–5.
- [24] T. L. Narasimhan and A. Chockalingam, "Channel hardening-exploiting message passing (CHEMP) receiver in large-scale MIMO systems," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 8, no. 5, pp. 847–860, 2014.
- [25] J. Goldberger and A. Leshem, "MIMO detection for high-order QAM based on a Gaussian tree approximation," *Information Theory, IEEE Transactions on*, vol. 57, no. 8, pp. 4973–4982, 2011.
- [26] S. Mohammed, A. Chockalingam, and B. S. Rajan, "Low-complexity near-map decoding of large non-orthogonal stbcs using pda," in *Information Theory, 2009. ISIT 2009. IEEE International Symposium on*. IEEE, 2009, pp. 1998–2002.
- [27] T. Datta, N. A. Kumar, A. Chockalingam, and B. S. Rajan, "A novel monte-carlo-sampling-based receiver for large-scale uplink multiuser MIMO systems," *Vehicular Technology, IEEE Transactions on*, vol. 62, no. 7, pp. 3019–3038, 2013.
- [28] Q. Zhou and X. Ma, "Element-based lattice reduction algorithms for large MIMO detection," *Selected Areas in Communications, IEEE Journal on*, vol. 31, no. 2, pp. 274–286, 2013.
- [29] C. Blum and A. Roli, "Metaheuristics in combinatorial optimization: Overview and conceptual comparison," *ACM Computing Surveys (CSUR)*, vol. 35, no. 3, pp. 268–308, 2003.
- [30] D. E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*, 1st ed. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc., 1989.
- [31] V. Vapnik, "Pattern recognition using generalized portrait method," *Automation and remote control*, vol. 24, pp. 774–780, 1963.
- [32] V. Vapnik and A. Chervonenkis, "A note on one class of perceptrons," *Automation and remote control*, vol. 25, no. 1, 1964.
- [33] B. Schölkopf and A. J. Smola, *Learning with kernels: Support vector machines, regularization, optimization, and beyond*. MIT press, 2002.

- [34] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A training algorithm for optimal margin classifiers," in *Proceedings of the fifth annual workshop on Computational learning theory*. ACM, 1992, pp. 144–152.
- [35] I. Guyon, B. Boser, and V. Vapnik, "Automatic capacity tuning of very large vc-dimension classifiers," *Advances in neural information processing systems*, pp. 147–147, 1993.
- [36] V. Vapnik, *The nature of statistical learning theory*. Springer Science & Business Media, 2013.
- [37] C. Cortes and V. Vapnik, "Support-vector networks," *Machine learning*, vol. 20, no. 3, pp. 273–297, 1995.
- [38] B. Schölkopf, C. Burges, and V. Vapnik, "Incorporating invariances in support vector learning machines," in *Artificial Neural Networks ICANN 96*. Springer, 1996, pp. 47–52.
- [39] V. Vapnik, S. E. Golowich, and A. Smola, "Support vector method for function approximation, regression estimation, and signal processing," in *Advances in Neural Information Processing Systems 9*. Citeseer, 1996.
- [40] V. N. Vapnik and S. Kotz, *Estimation of dependences based on empirical data*. Springer-verlag New York, 1982, vol. 40.
- [41] A. J. Smola and B. Schölkopf, "A tutorial on support vector regression," *Statistics and computing*, vol. 14, no. 3, pp. 199–222, 2004.
- [42] E. Osuna, R. Freund, and F. Girosi, "Training support vector machines: an application to face detection," in *Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on*. IEEE, 1997, pp. 130–136.
- [43] —, "An improved training algorithm for support vector machines," in *Neural Networks for Signal Processing [1997] VII. Proceedings of the 1997 IEEE Workshop*. IEEE, 1997, pp. 276–285.
- [44] J. Platt *et al.*, "Fast training of support vector machines using sequential minimal optimization," *Advances in kernel methodssupport vector learning*, vol. 3, 1999.
- [45] S. Yang and L. Hanzo, "Fifty years of mimo detection: The road to large-scale mimos," *Communications Surveys Tutorials, IEEE*, vol. PP, no. 99, pp. 1–1, 2015.
- [46] P. Bouboulis, S. Theodoridis, C. Mavroforakis, and L. Evaggelatos-Dalla, "Complex support vector machines for regression and quaternary classification."
- [47] I. Steinwart, D. Hush, and C. Scovel, "Training svms without offset," *The Journal of Machine Learning Research*, vol. 12, pp. 141–202, 2011.