

Advances in Parameter Estimation, Source Enumeration, and Signal Identification for Wireless Communications

by

©Mostafa Mohammadkarimi

A Thesis submitted to the School of Graduate Studies in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

**Faculty of Engineering and Applied Science Memorial University of
Newfoundland**

Memorial University of Newfoundland

May 2018

St. John's

Newfoundland

Abstract

Parameter estimation and signal identification play an important role in modern wireless communication systems. In this thesis, we address different parameter estimation and signal identification problems in conjunction with the Internet of Things (IoT), cognitive radio systems, and high speed mobile communications.

The focus of Chapter 2 of this thesis is to develop a new uplink multiple access (MA) scheme for the IoT in order to support ubiquitous massive uplink connectivity for devices with sporadic traffic pattern and short packet size. The proposed uplink MA scheme removes the Media Access Control (MAC) address through the signal identification algorithms which are employed at the gateway.

The focus of Chapter 3 of this thesis is to develop different maximum Doppler spread (MDS) estimators in multiple-input multiple-output (MIMO) frequency-selective fading channel. The main idea behind the proposed estimators is to reduce the computational complexity while increasing system capacity.

The focus of Chapter 4 and Chapter 5 of this thesis is to develop different antenna enumeration algorithms and signal-to-noise ratio (SNR) estimators in MIMO time-varying fading channels, respectively. The main idea is to develop low-complexity algorithms and estimators which are robust to channel impairments.

The focus of Chapter 6 of this thesis is to develop a low-complexity space-time block codes (STBC)s identification algorithms for cognitive radio systems. The goal is to design an algorithm that is robust to time-frequency transmission impairments.

Acknowledgements

Among all people, I would like to express my deepest gratitude and appreciation to my adviser Prof. Octavia A. Dobre. Her contribution to my education and professional development is invaluable. I am so grateful, and I consider myself lucky to have had such an adviser. I would also like to thank Prof. Moe Z. Win who gave me the chance to spend one year as a visiting student in his research group at MIT. I wish to express my appreciation and thanks to the members of the supervisory committee, Prof. Weimin, and Prof. Venkatesan for their kindness and supportiveness. Last, but not least, I would like to thank my parents for their supportiveness. They are the primary reason for the enjoyable years I have spent in the university.

Co-authorship Statement

I, Mostafa Mohammadkarimi, have the principal author status for all the manuscripts included in this thesis. However, all manuscripts in this work are co-authored by my supervisor Prof. Octavia A. Dobre in addition to other co-researchers, who have valuable contributions which facilitated the development of this work. The list of the manuscripts included in this thesis are described as below.

- M. Mohammadkarimi, O. A. Dobre, and M. Z. Win, “Massive Uncoordinated Multiple Access for the Internet of Things., *Submitted to IEEE Trans. Inf. Theory.*, Aug 2017.
- M. Mohammadkarimi, O. A. Dobre, and M. Z. Win, “An alternative Mechanism for MAC Address in IoT"., *Submitted to ICC.*, Aug 2017.
- M. Mohammadkarimi, E. Karami, O. A. Dobre, and M. Z. Win, “Doppler Spread Estimation in MIMO Frequency-selective Fading Channels,” *Submitted to Trans. Wireless Commun.*, Feb. 2017.
- M. Mohammadkarimi, E. Karami, O. A. Dobre, and M. Z. Win, “Number of transmit antennas detection using time-diversity of the fading channel,” *IEEE Trans. Signal Process.*, vol. 65, no. 15, pp. 4031–4046, Aug. 2017.
- M. Mohammadkarimi, O. A. Dobre, and M. Z. Win, “Non-Data-Aided SNR

Estimation for Multiple Antenna Systems,” in *Proc. GLOBECOM*, Washington, USA, Dec. 2016, pp. 1–5.

- M. Mohammadkarimi and O. A. Dobre, “A novel non-parametric method for blind identification of STBC codes,” in *Proc CWIT*, Canada, July 2015, pp. 97–100.
- M. Mohammadkarimi, E. Karami, and O. A. Dobre, “A novel algorithm for blind detection of the number of transmit antenna,” in *Proc. CROWNCOM*, Oct. 2015, pp. 441–450.
- M. Mohammadkarimi and O. A. Dobre, “Blind identification of spatial multiplexing and alamouti space-time block code via kolmogorov- smirnov (K-S) test,” *IEEE Commun. Lett.*, vol. 18, no. 10, pp. 1711–1714, Oct 2014.

Table of Contents

Abstract	ii
Acknowledgments	iii
Table of Contents	xii
List of Tables	xiii
List of Tables	xiii
List of Figures	xiv
List of Figures	xx
1 Introduction	1
1.1 Proposal Outline	3
2 Massive Uncoordinated Multiple Access (MA) for the Internet of Things	5
2.1 Introduction	5
2.1.1 Literature Review	7
2.1.2 Motivation	10
2.1.3 Problem Statement	11

2.1.4	Methodology	12
2.2	Uplink MA for IoT	13
2.2.1	System Model	13
2.3	IoT Identification	19
2.3.1	IoT Identification Formulation	19
2.3.2	Squared ℓ_2 norm DSA IoT Identification	25
2.3.2.1	Bayesian Squared ℓ_2 norm DSA IoT Identification Algorithm	30
2.3.2.2	ML Squared ℓ_2 norm DSA IoT Identification Algorithm	31
2.3.3	ℓ_1 norm DSA IoT Identification	33
2.3.4	$\ell_1 - \ell_2$ Mixed-norm PDSA IoT Identification Algorithm	37
2.3.5	Adaptive $\ell_1 - \ell_2$ Mixed-norm PDSA IoT Identification Algorithm	41
2.4	Data Detection	44
2.4.1	2MC-MUD Algorithm	45
2.4.2	Downlink Transmission	51
2.5	Simulation Results	51
2.5.1	Simulation Setup	51
2.5.2	Simulation Results	52
2.6	Conclusions and Directions for Future Research	55
2.6.1	Summary	56
2.6.2	Future research	58
Appendices		59
2.A	Proof of the Tuning Parameter for the ℓ_0 Quasi-norm	59
2.B	Proof of the Tuning Parameter for the $\ell_0 - \ell_0$ Quasi-norm	60
2.C	Proof of the Tuning Parameter for the Squared ℓ_2 Norm	61
2.D	Distribution of the ridge regression (RD) Estimator	61

2.E	Proof of the Optimum Bayesian’s Decision Rule	66
2.F	Proof of the maximum likelihood ratio (MLR) decision rule	66
2.G	Proof of the Correct Identification and False Alarm Rates for the ML Squared ℓ_2 norm	67
2.H	Approximation of the Tuning Parameter λ_{ℓ_1} for the ℓ_1 norm DSA IoT Identification Algorithm	69
2.I	Approximation of the Tuning Parameter λ_{ℓ_1} for the ℓ_1 norm DSA IoT Identification algorithm	70
2.J	Proof of the KKT Optimality Conditions for the $\ell_1 - \ell_2$ Mixed-norm PDSA IoT Identification Algorithms	71
2.K	Equivalence of $\ell_1 - \ell_2$ Mixed-norm Algorithm with Iterative Reweighed LS Estimator	72
3	Doppler Spread Estimation in MIMO Frequency-selective Fading Channels	73
3.1	Introduction	73
3.1.1	Literature Review	74
3.1.2	Motivation	75
3.1.3	Problem Statement	76
3.1.4	Methodology	77
3.2	Maximum Doppler Spread (MDS) Estimation	77
3.2.1	System Model	77
3.2.2	CRLB for MDS Estimation	78
3.2.2.1	DA-CRLB	79
3.2.2.2	NDA-CRLB	81
3.2.3	maximum likelihood estimator (MLE) for MDS	83
3.2.3.1	DA-MLE for MDS	84

3.2.3.2	NDA-MLE for MDS	85
3.2.4	NDA-MB Estimation of MDS	85
3.2.4.1	NDA-MBE for MDS in MISO Systems	86
3.2.4.2	NDA-MBE for MDS in MIMO Systems	92
3.2.4.3	Semi-blind NDA-MBE	95
3.3	Complexity Analysis	96
3.4	Simulations	97
3.4.1	Simulation Setup	98
3.4.2	Simulation Results	99
3.5	Conclusions and Directions for Future Research	105
3.5.1	Summary	106
3.5.2	Future research	107
Appendices		108
3.A	Proof of (3.28)	108
3.B	Proof of (3.34)	109
3.C	Proof of (3.35)	110
4	Antenna Enumeration in Time-varying Fading Channels	111
4.1	Introduction	111
4.1.1	Literature Review	112
4.1.2	Motivation	114
4.1.3	Problem Statement	114
4.1.4	Methodology	115
4.2	Antenna Enumeration	115
4.2.1	System model	115
4.2.2	Blind Antenna Enumeration Algorithm	116

4.2.3	Semi-blind Antenna Enumeration Algorithm	119
4.3	Performance Analysis and Threshold setting	121
4.3.1	Performance Analysis	121
4.3.2	Threshold setting	127
4.4	Simulations	129
4.4.1	Simulation Setup	129
4.4.2	Simulation Results	131
4.5	Conclusions and Directions for Future Research	139
4.5.1	Summary	139
4.5.2	Future research	141
Appendices		142
4.A	Proof of the Second-order Statistics	142
4.B	Proof of the Unbiased Estimators in (4.8)-(4.9)	143
4.C	First-order Taylor Expansion of the Decision Statistics in (4.11)	144
4.D	Proof of the Mean Vector	147
4.E	Proof of the Covariance Matrix	149
5	SNR and Noise Variance Estimation in MIMO Time-varying Channels	154
5.1	Introduction	154
5.1.1	Literature Review	155
5.1.2	Motivation	157
5.1.3	Problem Statement	158
5.1.4	Methodology	158
5.2	Moment-based SNR Estimation	159
5.2.1	Sysyem Model	159

5.2.2	M ₂ M ₄ SNR Estimator	160
5.2.3	M ₂ M ₄ M ₆ SNR Estimator	162
5.2.4	Extension to MIMO Systems	164
5.3	Simulations	165
5.3.1	Simulation Setup	165
5.3.2	Simulation Results	165
5.4	Conclusions and Directions for Future Research	169
5.4.1	Summary	169
5.4.2	Future research	170
6	Blind STBC Identification	171
6.1	Introduction	171
6.1.1	Literature Review	172
6.1.2	Motivation	173
6.1.3	Problem Statement	174
6.1.4	Methodology	174
6.2	Identification of SM and AL-STBC	175
6.2.1	Signal Model	175
6.2.2	K-S Based Identification Algorithm	176
6.3	Extension to a Larger Pool of STBCs	179
6.4	Simulations	184
6.4.1	Simulation Setup	185
6.4.2	Simulation Results	185
6.5	Conclusions and Directions for Future Research	192
6.5.1	Summary	192
6.5.2	Future research	193

7 Overall Conclusions	194
References	195
Bibliography	196

List of Tables

2.1	MAC Protocols for IoT.	8
3.1	Number of real additions, real multiplications, and complexity order of the proposed non-data-aided (NDA)-moment-based estimator (MBE).	96
6.1	The effect of the SNR on the P_c	187
6.2	The effect of the number symbols on the P_c	188
6.3	The effect of the modulation type on the P_c	189
6.4	The Effect of the timing offset (TO) on the P_c	189

List of Figures

2.1	Single-hop IoT network with sporadic traffic pattern.	13
2.2	Block diagram that illustrates packet transmission at IoT devices. . .	14
2.3	Received packets at the gateway. Active IoT devices are shown with different colors.	15
2.4	Block diagram of the proposed receiver at the gateway.	16
2.5	Underdetermined systems of linear equations for with $K_u = 7$, $K_a = 2$, $N_t = 8$, and $N_s = 6$	20
2.6	Different observation windows for IoT identification ($K_u = 7$, $N_s = 7$, $\alpha_{\max} = 2$, $\alpha_{\min} = 0$, $1 \leq l \leq 4$). Purple color is employed to show the packet of IoT devices which is zero for inactive and non-zero for active IoT devices.	21
2.7	The probability mass function of k_a for different probability of packet transmission, P_a , when $K_u = 1024$. The dashed lines show K_{al} and K_{au} for $P_a = 0.05$	23
2.8	The decision statistics of the maximum likelihood (ML) squared ℓ_2 norm device sparsity-aware (DSA) IoT identification algorithm, i.e., $\phi(\check{\mathbf{h}}_{k,l})$ versus k , for $l = 10$, $K_u = 2048$, $P_a = 0.01$, $P_k^{(f)} = 0.01$, $k = 0, 1, \dots, K_u - 1$, and at 10 dB SNR.	32

2.9	Block diagram of the proposed squared ℓ_2 norm DSA IoT identification algorithms. For Algorithm 1 and Algorithm 2, $\phi(\hat{\mathbf{h}}_{k,\bar{\alpha},l})$ is given by (2.49) and (2.53), respectively.	34
2.10	Block diagram of the proposed 2-mean clustering (2MC)-multiuser detection (MUD) algorithm	46
2.11	2M classifier for short IoT packet.	49
2.12	The correct identification rate, $P^{(c)}$, and false alarm rate, $P^{(f)}$, of the designed ML squared ℓ_2 norm DSA IoT identification algorithm (Algorithm 2) versus SNR for $K_u = 2048$, $P_a = 0.01$, and $l = 16$. The solid and dashed lines represent the results for $P_k^{(f)} = 0.01$ and $P_k^{(f)} = 0.02$, respectively.	53
2.13	The empirical correct identification rate, $P^{(c)}$, of the designed ML squared ℓ_2 norm DSA IoT identification algorithm (Algorithm 2) versus time for $K_u = 2048$, $P_a = 0.01$, $l = 16$, and at 10 dB SNR.	54
2.14	The correct identification rate, $P^{(c)}$, and false alarm rate, $P^{(f)}$, of the designed ℓ_1 norm DSA IoT identification algorithm (Algorithm 3) versus SNR for $K_u = 2048$, $P_a = 0.01$, and $l = 16$	54
2.15	Performance comparison of the proposed $\ell_1 - \ell_2$ mixed-norm (Algorithm 4) and adaptive $\ell_1 - \ell_2$ mixed-norm (Algorithm 5) packet-device sparsity-aware (PDSA) IoT identification algorithms.	55
2.16	packet error rate (PER) of the proposed 2MC-MUD algorithm versus SNR when the proposed $\ell_1 - \ell_2$ mixed-norm PDSA IoT identification algorithm (Algorithm 5) is employed at the gateway.	56
3.1	Illustration of the non-linear LS regression for the uniformly sampled normalized squared AF for $f_D T_s = 0.02$ and $f_D T_s = 0.005$, with $n_t = 1$, $n_r = 2$, $L = 1$, $u_s = 2$, and at $\gamma = 10$ dB.	94

3.2	Computational complexity comparison of the proposed NDA-MBE, the low-complexity data-aided (DA)-MLE in [1, 2], and the DA-COMAT estimator in [3].	97
3.3	Distribution of the estimated \hat{f}_D for the proposed NDA-MBE for $n_t = 2$, $n_r = 2$, and at $\gamma = 10$ dB.	98
3.4	The mean value of the estimated MDS by NDA-MBE for various SNR values for $n_t = 2$ and $n_r = 2$	100
3.5	The NRMSE of the proposed NDA-MBE versus $f_D T_s$ for different SNR values.	100
3.6	The effect of n_t on the performance of the proposed NDA-MBE for $n_r = 2$ and at $\gamma = 10$ dB.	101
3.7	The effect of n_r on the performance of the proposed NDA-MBE for $n_t = 2$ and at $\gamma = 10$ dB.	102
3.8	The effect of the parameter U_{\min} on the performance of the proposed NDA-MBE for different values of $f_d T_s$ and at $\gamma = 10$ dB.	102
3.9	The effect of the observation window size, N , on the performance of the proposed NDA-MBE for $n_t = 2$ and $n_r = 2$ in frequency-selective channel, and at $\gamma = 10$ dB.	103
3.10	Performance comparison of the proposed NDA-MBE, DA-CRLB [1, 2], the low-complexity DA-MLE in [1, 2], and the NDA-CCE in [4] in single-input single-output (SISO) frequency-flat Rayleigh fading channel for $N = 10^3$ at $\gamma = 10$ dB.	104
3.11	Performance comparison of the NDA-MBE, the derived DA-CRLB, and the derived DA-MLE in MIMO frequency-selective Rayleigh fading channel for $n_t = 2$, $n_r = 2$, $L = 5$, $N = 10^3$, and at $\gamma = 10$ dB.	105

3.12	Performance comparison of the proposed NDA-MBE, derived NDA-MLE, derived NDA-Cramer-Rao lower bound (CRLB), and DA-CRLB in SISO flat-fading channel for $N = 10$ and at $\gamma = 20$ dB.	106
4.1	The equivalent two-dimensional test under hypothesis H_3 for $\Gamma_2 = 2.5$ and $\Gamma_3 = 3.5$ at SNR $\gamma = -10$ dB.	123
4.2	Illustration of the estimated joint PDF of y_2 and y_3 under hypothesis H_3 , $f(y_2, y_3 H_3)$	124
4.3	The conditional probability of correct detection, $\mathbb{P}\{\hat{n}_t = n_t H_{n_t}\}$ versus SNR (γ) for different numbers of transmit antennas, n_t . Results are presented for the proposed blind antenna enumeration algorithm. . .	131
4.4	The conditional probability of correct detection, $\mathbb{P}\{\hat{n}_t = n_t H_{n_t}\}$, of the proposed semi-blind antenna enumeration algorithm versus SNR for different n_t and N_b values.	132
4.5	\mathbb{P}_D versus SNR (γ) for different threshold settings. Scenario 1: optimal thresholds $\Gamma_1, \Gamma_2, \Gamma_3$, $\Gamma_0 = -\infty$ and $\Gamma_4 = +\infty$; Scenario 2: $\Gamma_{n_t} = n_t + 1/2$, $n_t = 1, 2, 3$, $\Gamma_0 = -\infty$, $\Gamma_4 = +\infty$; and Scenario 3: $\Gamma_{n_t} = n_t + 1/2$, $n_t = 1, 2, 3, 4, \dots$, $\Gamma_0 = -\infty$. Results are presented for the proposed blind antenna enumeration algorithm.	133
4.6	The effect of an incorrect N_t value ($N_t=5$) on the conditional probability of correct detection of the proposed blind antenna enumeration algorithm with optimal threshold setting (Scenario 1), $\mathbb{P}\{\hat{n}_t = n_t H_{n_t}\}$, $n_t = 1, 2, 3$, when the true value of N_t is 3.	134
4.7	The effect of the carrier frequency offset (CFO) on the performance of the proposed blind antenna enumeration algorithm.	135
4.8	The effect of the modulation type on the performance of the proposed blind antenna enumeration algorithm.	136

4.9	The effect of the spatially correlated fading on the performance of the proposed blind antenna enumeration algorithm.	136
4.10	The effect of TO for root-raised-cosine (RRC) filter with $\beta = 0.35$ and low-pass 10th-order Butterworth filter with cutoff frequency $w_c = 0.2\pi$ at receive side. Results are presented for the proposed blind antenna enumeration algorithm.	137
4.11	The effect of the Doppler frequency, f_d , on the conditional probability of correct detection, $\mathbb{P}\{\hat{n}_t = n_t H_{n_t}\}$, for $N_c = \lfloor 9/16\pi f_d T_s \rfloor$, $NT_s = 1$ sec at SNR=0 dB. Results are presented for the proposed blind antenna enumeration algorithm.	138
4.12	The performance of the proposed blind antenna enumeration algorithm with noise power estimate (NPE) [5].	139
4.13	Performance comparison of the proposed blind antenna enumeration algorithm with $n_r = 1$ and the soft MDL and AIC approaches in [6] with $n_r = 5$, for $n_t = 2$	140
4.C.1	$\text{Var}\{u_2\}/\mu_{z_2}^2$ for different values of N_c and N_b at SNR=-10 dB.	147
4.C.2	$\mathbb{E}\{u_1 u_2\}/\mu_{z_2}^2$ for different values of N_c and N_b at SNR=-10 dB.	148
4.C.3	$\mathbb{E}\{\Upsilon\}$ for different values of N_c and N_b at SNR=-10 dB.	149
5.1	Distribution of the estimated SNR, $\hat{\gamma}$, for the proposed estimators with $n_t = 2$ and at $\gamma = 0$ dB.	166
5.2	The normalized bias of the M_2M_4 estimator versus SNR, γ , for $n_t = 1, 2, 3, 4$	166
5.3	NRMSE of the M_2M_4 estimator versus SNR, γ , for $n_t = 1, 2, 3, 4$	167
5.4	The normalized bias of the $M_2M_4M_6$ estimator versus SNR, γ , for $n_t = 1, 2, 3, 4$	168
5.5	NRMSE of the $M_2M_4M_6$ estimator versus SNR, γ , for $n_t = 1, 2, 3, 4$	168

6.1	Illustration of the event AL_1 . Solid lines are used to delimitate symbols which do not belong to the same block and dashed lines are employed to delimitate symbols which belong to the same block.	176
6.2	Flowchart of the proposed binary decision tree for the identification of SM, AL, and STBC3.	181
6.3	Illustration of the sequences \mathbf{y}_2 and \mathbf{z}_2 for the four different STBC3 events, corresponding to the four possible starting symbols which belong to an STBC block, for $K = 20$. Solid lines are used to delimitate symbols which do not belong to the same block and dashed lines are employed to delimitate symbols which belong to the same block. . . .	183
6.4	Illustration of the sequences \mathbf{y}_2 and \mathbf{z}_2 for the two different AL events, corresponding to the two possible starting symbols which belong to an AL block, for $K = 20$. Solid lines are used to delimitate symbols which do not belong to the same block and dashed lines are employed to delimitate symbols which belong to the same block.	184
6.5	The probability of correct identification, $\mathbb{P}\{\hat{\theta} = \theta \theta\}$, $\theta \in \{\text{SM}, \text{AL}\}$, versus SNR for the K-S-based algorithm over different Nakagami- m fading channels in the presence of impulsive noise, with QPSK modulation, $K = 2048$, $\Delta f = 0.01$, and $\beta = 0.99$	186
6.6	The effect of the confidence interval, β , on the probability of correct identification, $\mathbb{P}\{\hat{\theta} = \theta \theta\}$, $\theta \in \{\text{SM}, \text{AL}\}$, for the K-S-based algorithm over Nakagami- m fading channel, $m = 2$, in the presence of impulsive noise, with QPSK modulation, $K = 2048$, and $\Delta f = 0.01$. .	187

6.7	The effect of the normalized CFO, Δf , on the average probability of correct identification, P_c , for the ML [7], FOLP-C [8], and K-S-based algorithms over Nakagami- m fading channel, $m = 2$, in the presence of impulsive noise, with QPSK modulation, $K = 2048$, $\beta = 0.99$, and at SNR=10 dB.	188
6.8	Average probability of correct identification, P_c , $\theta \in \{\text{SM}, \text{AL}, \text{STBC3}\}$, versus SNR for diverse Nakagami- m fading channels.	191
6.9	Effect of the CFO normalized to the data rate, Δf , on P_c for $\theta \in \{\text{SM}, \text{AL}, \text{STBC3}\}$ in Nakagami- m fading channel, $m = 3$	191
6.10	Performance comparison of the proposed method, FOLP [8], and second-order cyclostationarity-based [9] for $\theta \in \{\text{SM}, \text{AL}, \text{STBC3}\}$ in Nakagami- m fading channel, $m = 3$ and $\Delta f = 0$	192

Notations and Symbols

AP	access point
AF	autocorrelation function
AIC	Akaike Information Criterion
AL-STBC	Alamouti Space-Time Block
BCH	Bose–Chaudhuri–Hocquenghem
BER	bit error rate
BIC	Bayesian Information Criterion
BLE	Bluetooth low energy
BPSK	binary phase-shift keying
BS	base station
CA	collision avoidance
CB	contention-based
CC	cyclic correlation
CCE	cyclic correlation estimator
CDMA	code division multiple access
CFO	carrier frequency offset
CP	carrier phase
CRLB	Cramer Rao Lower Bound
CSMA	carrier sensing multiple access

CSS	chirp spread spectrum
CV	cross-validation
CVPL	cross-validated partial likelihood
DA	data-aided
DD	decision-directed
DDE	decision-directed estimator
DOA	Direction-Of-Arrival
DS	direct sequence
DSA	device sparsity-aware
DA-CRLB	data-aided Cramer-Rao Lower Bound
DS-SS	direct-sequence spread spectrum
ECDF	empirical cumulative distribution function
EM	expectation maximization
FA	fixed assignment
FB	Feature-Based
FCC	fading channel coefficient
FDMA	frequency division multiple access
FIM	Fisher information matrix
GCV	generalized cross validation
FSA	frame slotted ALOHA
i.i.d.	independent and identically distributed
ICT	Information Theoretic Criterion
IoT	Internet of Things
KKT	Karush Kuhn Tucker
KL	Kullback Leibler
K-S	Kolmogrov-Smirnov

LBTlisten-before-talk
LPWAlow-power wide area
LSleast-square
LTELong-Term Evolution
LTE-MLong Term Evolution for Machines
LWCSSleast within-cluster sum of squares
MA.....multiple access
MACmedia access control
MAPmaximum a posteriori probability
MB.....moment-based
MBE.....moment-based estimator
MDL.....minimum description length
MDSmaximum Doppler spread
MFmatched filter
MISOmultiple-input single-output
MIMOMultiple-Input Multiple-Output
MLmaximum likelihood
MLEmaximum likelihood estimator
MMSEminimum mean squared error
MSEmean square error
MTC.....machine-type communications
MUDmultiuser detection
MUSICmultiple signal classifier
NBnarrow band
NDAnon-data aided
NPEnoise power estimate

NRMSE normalized root-mean-square error
OFDM Orthogonal Frequency-Division Multiplexing
PDF probability density function
PDSA packet-device sparsity-aware
PER packet error rate
PHY Physical
PMF probability mass function
P-LS penalized-LS
PSD Power Spectral Density
RA random access
RD ridge regression
RFID radio frequency identification
RMSE Root-Mean Square Error
RPMA random phase multiple access
R-FTDM random frequency-time division multiplexing
SIMO single-input multiple-output
SISO Single-Input Single-Output
SM Spatial Multiplexing
SNR Signal-to-Noise Ratio
SS spread spectrum
SSSR simultaneous sparse signal reconstruction
SSR sparse signal reconstruction
STBC Space-Time Block Code
TDMA time-division multiple access
TO Timing Offset
TOA Time-Of-Arrival

WCSS within-cluster sum of squares
2MC 2-mean clustering
3GPP 3rd Generation Partnership Project

Notations and Symbols

Random variables are displayed in sans serif, upright fonts; their realizations in serif, italic fonts. Vectors and matrices are denoted by bold lowercase and uppercase letters, respectively. For example, a random variable and its realization are denoted by x and x ; a random vector and its realization are denoted by \mathbf{x} and \mathbf{x} ; a random matrix and its realization are denoted by \mathbf{X} and \mathbf{X} , respectively. Sets and random sets are denoted by upright sans serif and calligraphic font, respectively. For example, a random set and its realization are denoted by X and \mathcal{X} , respectively. The matrix \mathbf{I}_n is the identity matrix of size n , and $\mathbf{0}_q \triangleq [0 \ 0 \ \dots \ 0]_{q \times 1}^\dagger$. The indicator function is defined as

$$\mathbb{I}\{x\} \triangleq \begin{cases} 1, & \text{if } x \text{ is true} \\ 0, & \text{otherwise.} \end{cases}$$

The sign operator is

$$(x)_+ \triangleq \begin{cases} x, & \text{if } x > 0 \\ 0, & \text{otherwise.} \end{cases}$$

The ℓ_0 quasi-norm of vector $\mathbf{a}_j \triangleq [a_{0,j} \ a_{1,j} \ \dots \ a_{m-1,j}]^\dagger$ and the $\ell_0 - \ell_0$ quasi-norm of matrix $\mathbf{A} \triangleq [\mathbf{a}_0 \ \mathbf{a}_1 \ \dots \ \mathbf{a}_{n-1}]$ are respectively defined as

$$\|\mathbf{a}_j\|_0 \triangleq \mathbf{card}\left(\left\{i \in \{0, 1, \dots, m\} \mid a_{i,j} \neq 0\right\}\right),$$

and

$$\|\mathbf{A}\|_0 \triangleq \mathbf{card}\left(\left\{i \in \{0, 1, \dots, m\} \mid \exists j, j = 0, 1, \dots, n-1, a_{i,j} \neq 0\right\}\right).$$

The Frobenius and ℓ_q norm of vector \mathbf{a} , $q > 0$ are defined as

$$\|\mathbf{a}\|_{\text{F}} \triangleq \sqrt{\sum_i |a_i|^2},$$

and

$$\|\mathbf{a}\|_q \triangleq \left(\sum_i |a_i|^q\right)^{1/q},$$

respectively, The Frobenius and $\ell_p - \ell_q$ mixed-norm of matrix \mathbf{A} are defined as

$$\|\mathbf{A}\|_{\text{F}} \triangleq \sqrt{\sum_i \sum_j |a_{i,j}|^2},$$

$$J_{p,q}(\mathbf{A}) \triangleq \sum_i \left(\sum_j |a_{i,j}|^q\right)^{\frac{p}{q}},$$

respectively. The trace of an n -by- n square matrix \mathbf{A} is $\text{tr}(\mathbf{A}) \triangleq \sum_{i=0}^{n-1} a_{ii}$, where a_{ii} is the i th element in the i th column of \mathbf{A} . The inverse and determinant of matrix \mathbf{A} are denoted by \mathbf{A}^{-1} and $\det(\mathbf{A})$, respectively. For a square matrix \mathbf{A} , $\text{diag}(\mathbf{A})$ denotes the vector of the diagonal elements of matrix \mathbf{A} .

Complex Gaussian distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ is denoted by $\mathcal{N}_c(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

Throughout the thesis, $(\cdot)^*$ is used for the complex conjugate, $(\cdot)^\dagger$ is used for transpose, $(\cdot)^{\text{H}}$ is used for Hermitian, $|\cdot|$ represents the absolute value operator, $[\cdot]$

is the floor function, $\delta_{i,j}$ denotes the Kronecker delta function, \otimes is the Kronecker product, $\mathbb{E}\{\cdot\}$ is the statistical expectation, and \hat{x} is an estimate of x .

For simplicity of notation, we define

$$\sum_{f_1:f_i}^F \triangleq \sum_{f_1=1}^F \sum_{\substack{f_2=1 \\ f_2 \neq f_1}}^F \cdots \sum_{\substack{f_i=1 \\ f_i \neq \{f_1, \dots, f_{i-1}\}}}^F .$$

Chapter 1

Introduction

Parameter estimation [10] and signal identification [11] play critical roles in accurately describing behavior of a communication system. Regarding parameter estimation, the goal is to estimate a vector of real- or complex-valued from the observation samples. In general, it is categorized as non-parametric and parametric approaches [10, 12]. With the parametric approaches, not only a vector of real- or complex-valued is estimated, but also estimation of one or several integer-valued parameters is required. For instance, estimating the number of sinusoids in white noise, the number of signal sources impinging on a sensor array, the number of multipath components of fading channel, and the orders of an autoregressive moving average model [13–15]. On the other hand, signal identification aims to determine the type of the transmitted signal, such as modulation format and type of the transmitting code [11, 16].

In this thesis, I address different parameter estimation and signal identification problems related to modern wireless communication systems, such as the Internet of Things (IoT), cognitive radio, and high speed mobile communications [17–19].

Firstly, a new multiple access (MA) scheme based on joint parametric parameter estimation and signal identification is proposed. The proposed MA scheme is devel-

oped to support massive uplink connectivity in the IoT applications with sporadic and short packet transmission. The proposed MA scheme is designed to reduce control signaling associated with the Physical (PHY) and Media Access Control (MAC) layers to increase capacity and spectral efficiency of the system. Hence, the preambles and pilots associated with the MAC address and exploited for parameter estimation are removed. Since the IoT devices do not use MAC address to identify themselves to the gateway, the proposed alternative approach to the MAC address requires IoT identification algorithm at the gateway to determine the active IoT devices before data detection. Because the number of active IoT devices and their identity are not known at the gateway, we have a joint parametric parameter estimation and signal identification problem.

Regarding parametric parameter estimation, the problem of antenna enumeration in time-varying fading channel is also investigated . Challenged by antenna enumeration in time-varying fading channel, different algorithms for counting the number of transmit antennas, applicable to mobile cognitive radio and adaptive wireless communication systems are proposed. In order to determine the number of transmit antennas, which has discrete values, the antenna enumeration problem is formulated as a multiple-hypothesis testing problem and a thresholds setting mechanism is developed. As performance measure in parametric parameter estimation problem, the probability of correct number of transmit antennas detection is also derived.

In conjunction with non-parametric parameter estimation, the problem of maximum Doppler spread (MDS) estimation in multiple-input multiple-output (MIMO) frequency-selective fading channel is studied. Estimation of MDS is of importance since it represents a key parameter in determining the rate of change in wireless communication channels [20, 21]; thus, its knowledge is required in adaptive transceivers and in cellular communications [22, 23]. As a bench-mark for performance comparison in non-

parametric parameter estimation, the data-aided (DA) and non-data-aided (NDA) Cramer-Rao lower bounds (CRLB)s for MDS estimation in MIMO frequency-selective fading channel are derived. Also, a new low-complexly NDA MDS estimator for fast-varying fading channel is proposed. Regarding non-parametric parameter estimation, I also study the problem of NDA signal-to-noise ratio (SNR) and noise variance estimation in MIMO time-varying channel as non-parametric parameter estimation problems. Challenged by SNR estimation in time-varying channel, different low-complexity algorithms are developed through the statistical moment-based approach.

Regarding signal identification, the process of identifying the type of the transmitted space-time block code (STBC) from a pool of candidates [11, 16], called STBC identification, finds applications to modern wireless communication systems, such as cognitive radio and adaptive communications [24, 25]. Challenged by the imperfect time-frequency synchronization and non-Gaussian noise, the STBC identification problem is investigated, and a new algorithm through the Kolmogrov-Smirnov (K-S) test and binary decision tree is proposed. As a performance measure in signal identification problems, the probability of correct identification is provided for the proposed algorithm.

It should be mentioned that the five mentioned problems will be addressed in independent chapters, and introduction and literature review for each problem will be independently provided.

1.1 Proposal Outline

Chapter 2 deals with uplink MA for IoT. Chapter 3 studies the problem of MDS estimation in MIMO frequency-selective fading channel. Chapter 4 investigates the problem of antenna enumeration in time-varying fading channel. Chapter 5 deals with

SNR estimation in MIMO time-varying fading channel. Chapter 6 investigates the problem of STBC identification. Finally, in Chapter 7, overall conclusions are drawn.

Chapter 2

Massive Uncoordinated Multiple Access (MA) for the Internet of Things

2.1 Introduction

Massive uplink connectivity is the key factor in the realization of the Internet of Things (IoT), as part of 5G wireless communication systems [17, 26, 27]. IoT is a recent communication paradigm that enables the objects of everyday life to efficiently communicate with one another and with the users through wireless networks connected to the Internet. In general, IoT networks comprise a collection of connected objects, embedding electronics, software, sensors, and wireless connectivity protocols that collect and exchange information through wireless networks. Through interaction of a wide variety of physical devices or things, such as home appliances, sensors, surveillance cameras, and actuators, IoT fosters the development of applications in many different domains, such as smart cities, health care, transportation and auto-

motive environment, as well as utilities [28–30].

The development of IoT is an extremely challenging task, as several issues concerning the layers of the protocol stack, from the physical layer transmission to data representation, need to be addressed [31–33]. In conjunction with the Physical (PHY) and Media Access Control (MAC) layers, the massive connection of IoT devices is an important challenge. Typically, the number of IoT devices allocated to a single gateway is in orders of magnitude above what current communication networks are capable to support.

Furthermore, in the majority of the IoT applications, IoT devices do not transmit continuously. Such transmissions are characterized as sporadic, in which updates are infrequently transmitted to the gateway, whenever a measured value changes. Hence, small packets are expected to carry critical payload in IoT [34, 35]. For example, Sigfox as one of the most adopted solutions for IoT can support a maximum packet payload of 12 bytes [36].

The design of the current wireless communication systems relies on the assumption that the control signaling related to PHY and MAC layers is of negligible size compared to the payload. Thus, transmission of control signaling does not affect the overall system performance. However, in IoT applications with short packet transmission, the control signaling may be of the same size as the payload. Hence, excessive control signaling, e.g., the preambles and pilots associated with the MAC address¹ and exploited for parameter estimation, significantly reduce capacity and spectral efficiency of the system [36–38].

Moreover, channel estimation is another challenge for sporadic communication, especially for an enormous number of connections in IoT. The standard channel estimation approaches are often based on the assumption that devices are active

¹This represents the hardware identification address.

over long periods. However, if an IoT device only transmits every so often, such assumptions cannot longer be valid. Instead, channel estimation has to rely on a single transmission that may be very short, which constrains the number of pilots available to keep the overhead low [39].

In the context of IoT, the existing MAC protocols based on random access (RA), either ALOHA or carrier sensing multiple access (CSMA), suffer from congestion since the traffic load and the number of IoT devices is significantly large [40]. One solution to this problem is to employ smaller and denser cells. This, a lower number of IoT devices can be dedicated to each gateway/access point, and, the congestion problem is solved at the expense of higher cost due to the deployed infrastructure. In some cases this approach may not constitute a cost-effective solution given the capacity requirements of the majority of IoT applications. On the other hand, both RA and fixed assignment (FA) protocols employ excessive control signaling.

2.1.1 Literature Review

Different wireless technologies are currently under investigation for the realization of the IoT vision. These technologies are mainly categorized as: i) short-range, ii) cellular, and iii) low-power wide area (LPWA) as shown in Table 2.1. The existing MAC layer protocols associated with these technologies rely on hybrid schemes that employ contention- and schedule-based access mechanisms to support short packet transmission and massive uplink connectivity [41, 42].

IEEE 802.15.4, Bluetooth low energy (BLE), radio frequency identification (RFID), and Wi-Fi are short-range wireless technologies developed to support a potentially large number of IoT devices. Based on the topology of the network, the IEEE 802.15.4 standard employs a slotted CSMA with collision avoidance (CA) for single-hop and

Table 2.1: MAC Protocols for IoT.

MAC Protocol	Short-range				Cellular		LPWA			
	BLE	ZigBee	RFID	WiFi	LTE-M	NB-IoT	Sigfox	INGENU	LoRax	Weightless
Fixed assignment	TDMA	×								×
	FDMA				×	×				
	CDMA					×			×	
	Time slotted reservation				×	×			×	
Random access	Pure ALOHA						×		×	×
	Slotted ALOHA			×	×	×				
	Slotted CSMA-CS		×							
	Non-slotted CSMA-CS		×		×					
Spread spectrum	FH-SS	×								×
	DS-SS		×		×			×		
	Chirp-SS		×						×	

a contention-based MAC employing a simple non-slotted CSMA/CA mechanism for multi-hop topologies [43–45]. The BLE or smart Bluetooth is a modified version of the classic Bluetooth intended to provide considerable reduced power consumption and cost while maintaining a similar communication range. The MAC protocol in BLE is based on a time slotted access mechanism with a time division multiplexing technique applied to coordinate the medium access [46].

The ability to uniquely identify a large number of devices is critical for the success of IoT. This ability is provided by RFID through a contention-based MAC protocol which relies on uncoordinated frame slotted ALOHA (FSA) [47, 48].

The WiFi alliance supports the IEEE 802.11 family of standards. The recent standard proposed by WiFi, i.e, the IEEE 802.11ah, relies on a CSMA/CA scheme with a slotted binary exponential backoff mechanism for retransmissions in case of collision. The developed MAC protocol in the IEEE 802.11ah leads to an increased number of IoT devices supported by a single access point [49–51].

Ubiquitous infrastructure, large coverage, and mobility capability enable existing wireless cellular networks to support massive connectivity in IoT [26, 52, 53]. Among

the existing cellular networks, the Long-Term Evolution (LTE) provides a suitable connectivity for the IoT applications [54]. Recently, the 3rd Generation Partnership Project (3GPP) introduced Long Term Evolution for Machines (LTE-M) and narrow band (NB)-IoT as cellular solutions for IoT. The multiple access (MA) protocols in LTE-M can be either RA or FA. In the delay-constrained IoT applications, MA is performed based on the FA protocol in which the base station (BS) allocates a channel to the request with higher probability of success. The RA protocol in LTE relies on FSA [55–58]. NB-IoT is a new NB radio channel access which employs the guard bands between channels in the LTE to increase the transmission coverage and support a huge number of IoT devices. NB-IoT utilizes pure ALOHA and frequency division multiple access (FDMA) depending on the coverage. Moreover, collision on the random access channel in NB-IoT is handled by use of overlaid code division multiple access (CDMA) [59–62].

A promising wireless technology for IoT applications, standing between short-range and cellular technologies, is the LPWA. This offers a unique set of features including wide-area connectivity for low-power and low-cost IoT devices. Its performance is optimized for maximum coverage and battery efficiency while supporting a huge number of IoT devices with a single gateway. Sigfox, INGENU, LoRa, and Weightless are some of the widely-deployed LPWA solutions for IoT [63–66]. Sigfox utilizes ultra NB technique to enable long-range communication for IoT applications with very low data rates. It relies on random frequency-time division multiplexing (R-FTDM) as a pure ALOHA MAC protocol, in which each IoT device asynchronously transmits at a frequency chosen randomly in the continuous available frequency band [67]. INGENU proposes a proprietary LPWA technology based on a patented MA scheme named random phase multiple access (RPMA). This is a variation of CDMA, in which the traditional CDMA is randomly delayed before

transmission [68]. The LoRa Alliance promotes the use of LoRa and LoRaWAN technologies for the IoT applications [69]. The PHY layer of LoRa is based on chirp spread spectrum (CSS) techniques and its MAC protocol varies for three different developed classes: class-A devices employ pure ALOHA along with the listen-before-talk (LBT) mechanism; class-B devices are also developed based on the LBT mechanism along with a beacon-enabled time-slotted communication scheme; class-C devices are always available for reception, except when transmitting [69, 70]. Weightless employs a master-slave architectural model and each MAC frame consists of a downlink part followed by an uplink one. The BS (master) allocates uplink transmission opportunities to the IoT devices (slaves). This allocation is transmitted in downlink slots, while transmissions occur in the uplink slots. The MAC protocol in Weightless is a combination of FDMA and time-division multiple access (TDMA) schemes. Moreover, Weightless specification employs various mechanisms based on pure ALOHA and direct-sequence spread spectrum (DS-SS) to reduce the increased number of collisions.

2.1.2 Motivation

After reviewing the existing MAC protocols for IoT in the literature, the following observations can be made:

The existing wireless communication systems have been mainly designed with the objective of providing substantial gain in terms of data rates. However, 5G will depart from this scheme, and its objective will not only be to provide services with higher data rates. One of the main goals of 5G is to support machine-type communications (MTC) to enable pervasive connections of the entire world in order to realize the IoT. One of the main challenges in the realization of IoT is the capability to support ubiquitous massive uplink connectivity for devices with sporadic traffic pattern and short packet size. In the long packet transmission, the payload contained in a packet is much larger

than the control signaling associated with the PHY and MAC layers. However, in the short packet transmission, the control signaling are is negligible in size compared to payload. Hence, transmitting control information significantly affects the overall system performance. Moreover, channel estimation is a challenging issue for sporadic communication, especially for enormous number of connections. The standard channel estimation approaches are often based on the assumption that devices are active over long periods. However, IoT devices only transmit every so often in a large number of applications. Thus, channel estimation has to rely on a single transmission that may be very short, which constrains the number of pilots available to keep the overhead low.

2.1.3 Problem Statement

The specific research problems which are studied in Chapter 2 of this thesis are presented as follows:

- A new uplink MA scheme for IoT applications with sporadic traffic pattern and short packet transmission is proposed. The main idea behind the proposed MA scheme is to reduce the control signaling while simultaneously supporting a massive number of IoT devices with a single gateway. The proposed MA scheme is designed based on the DS-SS technique with non-orthogonal spreading codebook capable of supporting undetermined systems;
- To reduce the control signaling associated with the MAC address, a unique spreading code is dedicated to each IoT device which is simultaneously used for spreading purpose and MAC address. In other words, instead of allocating a fragment of the IoT packet to the signaling associated with the MAC address, the unique spreading code is used as the IoT device identifier. Moreover, the

MA scheme relies on some statistics of the fading channel coefficient (FCC) to remove the need of preambles and pilots which are employed for channel and carrier phase (CP) estimation. The lack of preambles and pilots further reduces the control signaling;

- The proposed alternative approach to the MAC address requires IoT identification at the gateway to determine active IoT devices before data detection. Based on the sporadic traffic pattern of the IoT devices, and lack of knowledge about FCC and CP of the IoT devices, the device sparsity-aware (DSA) and packet-device sparsity-aware (PDSA) identification algorithms are developed;
- A new non-linear multiuser detection (MUD) algorithm for short packet transmission is designed. The designed MUD is employed in the proposed MA scheme in order to detect data of the active IoT devices identified through the IoT identification algorithms.

2.1.4 Methodology

Due to the sporadic traffic pattern of the IoT devices, the problem of the IoT identification in the proposed MA scheme is formulated as sparse signal reconstruction (SSR) and simultaneous sparse signal reconstruction (SSSR) problems. Moreover, I design the 2-mean clustering (2MC)-MUD algorithm based on differential coding and binary phase-shift keying (BPSK) modulation at IoT devices, and 2MC unsupervised machine learning algorithm along with differential decoding at the gateway.

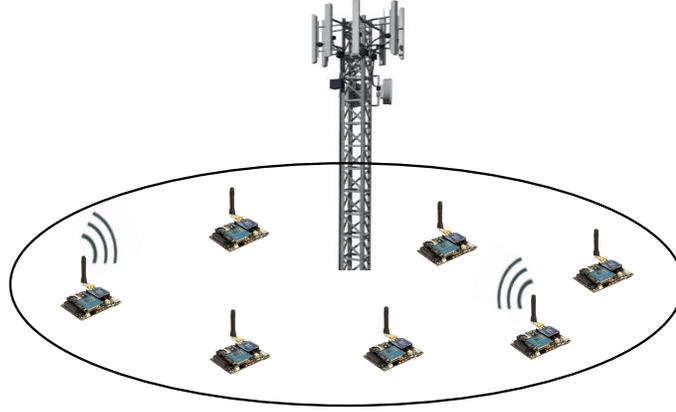


Fig. 2.1: Single-hop IoT network with sporadic traffic pattern.

2.2 Uplink MA for IoT

2.2.1 System Model

Consider K_u IoT devices communicating with a single IoT gateway in a single-hop communication, as shown in Fig. 2.1.

It is considered that IoT devices transmit data in short packets, where each packet carries only payload bits. The probability of packet transmission for each IoT device is assumed to be $P_a \ll 1$. Let us consider that $\mathcal{X}_u \triangleq \{0, 1, \dots, K_u - 1\}$ and \mathcal{X}_a denote the total and active IoT devices, respectively. As illustrated in Fig. 2.2, in each IoT device, the payload bits \mathbf{d}_k , $k \in \mathcal{X}_a$, are encoded by the channel encoder to increase the reliability of packet transmission. Then, the encoded data is passed through the differential encoding block. Differential encoding is employed to remove the need of channel estimation in the MUD at the IoT gateway. After differential encoding, the data is multiplied by a unique spreading waveform. It is considered that the spreading waveforms of the IoT devices do not change over time. Finally, the DS-SS signal is BPSK modulated and then transmitted.

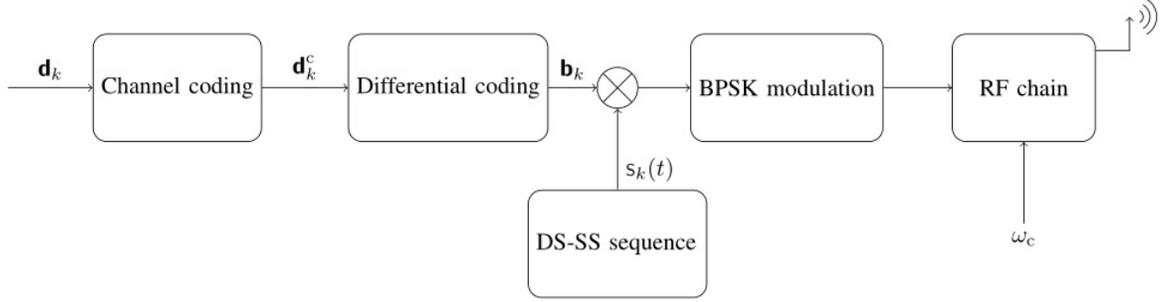


Fig. 2.2: Block diagram that illustrates packet transmission at IoT devices.

Let us consider that the maximum delay of the single-hop IoT network is τ_{\max} , i.e., $\tau_k \in [0, \tau_{\max}]$, $k \in \mathcal{X}_u$, where τ_k is delay of the k th IoT device. We consider that the IoT devices transmit their packet after receiving a beacon signal transmitted by the IoT gateway. This signal is periodically transmitted with period $T_t \geq N_s T_s + \tau_{\max}$, where N_s is the number of symbols per IoT packet, and T_s is the symbol duration. Fig. 2.3 illustrates the received IoT packets at the gateway. The received continuous-time baseband signal over frequency-flat fading channel in each period with respect to the timing reference of the gateway is modeled as

$$\begin{aligned}
 r(t) &= \sum_{k=0}^{K_u-1} \sum_{n=0}^{N_s-1} \check{g}_k \sqrt{p_k} e^{j\phi_k} \mathbf{b}_{k,n} s_k(t - nT_s - \tau_k) + w(t) \\
 &= \sum_{k=0}^{K_u-1} \sum_{n=0}^{N_s-1} \mathbf{g}_k \mathbf{b}_{k,n} s_k(t - nT_s - \tau_k) + w(t) \quad [0, T_t],
 \end{aligned} \tag{2.1}$$

where \check{g}_k , ϕ_k , and $\{\mathbf{b}_{k,n}, n = 0, 1, \dots, N_s - 1\}$ respectively denote the FCC, CP, and symbol stream of the k th IoT device, which are unknown at the gateway. It is considered that the envelope of the FCC, i.e., $|\check{g}_k|$ has a Rayleigh distribution. Without loss of generality, it is assumed that $\check{g}_k \sim \mathcal{N}(0, 1)$. The transmit power of the k th IoT devices is denoted by p_k which is known at the gateway, and the random variable $\mathbf{g}_k \triangleq \check{g}_k \sqrt{p_k} e^{j\phi_k}$. The symbol stream for the inactive IoT devices is modeled as transmitting zeros during the packet, i.e., $\mathbf{b}_{k,n} = 0$, $i = 0, 1, \dots, N_s - 1$ while active

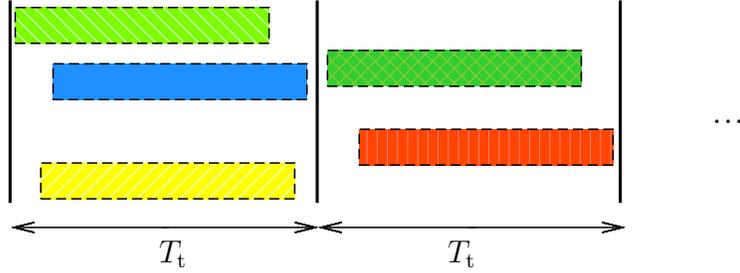


Fig. 2.3: Received packets at the gateway. Active IoT devices are shown with different colors.

IoT devices employ BPSK modulation. The DS-SS signaling waveform of the k th IoT device, $s_k(t)$, is given by

$$s_k(t) = \sum_{m=0}^{N_c-1} c_k^{(m)} \psi(t - mT_c) \quad t \in [0, T_s], \quad (2.2)$$

where T_c is the chip duration, $\mathbf{c}_k = [c_k^{(0)} \ c_k^{(1)} \ \dots \ c_k^{(N_c-1)}]^\dagger$ is the spreading sequence of $\{+1, -1\}$ assigned to the k th IoT device, and $\psi(t)$ is the chip waveform with unit power. It is assumed that $s_k(t)$, $k \in \mathcal{X}_u$, and $\psi(t)$ are rectangular pulses confined within $[0, T_s]$ and $[0, T_c]$, respectively. The baseband additive complex Gaussian noise at the output of the receive filter with bandwidth $1/T_c$ is denoted by $w(t)$.

Fig. 2.4 shows the block diagram of the proposed receiver at the IoT gateway. As seen, the received baseband signal is passed through chip matched filter (MF) and sampled at the chip rate. The output of the sampled chip MF for the i th chip at the j th observation symbol is obtained as

$$\begin{aligned} r_j^{(i)} &\triangleq \int_{jT_s+iT_c}^{jT_s+(i+1)T_c} r(t) \psi(t - jT_s - iT_c) dt \\ &= \sum_{k=0}^{K_u-1} g_k \mathbf{u}_{k,j}^{(i)} + w_j^{(i)} \quad i = 0, 1, \dots, N_c - 1, \end{aligned} \quad (2.3)$$

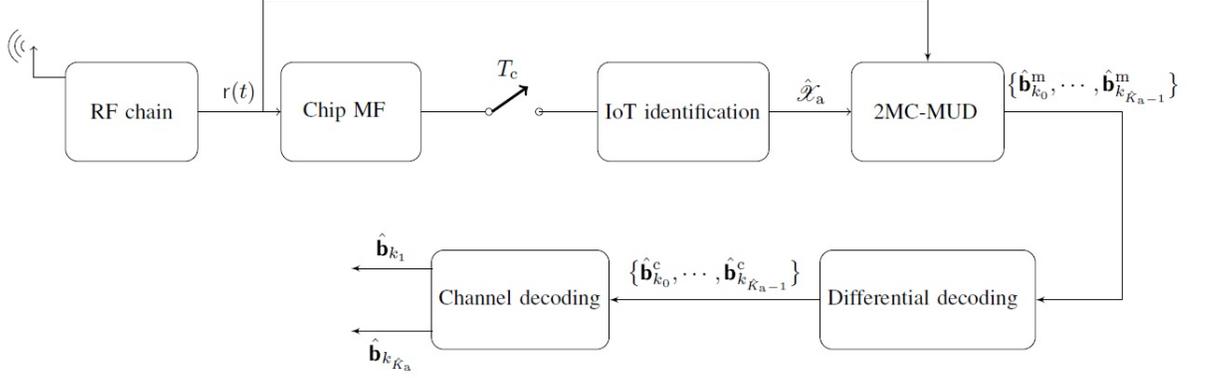


Fig. 2.4: Block diagram of the proposed receiver at the gateway.

where

$$w_j^{(i)} \triangleq \int_{jT_s+iT_c}^{jT_s+(i+1)T_c} w(t)\psi(t-jT_s-iT_c)dt, \quad (2.4)$$

and

$$u_{k,j}^{(i)} \triangleq \int_{jT_s+iT_c}^{jT_s+(i+1)T_c} \sum_{n=0}^{N_s-1} b_{k,n} s_k(t-nT_s-\tau_k)\psi(t-jT_s-iT_c)dt. \quad (2.5)$$

By employing (2.4), one can easily show that the joint probability density function (PDF) of the corresponding noise vector associated with the j th observation vector, i.e., $\mathbf{w}_j \triangleq [w_j^{(0)} \ w_j^{(1)} \ \dots \ w_j^{(N_c-1)}]^\dagger$ is characterized by $\mathbf{w}_j \sim \mathcal{N}_c(\mathbf{0}_{N_c}, \sigma_w^2 \mathbf{I})$ with $\sigma_w^2 \triangleq N_0/T_c$, where $N_0/2$ is the noise power spectral density of the white noise. The integral in (2.5) represents the area under the received signal waveform of the k th IoT device during the i th chip-matched filtering duration at the j th observation symbol. Let us define the delay of the k th IoT as

$$\tau_k \triangleq \alpha_k T_s + \beta_k T_c + \xi_k, \quad (2.6)$$

with $\alpha_k \triangleq \lfloor \tau_k/T_s \rfloor$, $\beta_k \triangleq \lfloor \tau_k/T_c \rfloor - \alpha_k N_c$, and $\xi_k \in [0, T_c)$. Based on the values of α_k ,

β_k , and ξ_k , $u_{k,j}^{(i)}$ in (2.5) is expressed as a function of $b_{k,j-\alpha_k}$, and $b_{k,j-\alpha_k-1}$ as [71–73]

$$\begin{aligned} u_{k,j}^{(i)} &\triangleq \sum_{n=0}^{N_s-1} \sum_{m=0}^{N_c-1} g_k c_k^{(m)} b_{k,n} \\ &\int_{jT_s+iT_c}^{jT_s+(i+1)T_c} \psi(t - nT_s - mT_c - \tau_k) \psi(t - jT_s - iT_c) dt \\ &= g_k b_{k,j-\alpha_k-1} x_k^{(i)} (1 - \xi_k) + g_k b_{k,j-\alpha_k} x_k^{(i)} (\xi_k), \end{aligned} \quad (2.7)$$

where

$$x_k^{(i)}(\nu) \triangleq \sum_{m=0}^{N_c-1} c_k^{(m)} \int_{iT_c}^{(i+1)T_c} \psi(t - mT_c - \nu T_c) \psi(t - iT_c) dt, \quad (2.8)$$

where $\nu \in [0, T_c)$. Equation (2.7) can be written in vector form as

$$\mathbf{u}_{k,j} = g_k b_{k,j-\alpha_k-1} \mathbf{x}_{k,0} + g_k b_{k,j-\alpha_k} \mathbf{x}_{k,1} \quad (2.9)$$

where $b_{k,j} = 0$, $j \notin [0, N_s - 1]$, and

$$\mathbf{u}_{k,j} \triangleq \left[u_{k,j}^{(0)} \quad u_{k,j}^{(1)} \quad \dots \quad u_{k,j}^{(N_c-1)} \right]^\dagger \quad (2.10a)$$

$$\mathbf{x}_{k,1} \triangleq \left[x_k^{(0)}(\xi_k) \quad x_k^{(1)}(\xi_k) \quad \dots \quad x_k^{(N_c-1)}(\xi_k) \right]^\dagger \quad (2.10b)$$

$$\mathbf{x}_{k,0} \triangleq \left[x_k^{(0)}(1 - \xi_k) \quad x_k^{(1)}(1 - \xi_k) \quad \dots \quad x_k^{(N_c-1)}(1 - \xi_k) \right]^\dagger. \quad (2.10c)$$

For the rectangular chip waveform pulse-shaping $\psi(t)$, one can easily obtain

$$\begin{bmatrix} \mathbf{x}_{k,1} \\ \mathbf{x}_{k,0} \end{bmatrix} = (1 - \xi_k) \begin{bmatrix} \mathbf{0}_{\beta_k} \\ \mathbf{c}_k \\ \mathbf{0}_{N_c-\beta_k} \end{bmatrix} + \xi_k \begin{bmatrix} \mathbf{0}_{\beta_k+1} \\ \mathbf{c}_k \\ \mathbf{0}_{N_c-\beta_k-1} \end{bmatrix}. \quad (2.11)$$

Let us define $\mathbf{X}_k \triangleq [\mathbf{x}_{k,0} \quad \mathbf{x}_{k,1}]$. By employing (2.3) and (2.9), the j th observation

where

$$\bar{\mathbf{R}} \triangleq \begin{bmatrix} \mathbf{r}_0 & \mathbf{r}_1 & \dots & \mathbf{r}_{N_t-1} \end{bmatrix} \quad (2.19a)$$

$$\bar{\mathbf{B}} \triangleq \begin{bmatrix} \mathbf{b}_0 & \mathbf{b}_1 & \dots & \mathbf{b}_{N_t-1} \end{bmatrix} \quad (2.19b)$$

$$\bar{\mathbf{W}} \triangleq \begin{bmatrix} \mathbf{w}_0 & \mathbf{w}_1 & \dots & \mathbf{w}_{N_t-1} \end{bmatrix} \quad (2.19c)$$

$$\bar{\mathbf{H}} \triangleq \begin{bmatrix} \mathbf{h}_0 & \mathbf{h}_1 & \dots & \mathbf{h}_{N_t-1} \end{bmatrix}. \quad (2.19d)$$

In (2.18), \mathbf{X} is referred to as dictionary.

As seen in Fig. 2.4, after chip-matched filtering and sampling, the IoT identification algorithm is applied to the measurement matrix $\bar{\mathbf{R}}$ to detect active IoT devices. The outcome of the IoT identification algorithm is a set of active IoT devices $\hat{\mathcal{X}}_a$. Then, the MUD algorithm is employed to detect the transmitted symbols of the IoT devices in $\hat{\mathcal{X}}_a$. Finally, after MUD, the bit streams related to the active IoT devices pass through differential and channel decoders, respectively.

2.3 IoT Identification

Node identification is the first step in the MA schemes where nodes do not use control signaling in order to identify themselves to the gateway. In this case, the gateway needs to determine the packet transmission state (PTS) of the nodes and detect data only for the active nodes. In this section, different IoT identification algorithms are developed.

2.3.1 IoT Identification Formulation

Let us write the observation model in (2.18) for an observation window with length l as

$$\mathbf{R} = \mathbf{X}\mathbf{G}\mathbf{B} + \mathbf{W} = \mathbf{X}\mathbf{H} + \mathbf{W}, \quad (2.20)$$

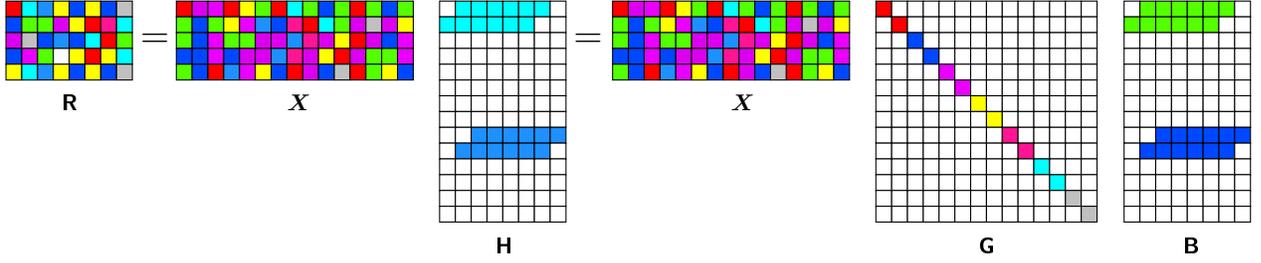


Fig. 2.5: Underdetermined systems of linear equations for with $K_u = 7$, $K_a = 2$, $N_t = 8$, and $N_s = 6$.

where

$$\mathbf{R} \triangleq \begin{bmatrix} \mathbf{r}_{\bar{\alpha}} & \mathbf{r}_{\bar{\alpha}+1} & \dots & \mathbf{r}_{\bar{\alpha}+l-1} \end{bmatrix} \quad (2.21a)$$

$$\mathbf{B} \triangleq \begin{bmatrix} \mathbf{b}_{\bar{\alpha}} & \mathbf{b}_{\bar{\alpha}+1} & \dots & \mathbf{b}_{\bar{\alpha}+l-1} \end{bmatrix} \quad (2.21b)$$

$$\mathbf{W} \triangleq \begin{bmatrix} \mathbf{w}_{\bar{\alpha}} & \mathbf{w}_{\bar{\alpha}+1} & \dots & \mathbf{w}_{\bar{\alpha}+l-1} \end{bmatrix} \quad (2.21c)$$

$$\mathbf{H} \triangleq \begin{bmatrix} \mathbf{h}_{\bar{\alpha}} & \mathbf{h}_{\bar{\alpha}+1} & \dots & \mathbf{h}_{\bar{\alpha}+l-1} \end{bmatrix}, \quad (2.21d)$$

where $1 \leq l \leq N_s + \alpha_{\min} - \alpha_{\max} - 1$, $\bar{\alpha}$ is an arbitrary positive integer as $\bar{\alpha} > \alpha_{\max} \triangleq \max\{\alpha_0, \alpha_1, \dots, \alpha_{K_u-1}\}$, and $\alpha_{\min} \triangleq \min\{\alpha_0, \alpha_1, \dots, \alpha_{K_u-1}\}$. Fig. 2.5 shows the underdetermined system of linear equations in (2.18), and Fig. 2.6 illustrates different observation windows for IoT identification in (2.20).

The activity of an IoT device is defined for an entire packet, i.e, the rows of \mathbf{H} corresponding to the active and inactive IoT devices are non-zero and zero, respectively. Thus, the problem of IoT identification for the k th IoT device, $k \in \mathcal{X}_u$, can be expressed as the following binary hypothesis testing problem

$$H_{1k} : \mathbf{h}_{k,\bar{\alpha},l} \neq \mathbf{0} \quad (2.22)$$

$$H_{0k} : \mathbf{h}_{k,\bar{\alpha},l} = \mathbf{0},$$

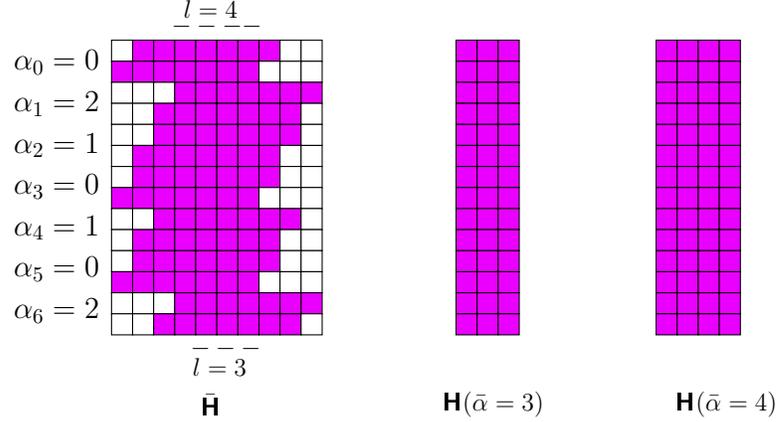


Fig. 2.6: Different observation windows for IoT identification ($K_u = 7$, $N_s = 7$, $\alpha_{\max} = 2$, $\alpha_{\min} = 0$, $1 \leq l \leq 4$). Purple color is employed to show the packet of IoT devices which is zero for inactive and non-zero for active IoT devices.

where

$$\mathbf{h}_{k,\bar{\alpha},l} \triangleq \left[\mathbf{h}_{k,\bar{\alpha}}^\dagger \quad \mathbf{h}_{k,\bar{\alpha}+1}^\dagger \quad \cdots \quad \mathbf{h}_{k,\bar{\alpha}+l-1}^\dagger \right]^\dagger, \quad (2.23a)$$

$$\mathbf{h}_{k,j} \triangleq \left[h_{k,j,0} \quad h_{k,j,1} \right]^\dagger, \quad (2.23b)$$

and H_{0k} and H_{1k} are the null and alternative hypothesis denoting that the k th IoT device is active and inactive, respectively.

As seen in (2.22), the IoT identification problem is formulated as K_u parallel binary hypothesis testing problems.

The first step in IoT identification is to reconstruct $\mathbf{h}_{k,\bar{\alpha},l}$, $k \in \mathcal{X}_u$, from the observation matrix in (2.20). However, (2.20) represents an underdetermined system of linear equations since $N_c < K_u$. Hence, it is not uniquely solvable.

Let us denote the number of active IoT devices by the random variable k_a . The

distribution of k_a is binomial, and

$$\mathbb{P}\{k_a = K_a\} = \binom{K_u}{K_a} P_a^{K_a} (1 - P_a)^{(K_u - K_a)}. \quad (2.24)$$

For $P_a \ll 1$, $\mathbb{P}\{k_a \ll K_u\} = 1$, and thus, \mathbf{B} and \mathbf{H} in (2.20) are sparse matrices. Moreover, the columns of $\mathbf{H}(\mathbf{B})$ share the same sparsity profiles. This sparse structure is referred to as block-sparse. The block-sparse structure of \mathbf{H} can be observed in Fig. 2.6. Fig. 2.7 shows the probability mass function of the number of active IoT devices for low values of P_a . As seen, this is condensed between a lower bound K_{al} and an upper bound K_{au} which are defined as $\mathbb{P}\{k_a > K_{al}\} < \epsilon$ and $\mathbb{P}\{k_a < K_{au}\} < \epsilon$, where ϵ is an arbitrary small value.

The sparse structure of \mathbf{H} can be employed to reconstruct columns of \mathbf{H} from the underdetermined linear observation model in (2.20). When each column of \mathbf{H} is individually reconstructed from its corresponding column in \mathbf{R} , it is referred to as SSR. The SSR of the columns of \mathbf{H} , i.e., \mathbf{h}_j , $\bar{\alpha} \leq j \leq \bar{\alpha} + l - 1$, is formulated as [74, 75]

$$\begin{aligned} & \underset{\mathbf{h}_j}{\text{minimize}} \quad \|\mathbf{r}_j - \mathbf{X}\mathbf{h}_j\|_F^2 \\ & \text{subject to} \quad \|\mathbf{h}_j\|_0 \leq 2K_{au}, \end{aligned} \quad (2.25)$$

where $\|\cdot\|_F$ and $\|\mathbf{h}_j\|_0$ are the Frobenius and ℓ_0 quasi-norm, respectively. From duality and the Karush Kuhn Tucker (KKT) optimality conditions [76, 77], (2.25) can be written as

$$\hat{\mathbf{h}}_j = \arg \min_{\mathbf{h}_j} \frac{1}{2} \|\mathbf{r}_j - \mathbf{X}\mathbf{h}_j\|_F^2 + \lambda_{\ell_0} \|\mathbf{h}_j\|_0, \quad (2.26)$$

where λ_{ℓ_0} is the tuning parameter which balances both approximation error and sparsity level of the solution. In Appendix 2.A, the value of λ_{ℓ_0} based on maximum a

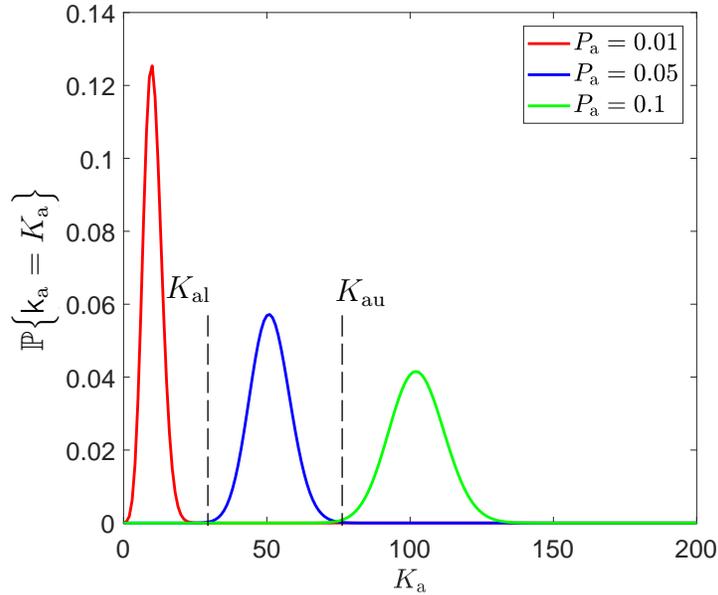


Fig. 2.7: The probability mass function of k_a for different probability of packet transmission, P_a , when $K_u = 1024$. The dashed lines show K_{al} and K_{au} for $P_a = 0.05$.

posteriori probability (MAP) criterion is obtained as

$$\lambda_{\ell_0} = \frac{\sigma_w^2}{2} \ln \left(\frac{4(1 - P_a)}{P_a} \right). \quad (2.27)$$

The SSR is based on the principle that, through optimization, the sparsity of a signal can be exploited to reconstruct it from far fewer samples than required by the Shannon-Nyquist sampling theorem.

The ℓ_0 -minimization in (2.26) is both numerically unstable and NP-hard since the ℓ_0 quasi-norm is a discrete-value function. One approach to the SSR is to replace the ℓ_0 quasi-norm by a convex function with common sparsity profile that leads to a solution very close to the one of the original problem. This method is called convex relaxation and converts the combinatorial problem in (2.26) into a convex optimization problem which can be solved in polynomial time. Different convex functions can be employed to relax $\|\mathbf{h}_j\|_0$ in (2.26). A common family of convex functions is the ℓ_q norm, given

as

$$\|\mathbf{h}_j\|_q = \left(\sum_{k=0}^{K_u-1} \sum_{f=0}^1 |h_{k,j,f}|^q \right)^{\frac{1}{q}}. \quad (2.28)$$

The recovered vectors by the ℓ_q norm minimization can be employed to infer the active IoT set \mathcal{X}_a through the parallel binary hypothesis testing problems in (2.22). We refer to the identification based on SSR as DSA IoT identification.

On the other hand, the block-sparse structure of \mathbf{H} can be employed to improve the reconstruction of \mathbf{H} in (2.25). This method of signal reconstruction is referred to as SSSR. Opposite to SSR, the SSSR simultaneously exploits the column sparsity along with the block-sparse structure in the optimization problem in order to reconstruct matrix \mathbf{H} . The SSSR of \mathbf{H} , given the received signal matrix \mathbf{R} and the dictionary \mathbf{X} is expressed as [78]

$$\begin{aligned} & \underset{\mathbf{H}}{\text{minimize}} \quad \|\mathbf{R} - \mathbf{X}\mathbf{H}\|_{\text{F}}^2 \\ & \text{subject to} \quad \|\bar{\mathbf{H}}\|_0 \leq 2K_{\text{au}}. \end{aligned} \quad (2.29)$$

From duality and the KKT optimality conditions, (2.29) can be rewritten as

$$\hat{\mathbf{H}} = \arg \min_{\mathbf{H}} \frac{1}{2} \|\mathbf{R} - \mathbf{X}\mathbf{H}\|_{\text{F}}^2 + \lambda_{\ell_0}^{\ell_0} \|\mathbf{H}\|_0, \quad (2.30)$$

where $\lambda_{\ell_0}^{\ell_0} \geq 0$ is the tuning parameter, and $\|\mathbf{H}\|_0$ is the $\ell_0 - \ell_0$ quasi-norm of \mathbf{H} . In Appendix 2.B, the value of $\lambda_{\ell_0}^{\ell_0}$ based on MAP criterion is obtained as

$$\lambda_{\ell_0}^{\ell_0} = \frac{\sigma_w^2}{2l} \ln \left(\frac{2^{l+1}(1 - P_a)}{P_a} \right). \quad (2.31)$$

Similar to the ℓ_0 -minimization in (2.25), the $\ell_0 - \ell_0$ -minimization in (2.29) is unstable and NP-hard. Therefore, the quasi-norm $\|\mathbf{H}\|_0$ is replaced with the $\ell_p - \ell_q$

($p, q \geq 1$) mixed-norm as

$$J_{p,q}(\mathbf{H}) = \sum_{k=0}^{K_u-1} \left\| \mathbf{h}_{k,\bar{\alpha},l} \right\|_q^p \quad (2.32)$$

to convert the combinatorial problem in (2.30) into a convex optimization problem.

Similar to the DSA IoT identification, we refer to the identification based on SSSR, as to PDSA IoT identification. With the PDSA approach, the recovered matrix is employed to infer the active IoT set.

2.3.2 Squared ℓ_2 norm DSA IoT Identification

The proposed squared ℓ_2 norm DSA IoT identification algorithm replaces the ℓ_0 quasi-norm in (2.26) with the squared ℓ_2 norm as

$$\hat{\mathbf{h}}_j = \arg \min_{\mathbf{h}_j} \frac{1}{2} \left\| \mathbf{r}_j - \mathbf{X} \mathbf{h}_j \right\|_2^2 + \lambda_{\ell_2} \left\| \mathbf{h}_j \right\|_2^2, \quad (2.33)$$

The squared ℓ_2 norm convex relaxation formulates the IoT identification problem as a ridge regression (RD) estimation problem as in (2.33) followed by K_u parallel binary hypothesis testing problems.

The optimal solution of (2.33) is obtained as [79]

$$\hat{\mathbf{h}}_j = \left(\mathbf{X}^\dagger \mathbf{X} + 2\lambda_{\ell_2} \mathbf{I} \right)^{-1} \mathbf{X}^\dagger \mathbf{r}_j, \quad (2.34)$$

which is a simple linear estimator of \mathbf{r}_j that shrinks ordinary least-squares (LS) estimates towards zero.

As mentioned above, λ_{ℓ_2} in (2.33) balances both approximation error and sparsity level of the solution. The optimal value of tuning parameter, λ_{ℓ_2} , can be obtained through cross validation and generalized cross validation [80–84]. The latter is a

method of model selection that is intuitively simple and widely employed; in this case, the optimal value of λ_{ℓ_2} is obtained as [82]

$$\lambda_{\ell_2}^{\text{op}} = \arg \min_{\lambda_{\ell_2}} \frac{2 \|\mathbf{r}_j - \mathbf{Q}\mathbf{r}_j\|_2^2}{[\text{tr}(\mathbf{I} - \mathbf{Q})]^2}, \quad (2.35)$$

where $\mathbf{Q} \triangleq \mathbf{X} \left(\mathbf{X}^\dagger \mathbf{X} + 2\lambda_j \mathbf{I} \right)^{-1} \mathbf{X}^\dagger$. As seen, $\lambda_{\ell_2}^{\text{op}}$ is obtained at the expense of high computational complexity to solve the minimization problem in (2.35). However, according to Appendix 2.C, a conservative choice of the tuning parameter λ_{ℓ_2} in terms of minimum mean square error (MMSE) can be approximated as

$$\lambda_{\ell_2}^{\text{op}} \approx \frac{\sigma_w^2 \text{tr}[\bar{\Sigma}_{\mathbf{X}}^{-1}]}{P_a (\mathbf{p}^\dagger \otimes \mathbf{1}^\dagger) \bar{\Lambda}_{\mathbf{X}} + 3 \text{tr}[\bar{\Sigma}_{\mathbf{X}}^{-2}]}, \quad (2.36)$$

where $\mathbf{p} \triangleq [p_0 \ p_1 \ \cdots \ p_{K_u-1}]^\dagger$, $\bar{\Sigma}_{\mathbf{X}} \triangleq \mathbf{X}^\dagger \mathbf{X}$, and $\bar{\Lambda}_{\mathbf{X}} \triangleq \text{diag}(\bar{\Sigma}_{\mathbf{X}}^{-1})$. As seen in (2.36), $\lambda_{\ell_2}^{\text{op}}$ is inversely proportional to P_a .

By substituting $\mathbf{r}_j = \mathbf{X}\mathbf{h}_j + \mathbf{w}_j$ in (2.12) into (2.34), $\hat{\mathbf{h}}_j$ can be written as a linear function of \mathbf{h}_j as

$$\hat{\mathbf{h}}_j = \mathbf{\Omega} \mathbf{h}_j + \mathbf{w}'_j, \quad (2.37)$$

where

$$\mathbf{\Omega} \triangleq \begin{bmatrix} \Omega_{0,0} & \Omega_{0,1} & \cdots & \Omega_{0,2K_u-1} \\ \Omega_{1,0} & \Omega_{1,1} & \cdots & \Omega_{1,2K_u-1} \\ \vdots & \vdots & \ddots & \vdots \\ \Omega_{2K_u-1,0} & \Omega_{2K_u-1,1} & \cdots & \Omega_{2K_u-1,2K_u-1} \end{bmatrix} = \mathbf{I} - 2\lambda_{\ell_2} (\bar{\Sigma}_{\mathbf{X}} + 2\lambda_{\ell_2} \mathbf{I})^{-1}, \quad (2.38)$$

and

$$\mathbf{w}'_j \triangleq \begin{bmatrix} w'_{0,j,0} \\ w'_{0,j,1} \\ \vdots \\ w'_{K_u-1,j,0} \\ w'_{K_u-1,j,1} \end{bmatrix} = \left(\bar{\Sigma}_{\mathbf{X}} + 2\lambda_{\ell_2} \mathbf{I} \right)^{-1} \mathbf{X}^\dagger \mathbf{w}_j. \quad (2.39)$$

In (2.39), \mathbf{w}'_j is zero-mean complex Gaussian colored noise vector with covariance matrix given by

$$\Sigma^{\mathbf{w}'_j} \triangleq \begin{bmatrix} \Sigma_{0,0}^{\mathbf{w}'} & \Sigma_{0,1}^{\mathbf{w}'} & \cdots & \Sigma_{0,2K_u-1}^{\mathbf{w}'} \\ \Sigma_{1,0}^{\mathbf{w}'} & \Sigma_{1,1}^{\mathbf{w}'} & \cdots & \Sigma_{1,2K_u-1}^{\mathbf{w}'} \\ \vdots & \vdots & \ddots & \vdots \\ \Sigma_{2K_u-1,0}^{\mathbf{w}'} & \Sigma_{2K_u-1,1}^{\mathbf{w}'} & \cdots & \Sigma_{2K_u-1,2K_u-1}^{\mathbf{w}'} \end{bmatrix} = \mathbb{E} \left\{ \mathbf{w}'_j (\mathbf{w}'_j)^{\text{H}} \right\} = \sigma_w^2 \left(\bar{\Sigma}_{\mathbf{X}} + 2\lambda_{\ell_2} \mathbf{I} \right)^{-2} \bar{\Sigma}_{\mathbf{X}}, \quad (2.40)$$

where $\mathbb{E} \left\{ w'_{k_1,j_2,f_1} (w'_{k_2,j_2,f_2})^* \right\} = \Sigma_{2k_1+f_1, 2k_2+f_2}^{\mathbf{w}'_j}$.

The elements of $\hat{\mathbf{h}}_j$ in (2.37) associated with the k th IoT device, i.e., $\hat{h}_{k,j,0}$ and $\hat{h}_{k,j,1}$ can be written in a summation form as

$$\begin{aligned} \hat{h}_{k,j,f} &= \Omega_{2k+f, 2k+f} h_{k,j,f} + \Omega_{2k+f, 2k+\bar{f}} h_{k,j,\bar{f}} \\ &+ \sum_{n \neq k} \left\{ \Omega_{2k+f, 2n+f} h_{n,j,f} + \Omega_{2k+f, 2n+\bar{f}} h_{n,j,\bar{f}} \right\} + w'_{k,j,f}, \end{aligned} \quad (2.41)$$

where $f, \bar{f} \in \{0, 1\}$ and $\bar{f} \triangleq f + (-1)^f$. The second term on the right-hand side of (2.41) represents the effect of interference caused by the existing active IoT devices

in the network. Let us define

$$\hat{\mathbf{h}}_{k,\bar{\alpha},l} \triangleq \left[\hat{\mathbf{h}}_{k,\bar{\alpha}}^\dagger \quad \hat{\mathbf{h}}_{k,\bar{\alpha}+1}^\dagger \quad \cdots \quad \hat{\mathbf{h}}_{k,\bar{\alpha}+l-1}^\dagger \right]^\dagger, \quad (2.42a)$$

$$\hat{\mathbf{h}}_{k,j} \triangleq \left[\hat{h}_{k,j,0} \quad \hat{h}_{k,j,1} \right]^\dagger. \quad (2.42b)$$

In order to identify the transmission state of the k th IoT device based on the reconstructed signal $\hat{\mathbf{h}}_{k,\bar{\alpha},l}$ in (2.42), the joint PDF of the random vector $\hat{\mathbf{h}}_{k,\bar{\alpha},l}$ is needed. In Lemma. (2.3.1), the joint PDF of $\hat{\mathbf{h}}_{k,\bar{\alpha},l}$ is derived.

Lemma 2.3.1. *The distribution of the random vector $\hat{\mathbf{h}}_{k,\bar{\alpha},l}$ in (2.42) under hypothesis H_{1k} and H_{0k} can be approximated by joint complex Gaussian distribution as*

$$p\left(\hat{\mathbf{h}}_{k,\bar{\alpha},l} | H_{tk}\right) \sim \mathcal{N}_c\left(\mathbf{0}, \Sigma^{tk}\right) \quad t \in \{0, 1\}, \quad (2.43)$$

where

$$\Sigma^{tk} \triangleq \begin{bmatrix} \Sigma_{0,0}^{tk} & \Sigma_{0,1}^{tk} & \cdots & \Sigma_{0,2l-1}^{tk} \\ \Sigma_{1,0}^{tk} & \Sigma_{1,1}^{tk} & \cdots & \Sigma_{1,2l-1}^{tk} \\ \vdots & \vdots & \ddots & \vdots \\ \Sigma_{2l-1,0}^{tk} & \Sigma_{2l-1,1}^{tk} & \cdots & \Sigma_{2l-1,2l-1}^{tk} \end{bmatrix}, \quad (2.44)$$

and

$$\begin{aligned} \Sigma_{2j'+f,2j'+f}^{tk} \triangleq \text{Var}\left\{\hat{h}_{k,j,f} | H_{tk}\right\} &= tp_k \left(\Omega_{2k+f,2k+f}^2 + \Omega_{2k+f,2k+\bar{f}}^2 \right) + \Sigma_{2k+f,2k+f}^{w'} \\ &+ P_a \sum_{n \neq k} p_n \left(\Omega_{2k+f,2n+f}^2 + \Omega_{2k+f,2n+\bar{f}}^2 \right), \end{aligned} \quad (2.45)$$

with

$$j' = j - \bar{\alpha} \quad (2.46)$$

$$\bar{\alpha} \leq j \leq \bar{\alpha} + l - 1,$$

and the off-diagonal elements of the covariance matrixes Σ^{tk} , i.e., $\Sigma_{2j'_1+f_1, 2j'_2+f_2}^{tk} \triangleq \text{Cov}\left\{\hat{h}_{k,j_1,f_1}, \hat{h}_{k,j_2,f_2} \middle| H_{tk}\right\}$, $2j'_1 + f_1 \neq 2j'_2 + f_2$, are given in (2.47).

$$\begin{aligned} \Sigma_{2j'_1+f_1, 2j'_2+f_2}^{tk} &= \left(\Sigma_{2j'_2+f_2, 2j'_1+f_1}^{tk}\right)^* = \text{Cov}\left\{\hat{h}_{k,j_1,f_1}, \hat{h}_{k,j_2,f_2} \middle| H_{tk}\right\} \quad 2j'_1 + f_1 < 2j'_2 + f_2 \\ &= \delta_{j_1, j_2} \delta_{f_1, 0} \delta_{f_2, 1} \left[tp_k \sigma_{g_k}^2 \left(\Omega_{2k, 2k} \Omega_{2k+1, 2k} + \Omega_{2k+1, 2k+1} \Omega_{2k, 2k+1} \right) \right. \\ &\quad \left. + P_a \sum_{n \neq k} p_n \left(\Omega_{2k, 2n+1} \Omega_{2k+1, 2n} + \Omega_{2k+1, 2n+1} \Omega_{2k, 2n+1} \right) + \Sigma_{2k, 2k+1}^{w'} \right] \\ &\quad + \delta_{j_2 - j_1, 1} \delta_{f_1, 0} \delta_{f_2, 0} \left[tp_k \left(\Omega_{2k, 2k} \Omega_{2k, 2k+1} \right) + P_a \sum_{n \neq k} p_n \left(\Omega_{2k, 2n} \Omega_{2k, 2n+1} \right) \right] \\ &\quad + \delta_{j_2 - j_1, 1} \delta_{f_1, 1} \delta_{f_2, 1} \left[tp_k \left(\Omega_{2k+1, 2k+1} \Omega_{2k+1, 2k} \right) + P_a \sum_{n \neq k} p_n \left(\Omega_{2k+1, 2n+1} \Omega_{2k+1, 2n} \right) \right] \\ &\quad + \delta_{j_2 - j_1, 1} \delta_{f_1, 0} \delta_{f_2, 1} \left[tp_k \left(\Omega_{2k, 2k+1} \Omega_{2k+1, 2k} \right) + P_a \sum_{n \neq k} p_n \left(\Omega_{2k, 2n+1} \Omega_{2k+1, 2n} \right) \right] \\ &\quad + \delta_{j_2 - j_1, 1} \delta_{f_1, 1} \delta_{f_2, 0} \left[tp_k \left(\Omega_{2k, 2k} \Omega_{2k+1, 2k+1} \right) + P_a \sum_{n \neq k} p_n \left(\Omega_{2k, 2n} \Omega_{2k+1, 2n+1} \right) \right] \end{aligned} \quad (2.47)$$

Proof in Appendix 2.D.

Let us define

$$d_k = \begin{cases} H_{1k}, & \text{if } \phi(\hat{\mathbf{h}}_{k, \bar{\alpha}, l}) \geq \theta_k \\ H_{0k}, & \text{if } \phi(\hat{\mathbf{h}}_{k, \bar{\alpha}, l}) < \theta_k \end{cases}, \quad (2.48)$$

where $\phi(\cdot, \cdot)$ is an arbitrary continuous function and θ_k is a threshold for the k th IoT device.

Algorithm 1 Bayesian Squared ℓ_2 norm DSA IoT Identification Algorithm

Input: \mathbf{X} , \mathbf{R} , λ_{ℓ_2} , Σ^{0k} , Σ^{1k} , and θ_k for $k \in \mathcal{X}_u$
Output: Active IoT set $\hat{\mathcal{X}}_a$
Initialization: $\hat{\mathcal{X}}_a = \emptyset$

- 1: **for** $k = 0, 1, \dots, K_u - 1$ **do**
- 2: Obtain $\hat{\mathbf{h}}_{k,\bar{\alpha},l}$ in (2.42) by employing (2.34)
- 3: Compute $\phi(\hat{\mathbf{h}}_{k,\bar{\alpha},l})$ in (2.49)
- 4: Identify the transmission state of the k th IoT device through (2.48)
- 5: **if** $d_k = H_{1k}$ **then**
- 6: $\hat{\mathcal{X}}_a \leftarrow \{\hat{\mathcal{X}}_a, k\}$
- 7: **end if**
- 8: **end for**

2.3.2.1 Bayesian Squared ℓ_2 norm DSA IoT Identification Algorithm

By applying the optimum Bayesian's decision rule to the reconstructed vector $\hat{\mathbf{h}}_{k,\bar{\alpha},l}$, $k \in \mathcal{X}_u$, in (2.42), the Bayesian squared ℓ_2 norm DSA IoT identification algorithm is derived as in (2.48) for

$$\phi(\hat{\mathbf{h}}_{k,\bar{\alpha},l}) = \hat{\mathbf{h}}_{k,\bar{\alpha},l}^H \left((\Sigma^{0k})^{-1} - (\Sigma^{1k})^{-1} \right) \hat{\mathbf{h}}_{k,\bar{\alpha},l}, \quad (2.49)$$

and

$$\theta_k = \ln \left(\frac{(1 - P_a) \det(\Sigma^{1k})}{P_a \det(\Sigma^{0k})} \right). \quad (2.50)$$

Proof in Appendix 2.E.

The optimum Bayesian's decision rule in (2.49) minimizes the Bayesian risk, i.e., $\mathbb{P}\{d_k = H_{1k}|H_{0k}\}(1 - P_a) + \mathbb{P}\{d_k = H_{0k}|H_{1k}\}P_a$. A formal description of the proposed Bayesian squared ℓ_2 norm DSA IoT identification algorithm is summarized in Algorithm 1.

2.3.2.2 ML Squared ℓ_2 norm DSA IoT Identification Algorithm

As seen in (2.47), each four successive elements of the reconstructed vector $\bar{\mathbf{h}}_{k,\bar{\alpha},l}$, i.e., $\hat{h}_{k,j,0}$, $\hat{h}_{k,j,1}$, $\hat{h}_{k,j+1,0}$, and $\hat{h}_{k,j+1,1}$, $\bar{\alpha} \leq j \leq \bar{\alpha} + l - 2$, are correlated random variables. Hence, they provide correlated information about the transmission state of the k th IoT device. Through downsampling of $\bar{\mathbf{h}}_{k,\bar{\alpha},2l}$, an independent and identically distributed (i.i.d.) random vector can be obtain which can be used to develop a low-complexity IoT identification algorithm.

Let us consider the i.i.d. random vector $\check{\mathbf{h}}_{k,l}$, defined as

$$\check{\mathbf{h}}_{k,l} \triangleq \left[\check{h}_{k,0} \check{h}_{k,2} \cdots \check{h}_{k,2(l-1)} \right]^\dagger, \quad (2.51)$$

where

$$\check{h}_{k,j'} \triangleq \hat{h}_{k,j,1} \mathbb{I} \left\{ \beta_k < \frac{N_c}{2} \right\} + \hat{h}_{k,j,0} \mathbb{I} \left\{ \beta_k \geq \frac{N_c}{2} \right\}, \quad (2.52)$$

where β_k is the chip-delay of the k th IoT device given in (2.6) and $j' = j - \bar{\alpha}$, $j \in \{\bar{\alpha}, \bar{\alpha} + 2 \cdots \bar{\alpha} + 2(l-1)\}$. In order to identify the transmission state of the k th IoT device, $k \in \mathcal{X}_u$, through the reconstructed vector $\check{\mathbf{h}}_{k,l}$, the maximum likelihood ratio (MLR) test can be employed. The MLR test maximizes the correct identification rate of the k th IoT devices, i.e., $P_k^{(c)} \triangleq \mathbb{P}\{d_k = H_{1k} | H_{1k}\}$ subject to a constraint on the maximum allowable false alarm rate, i.e., $P_k^{(f)} \triangleq \mathbb{P}\{d_k = H_{1k} | H_{0k}\}$ [85–87].

By employing Lemma (2.3.1), the decision rule in (2.48) for the MLR test of the reconstructed vector $\check{\mathbf{h}}_{k,l}$, $k \in \mathcal{X}_u$, is obtained as (*Proof* in Appendix 2.F)

$$\phi\left(\check{\mathbf{h}}_{k,l}\right) = \frac{\sqrt{2}}{\sqrt{\sum_{2k+f, 2k+f}^{0k}}} \sum_{i=0}^{l-1} \left| \check{h}_{k,2i} \right|^2 \quad (2.53)$$

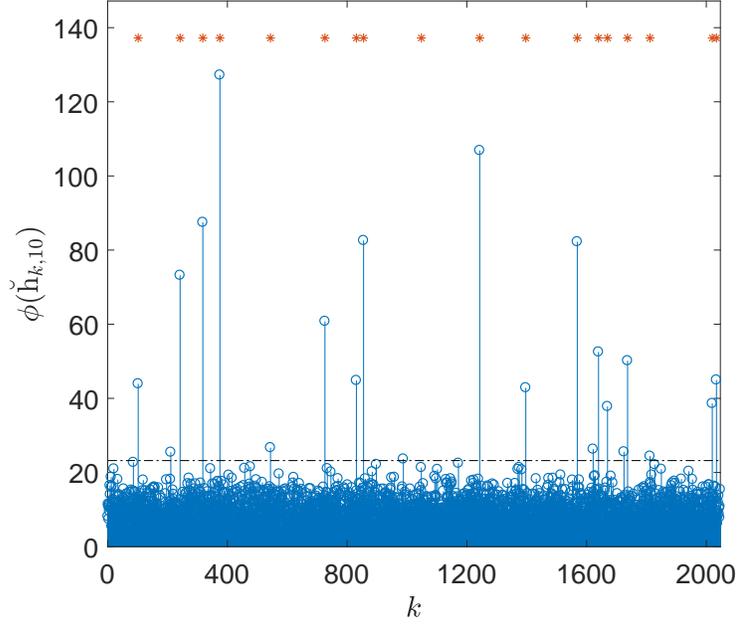


Fig. 2.8: The decision statistics of the ML squared ℓ_2 norm DSA IoT identification algorithm, i.e. $\phi(\check{\mathbf{h}}_{k,l})$ versus k , for $l = 10$, $K_u = 2048$, $P_a = 0.01$, $P_k^{(f)} = 0.01$, $k = 0, 1, \dots, K_u - 1$, and at 10 dB SNR.

for $\theta_k = \theta'_k$, where θ'_k is set based on a predefined false alarm rate as

$$\theta'_k = 2\Gamma^{-1}\left(l, (1 - P_k^{(f)})\Gamma(l)\right), \quad (2.54)$$

where $f = 1$ for $\beta_k < \frac{N_c}{2}$, and $f = 0$ for $\beta_k \geq \frac{N_c}{2}$, and $\Gamma(l) = \int_0^\infty \exp(-t)t^{(l-1)}dt$ and $\Gamma^{-1}(a, b)$ are complete gamma function and inverse lower incomplete gamma function, respectively (*Proof* in Appendix 2.G).

Fig. 2.8 shows the decision statistics of the squared ℓ_2 norm DSA IoT identification algorithm, i.e. $\phi(\check{\mathbf{h}}_{k,l})$ versus k , for $l = 10$ and at 10 dB SNR. As seen, $\phi(\check{\mathbf{h}}_{k,l})$ exceeds the threshold for the active IoT devices.

By using (2.54), one can obtain the correct identification rate for the k th IoT

Algorithm 2 ML Squared ℓ_2 norm DSA IoT Identification Algorithm

Input: \mathbf{X} , \mathbf{R} , $P_k^{(f)}$, $k \in \mathcal{X}_u$
Output: Active IoT set $\hat{\mathcal{X}}_a$
Initialization: $\hat{\mathcal{X}}_a = \emptyset$

- 1: **for** $k = 0, 1, \dots, K_u - 1$ **do**
- 2: Obtain θ'_k by employing (2.54)
- 3: Obtain $\mathbf{h}_{k,l}$ in (2.51) by employing (2.34) and (2.52)
- 4: Compute $\phi(\check{\mathbf{h}}_{k,l})$ in (2.53)
- 5: Identify the transmission state of the k th IoT device by employing (2.48) for $\phi(\check{\mathbf{h}}_{k,l})$ and θ'_k in (2.54)
- 6: **if** $d_k = H_{1k}$ **then**
- 7: $\hat{\mathcal{X}}_a \leftarrow \{\hat{\mathcal{X}}_a, k\}$
- 8: **end if**
- 9: **end for**

device as

$$P_k^{(c)} = \mathbb{P}\{d_k = H_{1k} | H_{1k}\} = 1 - \frac{\Gamma\left(l, \sqrt{\frac{\sum_{2k+f, 2k+f}^{0k} \theta'_k}{4\sum_{2k+f, 2k+f}^{1k}}}\right)}{\Gamma(l)}, \quad (2.55)$$

where $\Gamma(a, b) = \int_b^\infty \exp(-t)t^{(a-1)}dt$ is the lower incomplete gamma function, and θ'_k is given in (2.54) (*Proof* in Appendix 2.G).

A formal description of the proposed ML squared ℓ_2 norm DSA IoT identification algorithm is summarized in Algorithm 2. Also, the block diagram of the proposed squared ℓ_2 norm DSA IoT identification algorithm for both Bayesian and ML rule is shown in Fig. 2.9.

2.3.3 ℓ_1 norm DSA IoT Identification

The ℓ_1 norm DSA IoT identification algorithm formulates the SSR as

$$\hat{\mathbf{h}}_j = \arg \min_{\mathbf{h}_j} \frac{1}{2} \|\mathbf{r}_j - \mathbf{X}\mathbf{h}_j\|_F^2 + \lambda_{\ell_1} \|\mathbf{h}_j\|_1, \quad (2.56)$$

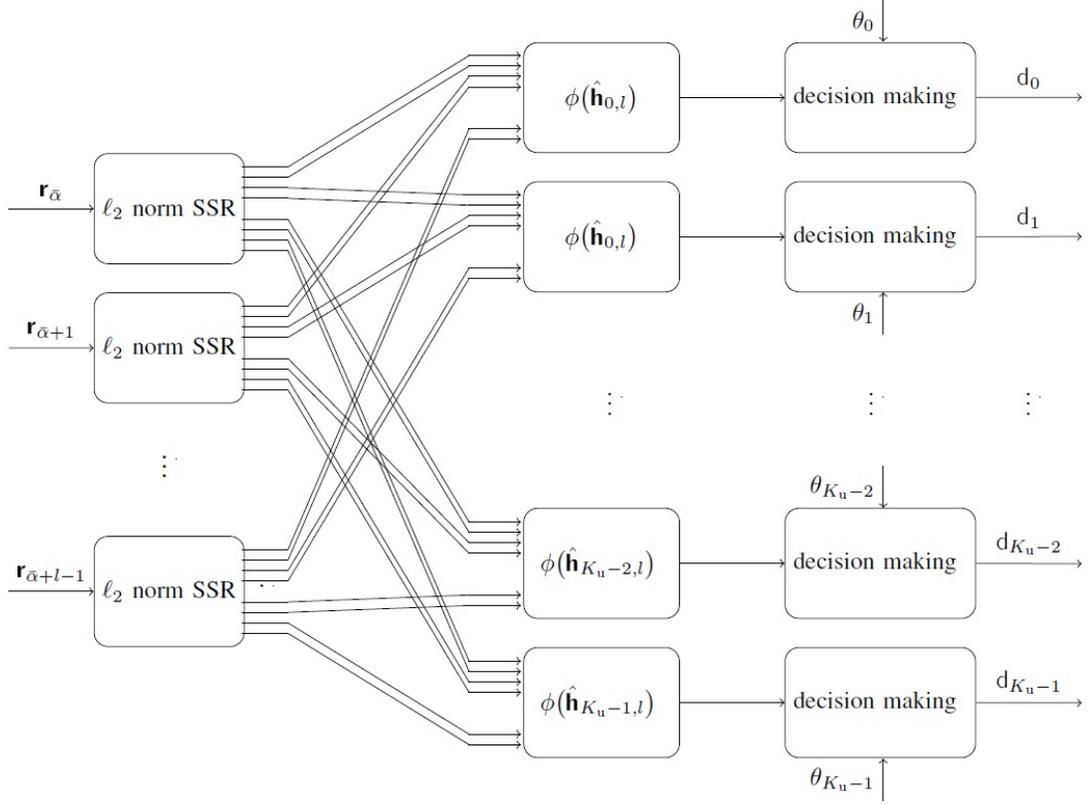


Fig. 2.9: Block diagram of the proposed squared ℓ_2 norm DSA IoT identification algorithms. For Algorithm 1 and Algorithm 2, $\phi(\hat{\mathbf{h}}_{k,\bar{\alpha},l})$ is given by (2.49) and (2.53), respectively.

where $j = \bar{\alpha}, \bar{\alpha} + 1, \dots, \bar{\alpha} + l - 1$ and λ_{ℓ_1} is the tuning parameter.

Let us define

$$\ddot{\mathbf{X}} \triangleq \mathbf{X} \mathbf{P}^{\frac{1}{2}} \quad (2.57a)$$

$$\ddot{\mathbf{h}}_j \triangleq \mathbf{P}^{-\frac{1}{2}} \mathbf{h}_j, \quad (2.57b)$$

$$\ddot{\mathbf{h}}_{k,j} \triangleq \frac{\mathbf{h}_{k,j}}{\sqrt{p_k}}, \quad (2.57c)$$

where

$$\mathbf{P} \triangleq \begin{bmatrix} p_0 & & & \\ & p_1 & & \\ & & \ddots & \\ & & & p_{K_u-1} \end{bmatrix} \otimes \mathbf{I}_2. \quad (2.58)$$

In Appendix 2.I, the SSR problem in (2.56) for IoT identification is reformulated as

$$\hat{\mathbf{h}}_j = \arg \min_{\mathbf{h}_j} \frac{1}{2} \|\mathbf{r}_j - \ddot{\mathbf{X}} \ddot{\mathbf{h}}_j\|_{\mathbb{F}}^2 + \ddot{\lambda}_{\ell_1} \|\ddot{\mathbf{h}}_j\|_1, \quad (2.59)$$

where

$$\ddot{\lambda}_{\ell_1} = \frac{2\lambda_{\ell_0}}{\sqrt{\pi}} = \frac{\sigma_w^2}{\sqrt{\pi}} \ln \left(\frac{4(1 - P_a)}{P_a} \right), \quad (2.60)$$

with

While there is no closed-form solution for (2.56) when the dictionary $\ddot{\mathbf{X}}$ is not an orthogonal matrix, it can be solved through quadratic programming.

By using the Landweber iterative algorithm [88], $\ddot{\mathbf{h}}_j$ is reconstructed in such a way that the elements of the reconstructed vector corresponding to the inactive IoT devices are zero without using the binary hypothesis testing in (2.48). The Landweber algorithm involves a gradient descent method with fixed step followed by a threshold setting based on the KKT optimality conditions as

$$-\ddot{\mathbf{x}}_{k,f}^\dagger (\mathbf{r}_j - \ddot{\mathbf{X}} \ddot{\mathbf{h}}_j) + \ddot{\lambda}_{\ell_1} \frac{\ddot{h}_{k,j,f}}{|\ddot{h}_{k,j,f}|} = 0 \quad \text{if } k \in \mathcal{X}_a \quad (2.61a)$$

$$\left| \ddot{\mathbf{x}}_{k,f}^\dagger (\mathbf{r}_j - \ddot{\mathbf{X}} \ddot{\mathbf{h}}_j) \right| \leq \ddot{\lambda}_{\ell_1} \quad \text{if } k \notin \mathcal{X}_a, \quad (2.61b)$$

where $f \in \{0, 1\}$, $\ddot{\mathbf{x}}_{k,f} = p_k \mathbf{x}_{k,f}$, and $\mathbf{x}_{k,1}$ and $\mathbf{x}_{k,0}$ are given in (2.11).

Let us assume that the convergence of the Landweber algorithm for the j th observation symbol occurs at the t_j th iteration. The average energy of the reconstructed signal for the k th IoT device is given as

$$e_k = \frac{1}{2l} \sum_{j=\bar{\alpha}}^{\bar{\alpha}+l-1} \sum_{f=0}^1 \left| \ddot{h}_{k,j,f}^{[t_j+1]} \right|^2. \quad (2.62)$$

In order to maximize the correct identification rate of the IoT devices, by employing *a priori* knowledge of P_a , the active IoT set can be identified as

$$\hat{\mathcal{X}}_a = \{e_{k_{i_0}}, e_{k_{i_1}}, \dots, e_{k_{K_{\text{up}}}}\}, \quad (2.63)$$

where

$$e_{k_{i_0}} \geq e_{k_{i_1}} \geq \dots \geq e_{k_{K_{\text{up}}}} \geq \dots \geq e_{k_{K_u-1}} \quad (2.64a)$$

$$e_{k_{i_0}} \geq e_{k_{i_1}} \geq \dots \geq e_{k_{K_{\text{up}}}} > 0, \quad (2.64b)$$

and

$$\mathbb{P}\{k_a \leq K_{\text{up}}\} = \sum_{i=0}^{K_{\text{up}}} \binom{K_u}{i} P_a^i (1 - P_a)^{(K_u-i)} = 0.999. \quad (2.65)$$

A formal description of the proposed ℓ_1 norm DSA IoT identification algorithm based of the Landweber iterative algorithm for SSR is provided in Algorithm 3.

Algorithm 3 ℓ_1 norm DSA IoT Identification

Input: \mathbf{X} , \mathbf{R} , δ , P_a
Output: Active IoT set $\hat{\mathcal{X}}_a$
Initialization: $\hat{\mathbf{H}} = \mathbf{0}$, $loop = 1$, $t = 0$

- 1: **for** $j = \bar{\alpha}, \bar{\alpha} + 1, \dots, \bar{\alpha} + l - 1$ **do**
- 2: **while** $loop$ **do**
- 3: $\ddot{\mathbf{h}}_j^{[t+\frac{1}{2}]} \leftarrow \ddot{\mathbf{h}}_j^{[t]} + \ddot{\mathbf{X}}^\dagger \left(\mathbf{r}_j - \ddot{\mathbf{X}} \ddot{\mathbf{h}}_j^{[t]} \right)$
- 4: **for** $k = 0, 1, \dots, K_u - 1$ **do**
- 5: $\ddot{h}_{k,j,0}^{[t+1]} = \left(1 - \frac{\ddot{\lambda}_{\ell_1}}{\left| \ddot{h}_{k,j,0}^{[t+\frac{1}{2}]} \right|} \right)_+ \ddot{h}_{k,j,0}^{[t+\frac{1}{2}]}$
- 6: $\ddot{h}_{k,j,1}^{[t+1]} = \left(1 - \frac{\ddot{\lambda}_{\ell_1}}{\left| \ddot{h}_{k,j,1}^{[t+\frac{1}{2}]} \right|} \right)_+ \ddot{h}_{k,j,1}^{[t+\frac{1}{2}]}$
- 7: **end for**
- 8: $t \leftarrow t + 1$
- 9: **if** $\left\| \ddot{\mathbf{h}}_j^{[t+1]} - \ddot{\mathbf{h}}_j^{[t]} \right\|_2 \leq \delta$ **then**
- 10: $loop = 0$
- 11: **end if**
- 12: **end while**
- 13: **end for**
- 14: Compute \mathbf{e}_k for $k = 0, 1, \dots, K_u - 1$ by employing (2.62)
- 15: Obtain $\hat{\mathcal{X}}_a$ through (2.63), (2.64), and (2.65)

2.3.4 $\ell_1 - \ell_2$ Mixed-norm PDSA IoT Identification Algorithm

The proposed $\ell_1 - \ell_2$ mixed-norm PDSA IoT identification algorithm replaces $\ell_0 - \ell_0$ quasi-norm in (2.30) with the $\ell_1 - \ell_2$ mixed-norm as

$$\hat{\mathbf{H}} = \underset{\mathbf{H}}{\operatorname{argmin}} \quad \frac{1}{2} \left\| \mathbf{R} - \mathbf{X}\mathbf{H} \right\|_{\mathbb{F}}^2 + \lambda_{\ell_2}^{\ell_1} \sum_{k=0}^{K_u-1} \left\| \mathbf{h}_{k,\bar{\alpha},1} \right\|_2, \quad (2.66)$$

where $\lambda_{\ell_1}^{\ell_2}$ is the tuning parameter. To find the optimal value of $\lambda_{\ell_1}^{\ell_2}$, the K -fold cross validation (CV) is a popular method; however, this suffers from high computational complexity. Another method is the cross-validated partial likelihood, an adaptation of the CV [89, 90]; however, it still suffers from the high computational cost.

For the block-sparse matrix \mathbf{H} in (2.20), since FCC is fixed over a packet, we can

write

$$\lambda_{\ell_0}^{\ell_0} \|\mathbf{H}\|_0 = \frac{\lambda_{\ell_0}^{\ell_0}}{l} \sum_{k=0}^{K_u-1} \|\mathbf{h}_{k,\bar{\alpha},l}\|_0 = \frac{\lambda_{\ell_0}^{\ell_0}}{l} \sum_{k=0}^{K_u-1} \frac{\sqrt{2} \|\mathbf{h}_{k,\bar{\alpha},l}\|_2}{\sqrt{l} |g_k|}, \quad (2.67)$$

where $\mathbf{h}_{k,\bar{\alpha},l}$ is given in (2.23). In order to have the solution of (2.66) the same as the one of the original SSSR problem in (2.30), we need to have

$$\lambda_{\ell_0}^{\ell_0} \|\mathbf{H}\|_0 = \lambda_{\ell_2}^{\ell_1} \sum_{k=0}^{K_u-1} \|\mathbf{h}_{k,\bar{\alpha},l}\|_2. \quad (2.68)$$

Let us define

$$\ddot{\mathbf{h}}_{k,\bar{\alpha},l} \triangleq \frac{\mathbf{h}_{k,\bar{\alpha},l}}{\sqrt{p_k}}, \quad (2.69a)$$

$$\ddot{\mathbf{H}} \triangleq \mathbf{P}^{-\frac{1}{2}} \mathbf{H}. \quad (2.69b)$$

Since $g_k \sim \mathcal{N}_c(0, p_k)$, $k \in \mathcal{X}_u$, $|g_k|$ follows the Rayleigh distribution. By replacing $|g_k|$ with $\mathbb{E}\{|g_k|\} = \sqrt{\pi p_k/4}$ in (2.67), and then, by substituting (2.69a) into the result, (2.68) can be approximated as

$$\lambda_{\ell_0}^{\ell_0} \|\mathbf{H}\|_0 = \lambda_{\ell_2}^{\ell_1} \sum_{k=0}^{K_u-1} \|\mathbf{h}_{k,\bar{\alpha},l}\|_2 \approx \frac{2\lambda_{\ell_0}^{\ell_0}}{\sqrt{l}\pi} \sum_{k=0}^{K_u-1} \frac{\|\mathbf{h}_{k,\bar{\alpha},l}\|_2}{\sqrt{p_k}} = \frac{2\lambda_{\ell_0}^{\ell_0}}{\sqrt{l}\pi} \sum_{k=0}^{K_u-1} \|\ddot{\mathbf{h}}_{k,\bar{\alpha},l}\|_2. \quad (2.70)$$

Furthermore, we have

$$\|\mathbf{R} - \mathbf{X}\mathbf{H}\|_{\mathbb{F}}^2 = \|\mathbf{R} - \ddot{\mathbf{X}}\ddot{\mathbf{H}}\|_{\mathbb{F}}^2, \quad (2.71)$$

where $\ddot{\mathbf{X}}$ and $\ddot{\mathbf{H}}$ are given in (2.57a) and (2.69b), respectively.

By substituting (2.70) and (2.140) into (2.66), the SSSR problem is reformulated

as

$$\hat{\mathbf{H}} = \underset{\mathbf{H}}{\operatorname{argmin}} \frac{1}{2} \|\mathbf{R} - \ddot{\mathbf{X}}\ddot{\mathbf{H}}\|_{\mathbf{F}}^2 + \ddot{\lambda}_{\ell_2}^{\ell_1} \sum_{k=0}^{K_u-1} \|\ddot{\mathbf{h}}_{k,\bar{\alpha},l}\|_2, \quad (2.72)$$

where $\ddot{\lambda}_{\ell_2}^{\ell_1}$ is the tuning parameter given as

$$\ddot{\lambda}_{\ell_2}^{\ell_1} = \frac{2\lambda_{\ell_0}^{\ell_0}}{\sqrt{l\pi}} = \frac{\sigma_w^2}{l\sqrt{l\pi}} \ln \left(\frac{2^{l+1}(1-P_a)}{P_a} \right). \quad (2.73)$$

The optimal solution of (2.72) is obtained with the gradient of the objective function equal to zero as

$$-\ddot{\mathbf{X}}^\dagger(\mathbf{R} - \ddot{\mathbf{X}}\ddot{\mathbf{H}}) + \ddot{\lambda}_{\ell_2}^{\ell_1} \mathbf{F}^{-2} \ddot{\mathbf{H}} = \mathbf{0}, \quad (2.74)$$

where

$$\mathbf{F} \triangleq \begin{bmatrix} \|\ddot{\mathbf{h}}_{k,\bar{\alpha},l}\|_2^{\frac{1}{2}} & & & \\ & \ddots & & \\ & & & \|\ddot{\mathbf{h}}_{K_u-1,\bar{\alpha},l}\|_2^{\frac{1}{2}} \end{bmatrix} \otimes \mathbf{I}_2 \quad (2.75)$$

After simple mathematical manipulations, (2.74) can be written as

$$\left[(\ddot{\mathbf{X}}\mathbf{F})^{\mathbf{H}}(\ddot{\mathbf{X}}\mathbf{F}) + \ddot{\lambda}_{\ell_2}^{\ell_2} \mathbf{I} \right] \mathbf{F}^{-1} \ddot{\mathbf{H}} = (\ddot{\mathbf{X}}\mathbf{F})^{\mathbf{H}} \mathbf{R}. \quad (2.76)$$

From (2.76), the optimal solution of (2.74) can be expressed as

$$\ddot{\mathbf{H}} = \mathbf{F} \left[(\ddot{\mathbf{X}}\mathbf{F})^{\mathbf{H}}(\ddot{\mathbf{X}}\mathbf{F}) + \ddot{\lambda}_{\ell_2}^{\ell_2} \mathbf{I} \right]^{-1} (\ddot{\mathbf{X}}\mathbf{F})^{\mathbf{H}} \mathbf{R}. \quad (2.77)$$

Since \mathbf{F} depends on \mathbf{H} as in (2.75), (2.77) suggests to obtain the optimal solution

through an iterative way as

$$\begin{aligned}\ddot{\mathbf{H}}^{[t+1]} &= \mathbf{F}^{[t]} \left[(\ddot{\mathbf{X}}\mathbf{F}^{[t]})^{\text{H}} (\ddot{\mathbf{X}}\mathbf{F}^{[t]}) + \ddot{\lambda}_{\ell_2}^{\ell_1} \mathbf{I} \right]^{-1} (\ddot{\mathbf{X}}\mathbf{F}^{[t]})^{\text{H}} \mathbf{R} \\ &= \mathbf{F}^{[t]} \left[(\mathbf{V}^{[t]})^{\text{H}} \mathbf{V}^{[t]} + \ddot{\lambda}_{\ell_2}^{\ell_1} \mathbf{I} \right]^{-1} (\mathbf{V}^{[t]})^{\text{H}} \mathbf{R}\end{aligned}\quad (2.78)$$

until the convergence criterion is satisfied as

$$\|\ddot{\mathbf{H}}^{[t+1]} - \ddot{\mathbf{H}}^{[t]}\|_2 < \delta, \quad (2.79)$$

where δ is an arbitrary small value. Appendix 2.K demonstrates that the $\ell_1 - \ell_2$ mixed-norm algorithm in (2.78) can be considered as an iterative reweighted LS estimator.

Similar to the Landweber iterative algorithm utilized in the ℓ_1 norm DSA IoT identification algorithm, the designed $\ell_1 - \ell_2$ mixed-norm PDSA IoT identification algorithm employs the gradient descent algorithm and reconstructs the matrix $\ddot{\mathbf{H}}$ such that the two consecutive rows corresponding to the inactive IoT devices are zero when $\ddot{\mathbf{H}}$ is initialized with an appropriate value. Since the gradient descent algorithm does not guarantee convergence to the global minimum, some elements of the reconstructed $\ddot{\mathbf{h}}_{k,\bar{\alpha},l}$ may not be zero when k th IoT device is inactive, and vice versa. Accordingly, the l_1 out of $2l$ combining rules can be employed to make decision on the transmission state of the IoT devices as

$$d_k = \begin{cases} H_{1k}, & \text{if } \sum_{j=\bar{\alpha}}^{l+\bar{\alpha}-1} \sum_{f=0}^1 \mathbb{I}\{\hat{h}_{k,j,f} \neq 0\} \geq l_1 \\ H_{0k}, & \text{if } \sum_{j=\bar{\alpha}}^{l+\bar{\alpha}-1} \sum_{f=0}^1 \mathbb{I}\{\hat{h}_{k,j,f} \neq 0\} < l_1 \end{cases}, \quad (2.80)$$

A formal description of the proposed $\ell_1 - \ell_2$ mixed-norm PDSA IoT identification algorithm is summarized in Algorithm 4.

Algorithm 4 $\ell_1 - \ell_2$ Mixed-norm PDSA IoT Identification Algorithm

Input: \mathbf{X} , \mathbf{R} , $\ddot{\lambda}_{\ell_2}^{\ell_1}$, ℓ_1 , δ
Output: Active IoT set $\hat{\mathcal{X}}_a$
Initialization: $\hat{\mathcal{X}}_a = \emptyset$, $\ddot{\mathbf{H}}^{[0]} = \mathbf{0}$

- 1: **while** $\|\ddot{\mathbf{H}}^{[t]} - \ddot{\mathbf{H}}^{[t-1]}\|_{\text{F}} > \delta$ **do**
- 2: Update $\mathbf{F}^{[t]}$ by $\ddot{\mathbf{H}}^{[t-1]}$ as in (2.75)
- 3: Obtain $\mathbf{V}^{[t]} = \ddot{\mathbf{X}}\mathbf{F}^{[t]}$
- 4: $\ddot{\mathbf{H}}^{[t+1]} \leftarrow \ddot{\mathbf{H}}^{[t]}$ by employing (2.78)
- 5: **end while**
- 6: **for** $k = 0, 1, \dots, K_u - 1$ **do**
- 7: Obtain d_k through (2.80)
- 8: **if** $d_k = H_{1k}$ **then**
- 9: $\hat{\mathcal{X}}_a \leftarrow \{\hat{\mathcal{X}}_a, k\}$
- 10: **end if**
- 11: **end for**

2.3.5 Adaptive $\ell_1 - \ell_2$ Mixed-norm PDSA IoT Identification

Algorithm

By employing (2.67), the SSSR problem in (2.30) can be written as

$$\hat{\mathbf{H}} = \arg \min_{\mathbf{H}} \frac{1}{2} \|\mathbf{R} - \mathbf{X}\mathbf{H}\|_{\text{F}}^2 + \lambda_{\ell_0}^{\ell_0} \sum_{k=0}^{K_u-1} \frac{\sqrt{2} \|\mathbf{h}_{k,\bar{\alpha},1}\|_2}{\sqrt{l} |g_k|}, \quad (2.81)$$

where $\lambda_{\ell_0}^{\ell_0}$ is given in (2.31).

By using $\mathbb{E}\{|g_k|\} = \sqrt{\pi p_k/4}$, an approximation for the penalty term in (2.81) is given as

$$\lambda_{\ell_0}^{\ell_0} \sum_{k=0}^{K_u-1} \frac{\|\mathbf{h}_{k,\bar{\alpha},1}\|_2}{\sqrt{l} |g_k|} \approx \sum_{k=0}^{K_u-1} \frac{2\lambda_{\ell_0}^{\ell_0}}{\sqrt{\pi l p_k}} \|\mathbf{h}_{k,\bar{\alpha},1}\|_2 \quad (2.82)$$

By substituting (2.82) into (2.81), one can write

$$\begin{aligned}\hat{\mathbf{H}} &= \arg \min_{\mathbf{H}} \mathcal{L}(\mathbf{H}, \lambda_0, \lambda_1, \dots, \lambda_{K_u-1}) \\ &= \arg \min_{\mathbf{H}} \frac{1}{2} \|\mathbf{R} - \mathbf{X}\mathbf{H}\|_{\text{F}}^2 + \sum_{k=0}^{K_u-1} \lambda_k \|\mathbf{h}_{k,\bar{\alpha},1}\|_2,\end{aligned}\tag{2.83}$$

where

$$\lambda_k \triangleq \frac{2\lambda_{\ell_0}^{\ell_0}}{\sqrt{\pi l p_k}} = \frac{\sigma_w^2}{l\sqrt{l\pi p_k}} \ln \left(\frac{2^{l+1}(1-P_a)}{P_a} \right).\tag{2.84}$$

Similar as in Appendix 2.J, the KKT optimality conditions of the optimization problem in (2.83) are expressed as

$$-\boldsymbol{\Psi}_k + \lambda_k \frac{\mathbf{h}_{k,\bar{\alpha},l}}{\|\mathbf{h}_{k,\bar{\alpha},l}\|_2} = \mathbf{0} \quad \text{if } k \in \mathcal{X}_a\tag{2.85a}$$

$$\|\boldsymbol{\Psi}_k\|_2 < \lambda_k \quad \text{if } k \notin \mathcal{X}_a,\tag{2.85b}$$

where

$$\boldsymbol{\Psi}_k \triangleq \left[\Psi_{k,0,0} \ \Psi_{k,0,1} \ \cdots \ \Psi_{k,l-1,0} \ \Psi_{k,l-1,1} \right]^\dagger\tag{2.86a}$$

$$\left[\Psi_{k,0,f} \ \Psi_{k,1,f} \ \cdots \ \Psi_{k,l-1,f} \right] \triangleq \mathbf{x}_{k,f}^\dagger (\mathbf{R} - \mathbf{X}\mathbf{H}).\tag{2.86b}$$

Let us write $\boldsymbol{\Psi}_k$ as

$$\boldsymbol{\Psi}_k \triangleq \nabla_{\mathbf{h}_{k,\bar{\alpha},l}} \frac{1}{2} \|\mathbf{R} - \mathbf{X}\mathbf{H}\|_{\text{F}}^2 = \boldsymbol{\varphi}_k - \boldsymbol{\Lambda}_k \mathbf{h}_{k,\bar{\alpha},l}\tag{2.87}$$

Algorithm 5 : Adaptive $\ell_1 - \ell_2$ Mixed-norm PDSA IoT Identification Algorithm

Input: \mathbf{X} , \mathbf{R} , λ_k , $k = 0, 1, \dots, K_u - 1$, $\mathbf{\Lambda}_k$
Output: Active IoT set $\hat{\mathcal{X}}_a$
Initialization: $\hat{\mathcal{X}}_a = \emptyset$, $\mathbf{H}^{[0]} = \mathbf{0}$, $t = 1$

- 1: **while** $loop = 1$ **do**
- 2: **for** $k = 0, 1, \dots, K_u - 1$ **do**
- 3: **if** the KKT optimality conditions in (2.85) are not satisfied **then**
- 4: Compute $\boldsymbol{\varphi}_k^{[t-1]}$ for $\mathbf{H}_{-k}^{[t-1]}$ through (2.88) and (2.89)
- 5: $\boldsymbol{\varphi}_k^{[t]} \leftarrow \boldsymbol{\varphi}_k^{[t-1]}$
- 6: $\mathbf{h}_{k,\bar{\alpha},l}^{[t]} \leftarrow \mathbf{h}_{k,\bar{\alpha},l}^{[t-1]}$ through (2.91)
- 6: **end if**
- 7: $t \leftarrow t + 1$
- 8: **end for**
- 9: **if** $\|\mathbf{h}_{k,\bar{\alpha},l}^{[t]} - \mathbf{h}_{k,\bar{\alpha},l}^{[t-1]}\|_2 \leq \delta$ **then**
- 10: $loop = 0$
- 11: **end if**
- 12: **end while**
- 13: $\hat{\mathcal{X}}_a = \{k | \mathbf{h}_{k,\bar{\alpha},l}^{[t]} \neq \mathbf{0}, k = 0, 1, \dots, K_u - 1\}$

A formal description of the proposed adaptive $\ell_1 - \ell_2$ mixed-norm PDSA IoT identification algorithm is summarized in Algorithm 5.

2.4 Data Detection

While the LS estimator is the unbiased estimator with the smallest variance, the developed penalized-LS (P-LS) estimators in the IoT identification section trade off bias for variance. In other words, they employ a penalty term to balance both approximation error and sparsity level of the solution. Hence, the mean square error (MSE) of the estimated signal may be high so that it results in high bit error rate (BER) if they are used for data detection. Thus, the proposed biased P-LS estimators exhibit a good performance for the IoT identification as an inference problem; however, the estimated signals by the biased P-LS estimators result in higher BER in data

detection.

To detect the data stream of the active IoT devices, conventional suboptimal linear or non-linear MUD algorithms, such as decorrelating detector, MMSE detector, successive interference cancellation, and parallel interference cancellation can be applied to the identified active IoT set when FCC and CP are known at the gateway. However, in the defined short packet transmission scheme, the IoT devices do not use any pilot or preamble for the estimate of the FCC and CP. Hence, a new MUD blind to the FCC and CP is required.

In this section, we propose a new non-linear MUD algorithm which does not require estimates of the FCC and CP of the IoT devices. The proposed MUD algorithm is developed based on the differential coding used at the IoT devices and an unsupervised machine learning technique followed by differential decoding at the gateway.

2.4.1 2MC-MUD Algorithm

Fig. 2.10 illustrates the block diagram of the proposed 2MC-MUD algorithm. The output of the IoT identification algorithms in the previous section is a set of active IoT devices $\hat{\mathcal{X}}_a$. Without loss of generality, we assume that $\hat{\mathcal{X}}_a \triangleq \{k_0, k_1, \dots, k_{\hat{K}_a-1}\}$, where $\hat{K}_a \triangleq \mathbf{card}(\hat{\mathcal{X}}_a)$ is the cardinality of $\hat{\mathcal{X}}_a$.

Let us consider a bank of \hat{K}_a single-user MFs for the identified active IoT devices in $\hat{\mathcal{X}}_a$ as shown in Fig. 2.10. The output of the MF after synchronized sampling for

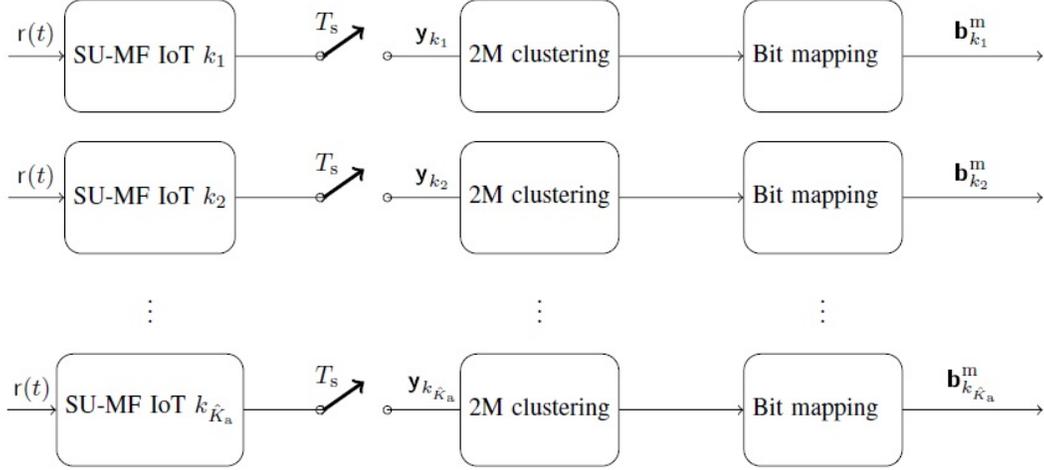


Fig. 2.10: Block diagram of the proposed 2MC-MUD algorithm

the k_n th IoT device is expressed as [91–93]

$$\begin{aligned}
y_{k_n}[i] &\triangleq \frac{1}{2} \int_{\tau_{k_n} + iT_s}^{\tau_{k_n} + (i+1)T_s} r(t) s_{k_n}(t - iT_s - \tau_{k_n}) dt \\
&= g_{k_j} b_{k_j}[i] + \sum_{k_j < k_n} g_{k_j} b_{k_j}[i+1] \rho_{k_n k_j} + \sum_{k_j < k_n} g_{k_j} b_{k_j}[i] \rho_{k_j k_n} + \sum_{k_j > k_n} g_{k_j} b_{k_j}[i] \rho_{k_n k_j} \\
&\quad + \sum_{k_j > k_n} g_{k_j} b_{k_j}[i-1] \rho_{k_j k_n} + w_{k_n}[i], \quad i = \{1, \dots, N_s\}
\end{aligned} \tag{2.92}$$

where

$$\rho_{k_n k_j} \triangleq \frac{1}{T_s} \int_0^{T_s} s_{k_n}(t) s_{k_j}(t) dt \tag{2.93}$$

and

$$w_{k_n}[i] \triangleq \frac{1}{2} \int_{\tau_{k_n} + iT_s}^{\tau_{k_n} + (i+1)T_s} w(t) s_{k_n}(t - iT_s - \tau_{k_n}) dt. \tag{2.94}$$

The output of the synchronized single-user MF in (2.92) for the k_n th IoT device

can be written as

$$y_{k_n}[i] = g_{k_n} b_{k_n}[i] + v_{k_n}[i], \quad i = 0, 1, \dots, N_s - 1, \quad (2.95)$$

where $v_{k_n}[i]$ represents the effect of noise and interference. The effects of FCC and CP are captured by the random variable g_{k_n} , $k_n \in \hat{\mathcal{X}}_a$, in (2.95) which is unknown at the gateway.

For data detection without any sign ambiguity, the phase of g_{k_n} , $k_n \in \hat{\mathcal{X}}_a$ is leastwise required to be known at the gateway. However, by employing differential coding at IoT devices, a MUD algorithm can be developed which removes the need for such *a priori* knowledge. Differential coding is a technique used to provide unambiguous signal reception in phase-shift-keying and quadrature amplitude modulation. Instead of encoding a bit sequence directly, the differential coding technique encodes the difference between the bit sequence as [94]

$$b_{k_n}[i] = b_{k_n}[i-1] \oplus b_{k_n}^c[i], \quad k_n \in \mathcal{X}_a, \quad (2.96)$$

where \oplus is the modulo-2 addition and $b_{k_n}^c[i]$ is the i th bit at the output of the channel encoder of the k_n th IoT device.

Since g_{k_n} , $k_n \in \mathcal{X}_a$, remains unchanged during an IoT packet, the received noisy BPSK symbols of the active IoT device k_n in (2.95) form two clusters corresponding to the transmitted bits 1 and 0. The main idea behind the proposed MUD is to extract the two clusters regardless of which cluster is labeled 1 or 0. By extraction of the two clusters and differential decoding at the gateways, the data stream of the active IoT device k_n is obtained without any prior knowledge about the FCC and CP.

By applying the 2MC algorithm to $y_{k_n}[i]$, $i = 0, 1, \dots, N_s - 1$, in (2.95), the two clusters are separated based on the nearest mean criterion disregard to the label. The

2MC minimizes the least within-cluster sum of squares (WCSS), i.e, then sum of the squared Euclidean distance.

Let us define $\mathcal{U} \triangleq \{0, 1, \dots, N_s - 1\}$. The 2MC algorithms partitions \mathcal{U} into two sets $\mathcal{U}_{k_n,0}$ and $\mathcal{U}_{k_n,1}$ for the active IoT device k_n by minimizing the WCSS as

$$\begin{aligned} \arg \min_{\mathcal{U}} \quad & \sum_{i \in \mathcal{U}_{k_n,0}} \left\| y_{k_n}[i] - \mu_{k_n,0} \right\|^2 + \sum_{i \in \mathcal{U}_{k_n,1}} \left\| y_{k_n}[i] - \mu_{k_n,1} \right\|^2, \\ \text{subject to} \quad & \mu_{k_n,0} = \frac{1}{\text{card}(\mathcal{U}_{k_n,0})} \sum_{i \in \mathcal{U}_{k_n,0}} y_{k_n}[i], \\ & \mu_{k_n,1} = \frac{1}{\text{card}(\mathcal{U}_{k_n,1})} \sum_{i \in \mathcal{U}_{k_n,1}} y_{k_n}[i]. \end{aligned} \quad (2.97)$$

The minimization problem in (2.97) can be solved by different methods. One of the most common algorithm is th Lloyd's algorithm which uses an iterative refinement technique. Given an initial mean values $\mu_{k_n,0}^{[0]}$ and $\mu_{k_n,1}^{[1]}$, the Lloyd's algorithm proceeds by alternating between the assignment and updating steps as

Assignment step: The element of $\mathcal{U}^{[t]}$ is assigned to $\mathcal{U}_{k_n,1}^{[t]}$ when

$$\mathcal{U}_{k_n,1}^{[t]} = \left\{ i : \left\| y_{k_n}[i] - \mu_{k_n,0}^{[t]} \right\|^2 \leq \left\| y_{k_n}[i] - \mu_{k_n,1}^{[t]} \right\|^2 \right\}. \quad (2.98)$$

Otherwise, it is assigned to $\mathcal{U}_{k_n,0}^{[t]}$.

Updating step: The mean of the the clusters $\mathcal{U}_{k_n,0}^{[t]}$ and $\mathcal{U}_{k_n,1}^{[t]}$ are updated as

$$\mu_{k_n,1}^{[t+1]} = \frac{1}{\text{card}(\mathcal{U}_{k_n,1}^{[t]})} \sum_{i \in \mathcal{U}_{k_n,1}^{[t]}} y_{k_n}[i], \quad (2.99a)$$

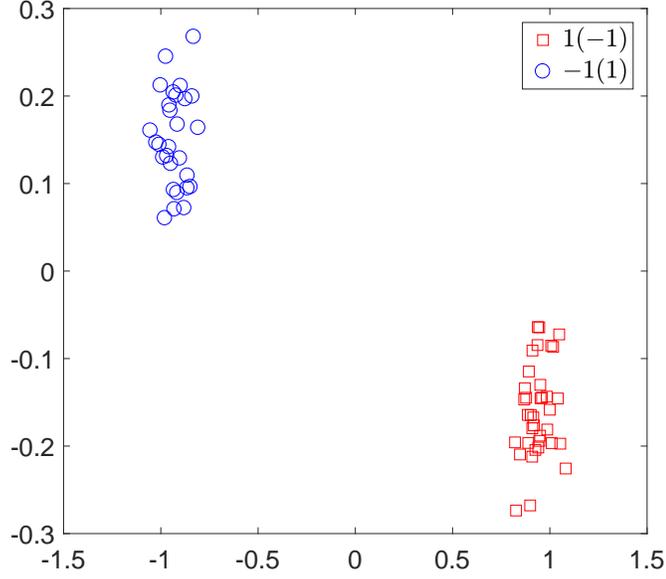


Fig. 2.11: 2M classifier for short IoT packet.

$$\boldsymbol{\mu}_{k_n,0}^{[t+1]} = \frac{1}{\text{card}(\mathcal{U}_{k_n,0}^{[t]})} \sum_{i \in \mathcal{U}_{k_n,0}^{[t]}} y_{k_n}[i]. \quad (2.99b)$$

The 2MC algorithm converges when the assignment step does not change. Fig. 2.11 shows the output of the 2MC algorithm for two active IoT devices. As seen, by employing the 2MC algorithm, the sequence at the output of the MF is partitioned into two clusters disregarding the label.

After partitioning \mathcal{U} into two clusters $\mathcal{U}_{k_n,0}$ and $\mathcal{U}_{k_n,1}$, $y_{k_n}[i]$, $i = 0, 1, \dots, N_s - 1$ is mapped into a binary sequence $\mathbf{b}_{k_n}^m \triangleq [b_{k_n}^m[0] \ b_{k_n}^m[1] \ \dots \ b_{k_n}^m[N_s - 1]]$ with elements as

$$b_{k_n}^m[i] = \mathbb{I}\{i \in \mathcal{U}_{k_n,1}\}. \quad (2.100)$$

Then, by applying differential decoding to the mapped binary sequence $\mathbf{b}_{k_n}^m$, the channel coded data stream for the active IoT device k_n is obtained as

$$\hat{\mathbf{b}}_{k_n}^c [i] = \mathbf{b}_{k_n}^m [i] \oplus \mathbf{b}_{k_n}^m [i - 1]. \quad (2.101)$$

Finally, $\hat{\mathbf{b}}_{k_n}^c \triangleq [\hat{\mathbf{b}}_{k_n}^c [0] \hat{\mathbf{b}}_{k_n}^c [1] \cdots \hat{\mathbf{b}}_{k_n}^c [N_c - 2]]$ is decoded by the channel decoder and the data stream of the active IoT device k_n is obtained. The proposed 2MC-MUD algorithm is summarized in Algorithm 6.

Algorithm 6 : 2MC-MUD

Input: $r(t)$, $\hat{\mathcal{X}}_a$, $\hat{K}_a = \text{card}(\hat{\mathcal{X}}_a)$

Output: $\hat{\mathbf{b}}_{k_n}$, $k_n \in \hat{\mathcal{X}}_a$

- 1: **for** $n = 0, 1, \dots, \hat{K}_a - 1$ **do**
 - 2: Set initial value for $\mathcal{U}_{k_n,1}^{[0]}$ and $\mathcal{U}_{k_n,0}^{[0]}$
 - 3: Obtain $y_{k_n}[i]$, $i = 0, 1, \dots, N_s - 1$, by employing (2.92)
 - 4: **while** $\mathcal{U}_{k_n,1}^{[t+1]} \neq \mathcal{U}_{k_n,1}^{[t]}$ **do**
 - 5: obtain $\mathcal{U}_{k_n,1}^{[t]}$ and $\mathcal{U}_{k_n,0}^{[t]}$ by employing (2.98)
 - 6: $\mu_{k_n,1}^{[t+1]} \leftarrow \mathcal{U}_{k_n,1}^{[t]}$ by employing (2.99a)
 - 7: $\mu_{k_n,0}^{[t+1]} \leftarrow \mathcal{U}_{k_n,0}^{[t]}$ by employing (2.99b)
 - 8: **end while**
 - 9: Obtain the binary mapped sequence $\mathbf{b}_{k_n}^m$ through (2.100)
 - 10: Apply differential decoding to $\mathbf{b}_{k_n}^m$ to obtain $\hat{\mathbf{b}}_{k_n}^c$ as
in (2.101)
 - 11: Apply channel decoding to $\hat{\mathbf{b}}_{k_n}^c$ to obtain $\hat{\mathbf{b}}_{k_n}$
 - 12: **end for**
-

2.4.2 Downlink Transmission

After packet transmission, the active IoT devices change their transmission mode into receive mode in the next observation slot in order to receive the acknowledgment packet transmitted by the gateway over the downlink channel. If the acknowledgment packet is not received by the active IoT device, the packet is retransmitted. This procedure continues until the reception of the acknowledgment packet by the IoT device. Based on this mechanism, the IoT devices are shortly active, either in the transmission mode or reception mode. Hence, the designed MA scheme is significantly power-efficient and capable of supporting low power IoT devices.

2.5 Simulation Results

In this section, we examine the performance of the designed IoT identification algorithms and 2MC-MUD algorithm through several simulation experiments.

2.5.1 Simulation Setup

Unless otherwise specified, an IoT network with $K_u = 2048$ and $P_a = 0.01$ was considered. It is assumed that the spreading sequence of the IoT devices is pseudo-random number codes with spreading factor $N_c = 512$ being known at the gateway. Each IoT packet is 15 bytes, and the delay of the IoT devices was generated through uniform distributions as $\alpha_k \sim \mathcal{U}[0, 5]$, $\beta_k \sim \mathcal{U}[0, 511]$, and $\epsilon_k \sim \mathcal{U}[0, 1]$.

For channel coding, the BCH encoder with codeword length 15 is applied to 10 messages each with 7 bits. The window size for IoT identification was considered $l = 16$ for $\bar{\alpha} = 7$. The effect of FCC and CP for each IoT device was modeled as independent circular complex Gaussian random variables with variance $\sigma_{g_k}^2 = 1$, $k \in \mathcal{X}_u$. The additive white noise was modeled as circular complex Gaussian random

variables with variance σ_w^2 , and the average system signal-to-noise ratio (SNR) was defined as

$$\gamma = \frac{P_a \sum_{k=0}^{K_u-1} p_k}{\sigma_w^2}.$$

The designed MA scheme was examined for equal power IoT devices.

The performance of the proposed DSA and PDSA IoT identification algorithms was evaluated in terms of system correct identification and system false alarm rates defined as

$$P^{(c)} = \frac{1}{K_a} \sum_{k \in \mathcal{X}_a} P_k^{(c)}$$

$$P^{(f)} = \frac{1}{K_u - K_a} \sum_{k \in \mathcal{X}_u - \mathcal{X}_a} P_k^{(f)}$$

respectively. $P^{(c)}$ and $P^{(f)}$, as performance measures, were obtained based on 1000 Monte Carlo trials. The performance of the designed 2MC-MUD algorithm was evaluated in terms of average packet error rate (PER).

2.5.2 Simulation Results

Fig. 2.12 depicts the correct identification rate, $P^{(c)}$, and false alarm rate, $P^{(f)}$, of the designed ML squared ℓ_2 norm DSA IoT identification algorithm (Algorithm 2) versus SNR for $P_k^{(f)} = 0.01$ and $P_k^{(f)} = 0.02$. As seen, $P^{(c)} > 0.93$ over a wide range of SNR. Also, as expected, $P^{(c)}$ increases when SNR and/or false alarm rate increases.

In Fig. 2.13, the empirical correct identification rate of the proposed ML squared ℓ_2 norm IoT identification algorithm (Algorithm 2) versus time is illustrated. As

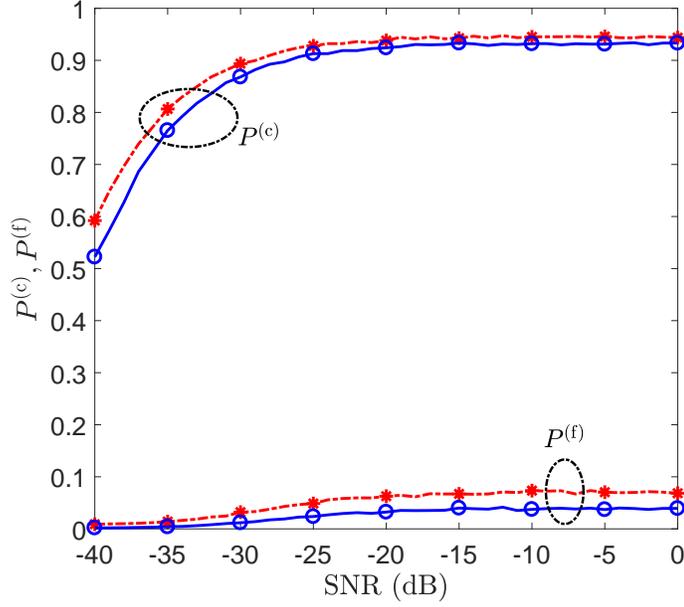


Fig. 2.12: The correct identification rate, $P^{(c)}$, and false alarm rate, $P^{(f)}$, of the designed ML squared ℓ_2 norm DSA IoT identification algorithm (Algorithm 2) versus SNR for $K_u = 2048$, $P_a = 0.01$, and $l = 16$. The solid and dashed lines represent the results for $P_k^{(f)} = 0.01$ and $P_k^{(f)} = 0.02$, respectively.

seen, the proposed algorithm accurately tracks the activity of the IoT devices in the network.

In Fig. 2.14, the correct identification, $P^{(c)}$, and false alarm, $P^{(f)}$, rates of the proposed ℓ_1 norm DSA IoT identification algorithm (Algorithm 3) versus SNR are illustrated. As seen, the proposed algorithm exhibits high $P^{(c)}$ and low $P^{(f)}$ over a wide range of SNRs.

Fig. 2.15 compares the performance of the proposed $\ell_1 - \ell_2$ mixed-norm (Algorithm 4) and adaptive $\ell_1 - \ell_2$ mixed-norm (Algorithm 5) PDSA IoT identification algorithms when $p_1 = p_2 = \dots = p_{1024} = 0.6$ and $p_{1025} = p_{1026} = \dots = p_{2048} = 1.4$. As seen, the adaptive $\ell_1 - \ell_2$ mixed-norm PDSA IoT identification algorithm outperforms the $\ell_1 - \ell_2$ mixed-norm algorithm. This behaviour can be explained, as the tuning

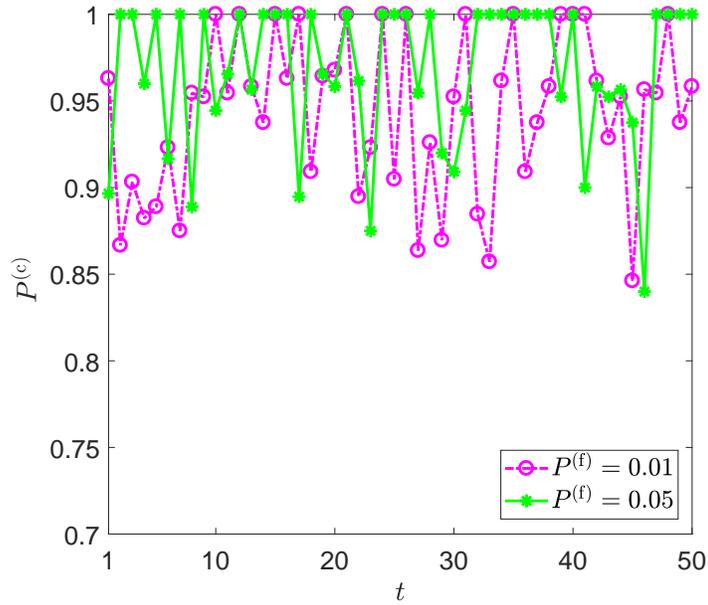


Fig. 2.13: The empirical correct identification rate, $P^{(c)}$, of the designed ML squared ℓ_2 norm DSA IoT identification algorithm (Algorithm 2) versus time for $K_u = 2048$, $P_a = 0.01$, $l = 16$, and at 10 dB SNR.

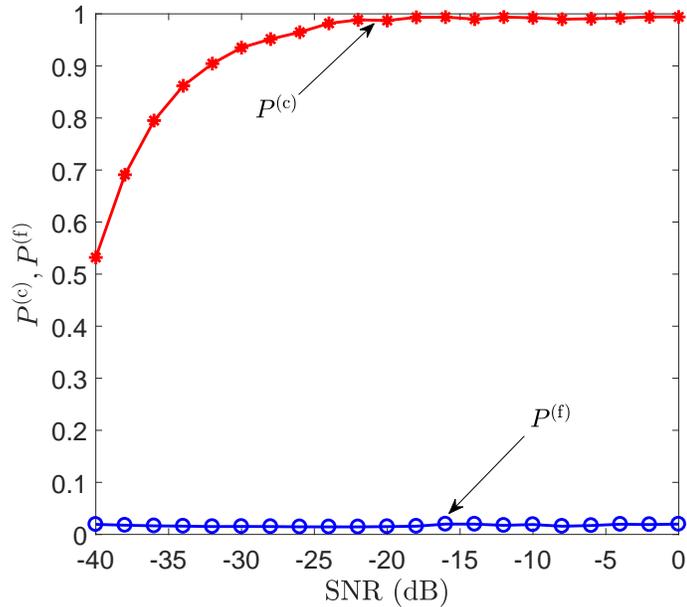


Fig. 2.14: The correct identification rate, $P^{(c)}$, and false alarm rate, $P^{(f)}$, of the designed ℓ_1 norm DSA IoT identification algorithm (Algorithm 3) versus SNR for $K_u = 2048$, $P_a = 0.01$, and $l = 16$.

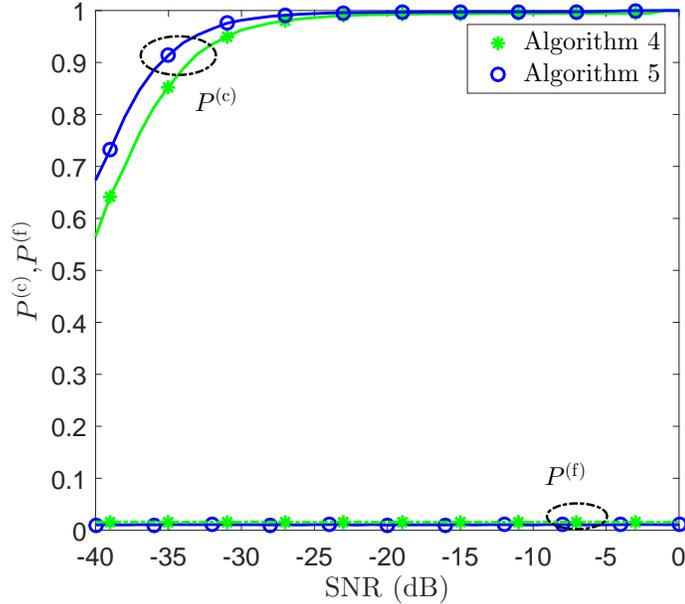


Fig. 2.15: Performance comparison of the proposed $\ell_1 - \ell_2$ mixed-norm (Algorithm 4) and adaptive $\ell_1 - \ell_2$ mixed-norm (Algorithm 5) PDSA IoT identification algorithms.

parameter for each IoT devices is set based on its transmit power for the proposed adaptive algorithm while this is not the case in the $\ell_1 - \ell_2$ mixed-norm algorithm.

Fig. 2.16 shows the PER of the proposed 2MC-MUD algorithm versus SNR when the proposed $\ell_1 - \ell_2$ mixed-norm PDSA IoT identification algorithm (Algorithm 5) is employed at the gateway. As seen, the proposed 2MC-MUD algorithm exhibits a good performance over a wide range of SNRs without FCC and CP estimation.

2.6 Conclusions and Directions for Future Research

In Chapter 2 of this thesis, uplink MA in IoT was studied. In this section, a summary of the main results in Chapter 2 is provided, and possible directions for future research are pointed out.

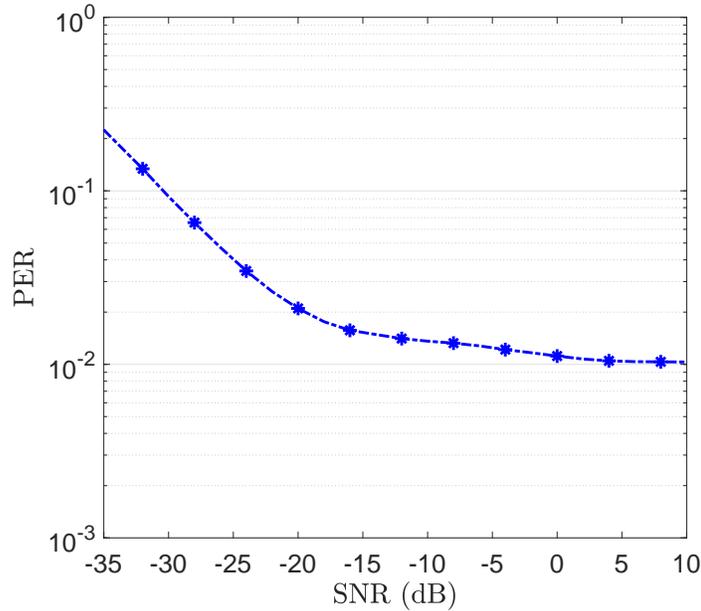


Fig. 2.16: PER of the proposed 2MC-MUD algorithm versus SNR when the proposed $\ell_1 - \ell_2$ mixed-norm PDSA IoT identification algorithm (Algorithm 5) is employed at the gateway.

2.6.1 Summary

- A new MA scheme for the uplink transmission of IoT was designed. The proposed MA scheme exhibits the following advantages:
 - It supports thousandths of uncoordinated IoT devices;
 - It supports sporadic traffic pattern and short packet transmission in IoT applications;
 - It was designed for underdetermined DS-SS. Thus, packet time on-air significantly reduces;
 - It removes the need of control signaling associated with the MAC address to reduce uplink packet overhead;
 - It removes preambles and pilot employed for parameter and channel estimation to reduce uplink overhead;

- It increases the spectral efficiency of the system by decreasing uplink overhead;
 - It is power-efficient and capable of supporting low-power IoT devices since the IoT devices are shortly active, either in transmission or reception mode;
 - It exhibits high scalability in terms of adding new IoT devices without negatively affecting the quality of existing services;
 - improves the protection against interfering and jamming signals. It also provides security of transmission if the codes are not known to the public.
- A new mechanism instead of the MAC address for IoT identification at the gateway was developed;
 - Since IoT devices do not use MAC address in order to identify themselves to the gateway, different IoT identification algorithms were designed to detect active IoT devices based on SSR and SSSR techniques;
 - The statistical performance analysis of the identification algorithms was presented, and closed-form expressions for the correct identification and false alarm rates were derived;
 - Approximate closed-form expressions for the tuning parameter employed in the non-linear optimization problems of the SSR and SSSR were derived based on the statistics of the FCC.
 - A new non-linear MUD algorithm, i.e, 2MC-MUD, was designed for short packet transmission over flat fading channel. The proposed 2MC-MUD algorithm exhibits the following advantages:
 - It supports both synchronous and asynchronous users irrespective to the traffic pattern.

- It does not require FCC and CP estimation;

2.6.2 Future research

The results in Chapter 2 of this thesis open interesting directions for a number of future research topics, as follows:

- An extension to the proposed uplink MA in order to support a large number of IoT devices in frequency-selective fading channel;
- An extension to the proposed massive uplink MA to support different traffic patterns;
- Developing a downlink MA scheme for the proposed uplink MA.

Appendix

2.A Proof of the Tuning Parameter for the ℓ_0 Quasi-norm

Let us write the MAP estimate of \mathbf{h}_j , $\bar{\alpha} \leq j \leq \bar{\alpha} + l - 1$ in (2.12) as

$$\begin{aligned}\hat{\mathbf{h}}_j &= \arg \max_{\mathbf{h}_j} p(\mathbf{h}_j | \mathbf{r}_j) \\ &= \arg \min_{\mathbf{h}_j} -\ln p(\mathbf{r}_j | \mathbf{h}_j) - \ln p(\mathbf{h}_j)\end{aligned}\tag{2.102}$$

Since \mathbf{w}_j in (2.12) is white Gaussian vector, one can write

$$p(\mathbf{r}_j | \mathbf{h}_j) = \frac{1}{(\pi\sigma_w^2)^{N_c}} \exp\left(-\frac{\|\mathbf{r}_j - \mathbf{X}\mathbf{h}_j\|_{\mathbb{F}}^2}{\sigma_w^2}\right).\tag{2.103}$$

For $p(\mathbf{h}_j)$ in (2.102), by employing (2.12), $\mathbf{g}_k \sim \mathcal{N}_c(0, p_k)$, $k \in \mathcal{X}_u$, $\mathbb{E}\{\mathbf{g}_{k_m} \mathbf{g}_{k_n}^*\} = p_{k_m} \delta_{k_m, k_n}$, $\|\mathbf{h}_{k,j}\|_0 = \|\mathbf{b}_{k,j}\|_0$, and the fact that \mathbf{G} and \mathbf{b}_j in (2.12) are independent, one obtains

$$p(\mathbf{h}_j) = \prod_{k=0}^{K_u-1} (1 - P_a)^{1 - \frac{\|\mathbf{h}_{k,j}\|_0}{2}} \left(\frac{P_a}{4}\right)^{\frac{\|\mathbf{h}_{k,j}\|_0}{2}} \frac{1}{\pi p_k} \exp\left(-\frac{|\mathbf{g}_k|^2}{p_k}\right),\tag{2.104}$$

where $\mathbf{h}_{k,j} = [\mathbf{h}_{k,j,0} \ \mathbf{h}_{k,j,1}]^\dagger$. By substituting (2.103) and (2.104) into (2.102), and after some mathematical manipulations, (2.26) and (2.27) are derived.

2.B Proof of the Tuning Parameter for the $\ell_0 - \ell_0$ Quasi-norm

Let us write MAP estimate of \mathbf{H} in (2.20) as

$$\begin{aligned} \hat{\mathbf{H}} &= \arg \max_{\mathbf{H}} p(\mathbf{H}|\mathbf{R}) \\ &= \arg \min_{\mathbf{H}} -\ln p(\mathbf{R}|\mathbf{H}) - \ln p(\mathbf{H}) \end{aligned} \quad (2.105)$$

Since \mathbf{W} is zero-mean i.i.d. Gaussian matrix, one can write

$$p(\mathbf{R}|\mathbf{H}) = \frac{1}{(\pi\sigma_w^2)^{N_{cl}}} \exp\left(-\frac{\|\mathbf{R} - \mathbf{X}\mathbf{H}\|_F^2}{\sigma_w^2}\right). \quad (2.106)$$

For $p(\mathbf{H})$ in (2.105), by employing (2.20), $\mathbf{g}_k \sim \mathcal{N}_c(0, p_k)$, $k \in \mathcal{X}_u$, $\mathbb{E}\{\mathbf{g}_{k_m}\mathbf{g}_{k_n}^*\} = p_k\delta_{k_m, k_n}$, $\|\mathbf{b}_{k,\bar{\alpha},l}\|_0 = \|\mathbf{h}_{k,\bar{\alpha},l}\|_0$, and the fact that \mathbf{G} and \mathbf{B} in (2.20) are matrixes with independent elements, one obtains

$$p(\mathbf{H}) = p(\mathbf{B})p(\mathbf{G}) = \prod_{k=1}^K (1 - P_a)^{1 - \frac{\|\mathbf{h}_{k,\bar{\alpha},l}\|_0}{2^l}} \left(\frac{P_a}{2^{l+1}}\right)^{\frac{\|\mathbf{h}_{k,\bar{\alpha},l}\|_0}{2^l}} \frac{1}{\pi p_k} \exp\left(-\frac{|\mathbf{g}_k|^2}{p_k}\right). \quad (2.107)$$

By substituting (2.106) and (2.107) into (2.105), and after some mathematical manipulations, (2.30) and (2.31) are derived.

2.C Proof of the Tuning Parameter for the Squared ℓ_2 Norm

In [95], it is shown that the optimal tuning parameter of the RD estimator for $\mathbf{r}_j = \mathbf{X}\mathbf{h}_j + \mathbf{w}_j$ in terms of MMSE is obtained as

$$\lambda_j^{\text{op}} \approx \frac{\sigma_w^2 \text{tr}[\bar{\Sigma}_{\mathbf{X}}^{-1}]}{\mathbf{h}_j^{\text{H}} \bar{\Sigma}_{\mathbf{X}}^{-1} \mathbf{h}_j + 3\text{tr}[\bar{\Sigma}_{\mathbf{X}}^{-2}]}, \quad (2.108)$$

where $\bar{\Sigma}_{\mathbf{X}} \triangleq \mathbf{X}^\dagger \mathbf{X}$. As observed, λ_j^{op} depends on \mathbf{h}_j which is unknown and needs to be estimated by the RD estimator. In this case, a reasonable approximation of (2.108) can be obtained by replacing $\mathbf{h}_j^{\text{H}} \bar{\Sigma}_{\mathbf{X}}^{-1} \mathbf{h}_j$ with its expected value.

Since the elements of \mathbf{h}_j are uncorrelated, by employing $\mathbb{E}\{|h_{k,j,0}|^2\} = \mathbb{E}\{|h_{k,j,1}|^2\} = P_a p_k$, $k \in \mathcal{X}_u$, one can easily write

$$\mathbb{E}\left\{\mathbf{h}_j^{\text{H}} \bar{\Sigma}_{\mathbf{X}}^{-1} \mathbf{h}_j\right\} = P_a (\mathbf{p}^\dagger \otimes \mathbf{1}^\dagger) \bar{\Lambda}_{\mathbf{X}}, \quad (2.109)$$

where $\mathbf{p} \triangleq [p_0 \ p_1 \ \dots \ p_{K_u-1}]^\dagger$, and $\bar{\Lambda}_{\mathbf{X}} \triangleq \text{diag}(\bar{\Sigma}_{\mathbf{X}}^{-1})$. Finally, by substituting (2.109) into (2.108), (2.36) is obtained.

2.D Distribution of the RD Estimator

Based on the central limit theorem for dependent random variables in [96], one can show that the distribution of $\hat{\mathbf{h}}_{k,\bar{\alpha},l}$ given in (2.42) and (2.41) can be approximated by a joint complex Gaussian distribution under hypothesis H_{0k} and H_{1k} .

By applying the statistical expectation to (2.41), and employing $\mathbb{E}\{h_{k,j,f} | H_{tk}\} = 0$ and $\mathbb{E}\{h_{k,j,\bar{f}} | H_{tk}\} = 0$, one easily obtain $\mathbb{E}\{\hat{\mathbf{h}}_{k,\bar{\alpha},l} | H_{tk}\} = \mathbf{0}$.

To obtain the diagonal elements of the covariance matrix Σ^{tk} , i.e., $\Sigma_{2j'+f, 2j'+f}^{tk}$,

$t \in \{0, 1\}$, let us consider the general rule in (2.110) for computing the variance of multiple complex random variables z_0, z_2, \dots, z_{n-1} .

$$\begin{aligned} \mathbb{V}\text{ar}\left\{\sum_{i=1}^n a_i z_i\right\} &= \sum_{i=0}^{n-1} |a_i|^2 \mathbb{V}\text{ar}\{z_i\} \\ &+ \sum_{i=0}^{n-1} \sum_{j \neq i} a_i a_j^* \text{Cov}\{z_i, z_j\}. \end{aligned} \quad (2.110)$$

Since $h_{k_1, j, f}$ and $h_{k_2, j, \bar{f}}$, and $w'_{k, j, f}$ $k_1, k_2 \in \mathcal{X}_u$, in (2.41), are zero-mean and uncorrelated random variables, by applying (2.110) to (2.41), one can write

$$\begin{aligned} \Sigma_{2j'+f, 2j'+f}^{tk} &= \mathbb{V}\text{ar}\left\{\hat{h}_{k, j, f} | H_{tk}\right\} = \mathbb{E}\left\{\left|\hat{h}_{k, j, f}\right|^2 | H_{tk}\right\} = t\Omega_{2k+f, 2k+f}^2 \mathbb{V}\text{ar}\left\{h_{k, j, f} | H_{tk}\right\} \\ &+ t\Omega_{2k+f, 2k+\bar{f}}^2 \mathbb{V}\text{ar}\left\{h_{k, j, \bar{f}} | H_{tk}\right\} + \sum_{n \neq k} \Omega_{2k+f, 2n+f}^2 \mathbb{V}\text{ar}\left\{h_{n, j, f} | H_{tk}\right\} \\ &+ \sum_{n \neq n} \Omega_{2k+f, 2n+\bar{f}}^2 \mathbb{V}\text{ar}\left\{h_{n, j, \bar{f}} | H_{tk}\right\} + \mathbb{V}\text{ar}\left\{w'_{k, j, f}\right\}. \end{aligned} \quad (2.111)$$

Further, by employing the law of total variance [97] and (2.17), $\mathbb{V}\text{ar}\{h_{k, j, f} | H_{1k}\}$, $f \in \{0, 1\}$ in (2.111) can be written as

$$\begin{aligned} \mathbb{V}\text{ar}\left\{h_{k, j, f} | H_{tk}\right\} &= \mathbb{V}\text{ar}\left\{g_k b_{k, j-\alpha_k-1+f} | H_{tk}\right\} = \mathbb{E}\left\{\mathbb{V}\text{ar}\left\{g_k b_{k, j-\alpha_k-1+f} | g_k, H_{tk}\right\}\right\} \\ &+ \mathbb{V}\text{ar}\left\{\mathbb{E}\left\{g_k b_{k, j-\alpha_k-1+f} | g_k, H_{tk}\right\}\right\} \end{aligned} \quad (2.112)$$

For $\bar{\alpha}+ \leq j \leq \bar{\alpha} + l - 1$, $\mathbb{V}\text{ar}\{b_{k, j-\alpha_k-1+f} | H_{tk}\} = t$, $f \in \{0, 1\}$, and one obtains

$$\mathbb{E}\left\{\mathbb{V}\text{ar}\left\{g_k b_{k, j-\alpha_k-1+f} | g_k, H_{tk}\right\}\right\} = \mathbb{E}\left\{|g_k|^2\right\} \mathbb{V}\text{ar}\left\{b_{k, j-\alpha_k-1+f} | H_{tk}\right\} = tp_k. \quad (2.113)$$

By substituting (2.113) and $\mathbb{E}\{g_k b_{k, j-\alpha_k-1+f} | g_k, H_{tk}\} = 0$, $f \in \{0, 1\}$, into (2.112),

one obtains

$$\text{Var}\left\{h_{k,j,f}|H_{tk}\right\} = \mathbb{E}\left\{|h_{k,j,f}|^2|H_{tk}\right\} = tp_k. \quad (2.114)$$

Similar to (2.112), for $n \neq k$ and $f \in \{0, 1\}$, we can write

$$\begin{aligned} \text{Var}\left\{h_{n,j,f}|H_{tk}\right\} &= \text{Var}\left\{g_n b_{n,j-\alpha_n-1+f}|H_{tk}\right\} = \mathbb{E}\left\{\text{Var}\left\{g_n b_{n,j-\alpha_n-1+f}|g_n, H_{tk}\right\}\right\} \\ &+ \text{Var}\left\{\mathbb{E}\left\{g_n b_{n,j-\alpha_n-1+f}|g_n, H_{tk}\right\}\right\} \end{aligned} \quad (2.115)$$

For $\bar{\alpha}+ \leq j \leq \bar{\alpha} + l - 1$, $\text{Var}\left\{b_{n,j-\alpha_n-1+f}|H_{tk}\right\} = P_a$, $f \in \{0, 1\}$, and one can write

$$\mathbb{E}\left\{\text{Var}\left\{g_n b_{n,j-\alpha_n-1+f}|g_n, H_{tk}\right\}\right\} = \mathbb{E}\left\{|g_n|^2\right\} \text{Var}\left\{b_{n,j-\alpha_n-1+f}|H_{tk}\right\} = P_a p_n. \quad (2.116)$$

By substituting (2.116) and $\mathbb{E}\left\{g_n b_{n,j-\alpha_n-1+f}|g_n, H_{tk}\right\} = 0$, $f \in \{0, 1\}$, into (2.115), one obtains

$$\text{Var}\left\{h_{n,j,f}|H_{tk}\right\} = \mathbb{E}\left\{|h_{n,j,f}|^2|H_{tk}\right\} = P_a p_n, \quad n \neq k \quad (2.117)$$

Finally, by substituting (2.114), (2.117), and $\text{Var}\left\{w'_{k,j,f}\right\} = \Sigma_{2k+f,2k+f}^{w'}$, into (2.111), (2.45) is obtained.

For the off-diagonal elements of the covariance Σ^{tk} , $t \in \{0, 1\}$, by using (2.41) and $\mathbb{E}\left\{\hat{h}_{k_1,j_1,f_1}\right\} = \mathbb{E}\left\{\hat{h}_{k_2,j_2,f_2}\right\} = 0$, $k_1, k_2 \in \mathcal{X}_u$ one can write (2.118). Based on the value of $j_1(j'_1)$, $j_2(j'_2)$, f_1 , and f_2 , six cases for the upper triangular off-diagonal elements of Σ^{tk} . i.e., $2j'_1 + f_1 < 2j'_2 + f_2$, can be considered. It is obvious that the lower upper triangular off-diagonal elements of Σ^{tk} , $t \in \{0, 1\}$ are easily obtain by applying the Hermitian operand to the upper triangular off-diagonal elements.

$$\begin{aligned}
\Sigma_{2j'_1+f_1, 2j'_2+f_2}^{tk} &= \text{Cov}\left\{\hat{h}_{k,j_1,f_1}, \hat{h}_{k,j_2,f_2}^* \middle| H_{tk}\right\} = \mathbb{E}\left\{\hat{h}_{k,j_1,f_1} \hat{h}_{k,j_2,f_2}^* \middle| H_{tk}\right\} \\
&= \mathbb{E}\left\{\left(t\left(\Omega_{2k+f_1, 2k+f_1} h_{k,j_1,f_1} + \Omega_{2k+f_1, 2k+\bar{f}_1} h_{k,j_1,\bar{f}_1}\right) + \sum_{n \neq k} \left\{\Omega_{2k+f_1, 2n+f_1} h_{n,j_2,f_1} + \Omega_{2k+f_1, 2n+\bar{f}_1} h_{n,j_1,\bar{f}_1}\right\} + w'_{k,j_1,f_1}\right)\right. \\
&\quad \left.\left(t\left(\Omega_{2k+f_2, 2k+f_2} h_{k,j_2,f_2}^* + \Omega_{2k+f_2, 2k+\bar{f}_2} h_{k,j_2,\bar{f}_2}^*\right) + \sum_{n \neq k} \left\{\Omega_{2k+f_2, 2n+f_2} h_{n,j_2,f_2}^* + \Omega_{2k+\bar{f}_2, 2n+\bar{f}_2} h_{n,j_2,\bar{f}_2}^*\right\} + w'_{k,j_2,f_2}\right) \middle| H_{tk}\right\}
\end{aligned} \tag{2.118}$$

By replacing $h_{k_2, j_2, 0}$ with $h_{k_1, j_1, 1}$, $k_1 = k_2$, $j_2 = j_1 + 1$, and employing

$$\mathbb{E}\left\{h_{k_1, j_1, f_1} h_{k_2, j_2, f_2}^* \middle| H_{tk}\right\} = 0, \quad |j_2 - j_1| > 1, \tag{2.119a}$$

$$\mathbb{E}\left\{h_{k_1, j_1, f_1} h_{k_2, j_2, f_2}^* \middle| H_{tk}\right\} = 0, \quad j_2 = j_1 + 1, f_1 = 0, \tag{2.119b}$$

$$\mathbb{E}\left\{h_{k_1, j_1, f_1} h_{k_2, j_2, f_2}^* \middle| H_{tk}\right\} = 0, \quad j_2 = j_1 + 1, f_1 = f_2 = 1, \tag{2.119c}$$

$$\mathbb{E}\left\{h_{k_1, j_1, f_1} h_{k_2, j_2, f_2}^* \middle| H_{tk}\right\} = 0, \quad j_2 = j_1, f_1 \neq f_2, \tag{2.119d}$$

$$\mathbb{E}\left\{w'_{k_1, j_1, f_1} h_{k_2, f_2, j_2}^* \middle| H_{tk}\right\} = 0, \tag{2.119e}$$

$$\mathbb{E}\left\{w'_{k_1, j_1, f_1} (w'_{k_2, j_2, f_2})^* \middle| H_{tk}\right\} = 0 \quad j_1 \neq j_2, \tag{2.119f}$$

one can write

- Case 1: $j_2 > j_1 + 1$

$$\Sigma_{2j'_1+f_1, 2j'_2+f_2}^{tk} = \text{Cov}\left\{\hat{h}_{k,j_1,f_1}, \hat{h}_{k,j_2,f_2} \middle| H_{tk}\right\} = 0. \tag{2.120}$$

- Case 2: $j_1 = j_2 = j$, $f_1 = 0$ and $f_2 = 1$

$$\begin{aligned}
\Sigma_{2j',2j'+1}^{1k} &= \text{Cov}\left\{\hat{h}_{k,j,0}, \hat{h}_{k,j,1} \middle| H_{tk}\right\} = t\left(\Omega_{2k,2k}\Omega_{2k+1,2k}\right)\mathbb{E}\left\{\left|h_{k,j,0}\right|^2 \middle| H_{tk}\right\} \\
&+ t\left(\Omega_{2k+1,2k+1}\Omega_{2k,2k+1}\right)\mathbb{E}\left\{\left|h_{k,j,1}\right|^2 \middle| H_{tk}\right\} + \sum_{n \neq k} \Omega_{2k,2n}\Omega_{2k+1,2n}\mathbb{E}\left\{\left|h_{n,j,0}\right|^2 \middle| H_{tk}\right\} \\
&+ \sum_{n \neq k} \Omega_{2k+1,2n+1}\Omega_{2k,2n+1}\mathbb{E}\left\{\left|h_{n,j,1}\right|^2 \middle| H_{tk}\right\} + \mathbb{E}\left\{w'_{k,j,0}\left(w'_{k,j,1}\right)^*\right\}
\end{aligned} \tag{2.121}$$

- Case 3: $j_1 = j$, $j_2 = j + 1$, and $f_1 = f_2 = 0$,

$$\begin{aligned}
\Sigma_{2j',2j'+2}^{tk} &= \text{Cov}\left\{\hat{h}_{k,j,0}, \hat{h}_{k,j+1,0} \middle| H_{1k}\right\} = t\left(\Omega_{2k,2k}\Omega_{2k,2k+1}\right)\mathbb{E}\left\{\left|h_{k,j,1}\right|^2 \middle| H_{tk}\right\} \\
&+ \sum_{n \neq k} \Omega_{2k,2n}\Omega_{2k,2n+1}\mathbb{E}\left\{\left|h_{n,j,1}\right|^2 \middle| H_{tk}\right\}
\end{aligned} \tag{2.122}$$

- Case 4: $j_1 = j$, $j_2 = j + 1$, and $f_1 = f_2 = 1$,

$$\begin{aligned}
\Sigma_{2j'+1,2j'+3}^{tk} &= \text{Cov}\left\{\hat{h}_{k,j,1}, \hat{h}_{k,j+1,1} \middle| H_{tk}\right\} = t\left(\Omega_{2k+1,2k+1}\Omega_{2k+1,2k}\right)\mathbb{E}\left\{\left|h_{k,j,1}\right|^2 \middle| H_{tk}\right\} \\
&+ \sum_{n \neq k} \Omega_{2k+1,2n+1}\Omega_{2k+1,2n}\mathbb{E}\left\{\left|h_{n,j,1}\right|^2 \middle| H_{tk}\right\}
\end{aligned} \tag{2.123}$$

- Case 5: $j_1 = j$, $j_2 = j + 1$, $f_1 = 0$, and $f_2 = 1$

$$\begin{aligned}
\Sigma_{2j',2j'+3}^{tk} &= \text{Cov}\left\{\hat{h}_{k,j,0}, \hat{h}_{k,j+1,1} \middle| H_{tk}\right\} = t\left(\Omega_{2k,2k+1}\Omega_{2k+1,2k}\right)\mathbb{E}\left\{\left|h_{k,j,1}\right|^2 \middle| H_{tk}\right\} \\
&+ \sum_{n \neq k} \Omega_{2k,2n+1}\Omega_{2k+1,2n}\mathbb{E}\left\{\left|h_{n,j,1}\right|^2 \middle| H_{tk}\right\}
\end{aligned} \tag{2.124}$$

- Case 6: $j_1 = j$, $j_2 = j + 1$, $f_1 = 1$, and $f_2 = 0$

$$\begin{aligned} \Sigma_{2j'+1,2j'+2}^{1k} &= \mathbb{Cov}\left\{\hat{\mathbf{h}}_{k,j,1}, \hat{\mathbf{h}}_{k,j+1,0} \middle| H_{tk}\right\} = \Omega_{2k,2k}\Omega_{2k+1,2k+1}\mathbb{E}\left\{\left|h_{k,j,1}\right|^2 \middle| H_{tk}\right\} \\ &+ \sum_{n \neq k} \Omega_{2k,2n}\Omega_{2k+1,2n+1}\mathbb{E}\left\{\left|h_{n,j,1}\right|^2 \middle| H_{tk}\right\} \end{aligned} \quad (2.125)$$

By substituting (2.114) and (2.117) into (2.121)-(2.125) and using $\mathbb{E}\left\{w'_{k_1,j_1,f_1}(w'_{k_2,j_2,f_2})^* \middle| H_{tk}\right\} = \Sigma_{2k_1+f_1,2k_2+f_2}^{w'} \delta_{j_1,j_2}$, (2.47) is obtained.

2.E Proof of the Optimum Bayesian's Decision Rule

Based on the Bayesian hypothesis testing, transmission state of the k th IoT device is identified as active, i.e., H_{1k} , if

$$\frac{p\left(\hat{\mathbf{h}}_{k,\bar{\alpha},l} \middle| H_{1k}\right)}{p\left(\hat{\mathbf{h}}_{k,\bar{\alpha},l} \middle| H_{0k}\right)} \geq \frac{1 - P_a}{P_a}. \quad (2.126)$$

By using Lemma 2.3.1, the term on the left-hand side of (2.126) is written as

$$\frac{\det(\pi \Sigma^{0k}) \exp\left(-\hat{\mathbf{h}}_{k,\bar{\alpha},l}^H \left(\Sigma^{1k}\right)^{-1} \hat{\mathbf{h}}_{k,\bar{\alpha},l}\right)}{\det(\pi \Sigma^{1k}) \exp\left(-\hat{\mathbf{h}}_{k,\bar{\alpha},l}^H \left(\Sigma^{0k}\right)^{-1} \hat{\mathbf{h}}_{k,\bar{\alpha},l}\right)}. \quad (2.127)$$

By substituting (2.127) into (2.126), taking the natural logarithms of the both sides of the inequality, and after some simplification, one obtains (2.49) and (2.50).

2.F Proof of the MLR decision rule

Since the elements of the reconstructed vector $\check{\mathbf{h}}_{k,l} \triangleq [\check{h}_{k,0} \check{h}_{k,2} \cdots \check{h}_{k,2(l-1)}]^\dagger$ under hypothesis H_{tk} , $t \in \{0, 1\}$ are i.i.d. zero-mean complex Gaussian random variables

with variance $\Sigma_{2k+f,2k+f}^{tk}$, where $f = 1$ for $\beta_k < \frac{N_c}{2}$, and $f = 0$ for $\beta_k \geq \frac{N_c}{2}$, by employing the MLR test, the transmission state of the k th IoT device is identified as active, i.e., $d_k = H_{1k}$, if

$$\begin{aligned} \frac{p(\check{\mathbf{h}}_{k,l}|H_{1k})}{p(\check{\mathbf{h}}_{k,l}|H_{0k})} &= \frac{\prod_{i=0}^{l-1} p(\check{h}_{k,2i}|H_{1k})}{\prod_{i=0}^{l-1} p(\check{h}_{k,2i}|H_{0k})} \geq \gamma_k \\ &= \frac{(\pi \Sigma_{2k+f,2k+f}^{0k})^l \exp\left(-\frac{\sum_{i=0}^{l-1} |\check{h}_{k,i}|^2}{\Sigma_{2k+f,2k+f}^{1k}}\right)}{(\pi \Sigma_{2k+f,2k+f}^{1k})^l \exp\left(-\frac{\sum_{i=0}^{l-1} |\check{h}_{k,i}|^2}{\Sigma_{2k+f,2k+f}^{0k}}\right)} \geq \gamma_k, \end{aligned} \quad (2.128)$$

where γ_k is a threshold. By taking the natural logarithms of the both sides of the inequality in (2.128), and after some mathematical manipulations, one obtains (2.53) for

$$\theta'_k \triangleq l\sqrt{2} \ln \left(\frac{\gamma_k^{\frac{1}{l}} \Sigma_{2k+f,2k+f}^{1k}}{\Sigma_{2k+f,2k+f}^{0k}} \right) \left(\frac{\Sigma_{2k+f,2k+f}^{1k} - \Sigma_{2k+f,2k+f}^{0k}}{\Sigma_{2k+f,2k+f}^{1k} (\Sigma_{2k+f,2k+f}^{0k})^{\frac{3}{2}}} \right), \quad (2.129)$$

where $f = 1$ for $\beta_k < \frac{N_c}{2}$, and $f = 0$ for $\beta_k \geq \frac{N_c}{2}$.

2.G Proof of the Correct Identification and False Alarm Rates for the ML Squared ℓ_2 norm

According to Lemma 2.3.1, the elements of $\check{\mathbf{h}}_{k,l} = [\check{h}_{k,0} \check{h}_{k,2} \cdots \check{h}_{k,2(l-1)}]^\dagger$ in (2.51) under hypothesis H_{tk} , $t \in \{0, 1\}$ are i.i.d. zero-mean complex Gaussian random variables with variance $\Sigma_{2k+f,2k+f}^{tk}$, where $f = 1$ for $\beta_k < \frac{N_c}{2}$, and $f = 0$ for $\beta_k \geq \frac{N_c}{2}$. Hence, $y_k \triangleq \phi(\check{\mathbf{h}}_{k,l})$ in (2.53) follows central Chi-squared (χ^2) distribution with $2l$ degrees of

freedom as

$$f_{y_k|H_{tk}}\left(y|H_{tk}\right) = \frac{y^{l-1}\exp\left(\frac{-y}{2}\right)}{2^l\Gamma(l)}, \quad (2.130)$$

where $t \in \{0, 1\}$, $f = 1$ for $\beta_k < \frac{N_c}{2}$, and $f = 0$ for $\beta_k \geq \frac{N_c}{2}$.

By employing (2.53), the false alarm rate of the k th IoT device is derived as

$$\begin{aligned} P_k^{(f)} &= \mathbb{P}\left\{d_k = H_{1k}|H_{0k}\right\} \\ &= \mathbb{P}\left\{\frac{\sqrt{2}}{\sqrt{\sum_{2k+f, 2k+f}^{0k}}} \sum_{l=0}^{l-1} |\check{h}_{k,l}|^2 \geq \theta'_k|H_{0k}\right\} \\ &= \int_{\theta'_k}^{+\infty} \frac{y^{l-1}\exp\left(\frac{-y}{2}\right)}{2^l\Gamma(l)} dy = 1 - \frac{\Gamma\left(l, \frac{\theta'_k}{2}\right)}{\Gamma(l)}, \end{aligned} \quad (2.131)$$

where $\Gamma(l) = \int_0^\infty \exp(-t)t^{(l-1)}dt$ and $\Gamma(a, b) = \int_0^b \exp(-t)t^{(a-1)}dt$ are complete and lower incomplete gamma functions, respectively. By applying the inverse upper incomplete gamma function to (2.131), (2.54) is obtained.

Similar to the false alarm rate in (2.131), by using (2.130) for $t = 1$, the correct identification rate of the k th IoT device is obtained as

$$\begin{aligned} P_k^{(d)} &= \mathbb{P}\left\{d_k = H_{1k}|H_{1k}\right\} \\ &= \mathbb{P}\left\{\frac{\sqrt{2}}{\sqrt{\sum_{2k+f, 2k+f}^{0k}}} \sum_{l=0}^{l-1} |\check{h}_{k,l}|^2 \geq \theta'_k|H_{1k}\right\} \\ &= \mathbb{P}\left\{\frac{\sqrt{2}}{\sqrt{\sum_{2k+f, 2k+f}^{1k}}} \geq \sqrt{\frac{\sum_{2k+f, 2k+f}^{0k}}{\sum_{2k+f, 2k+f}^{1k}}} \theta'_k|H_{1k}\right\} \\ &= \int_{\theta''_k}^{+\infty} \frac{y^{l-1}\exp\left(\frac{-y}{2}\right)}{2^l\Gamma(l)} dy = 1 - \frac{\Gamma\left(l, \sqrt{\frac{\sum_{2k+f, 2k+f}^{0k}}{4\sum_{2k+f, 2k+f}^{1k}}} \theta'_k\right)}{\Gamma(l)}. \end{aligned} \quad (2.132)$$

2.H Approximation of the Tuning Parameter λ_{ℓ_1} for the ℓ_1 norm DSA IoT Identification Algorithm

By replacing $\|\mathbf{h}_{k,j}\|_0$ with $\|\mathbf{h}_{k,j}\|_1/|g_k|$, one can write

$$\lambda_{\ell_0} \|\mathbf{h}_j\|_0 = \lambda_{\ell_0} \sum_{k=0}^{K_u-1} \|\mathbf{h}_{k,j}\|_0 = \lambda_{\ell_0} \sum_{k=0}^{K_u-1} \frac{\|\mathbf{h}_{k,j}\|_1}{|g_k|} \quad (2.133)$$

In order to have the solution of (2.56) the same as the one of the original SSR problem in (2.26), we need to have

$$\lambda_{\ell_1} \|\mathbf{h}_j\|_1 = \lambda_{\ell_0} \|\mathbf{h}_j\|_0 = \lambda_{\ell_0} \sum_{k=0}^{K_u-1} \frac{\|\mathbf{h}_{k,j}\|_1}{|g_k|}. \quad (2.134)$$

Since $g_k \sim \mathcal{N}_c(0, p_k)$, $k \in \mathcal{X}_u$, $|g_k|$ follows Rayleigh distribution. By replacing $|g_k|$ with $\mathbb{E}\{|g_k|\} = \sqrt{\pi p_k}/2$ in (2.134), one can write

$$\lambda_{\ell_1} \|\mathbf{h}_j\|_1 \approx \frac{2\lambda_{\ell_0}}{\sqrt{\pi}} \sum_{k=0}^{K_u-1} \frac{\|\mathbf{h}_{k,j}\|_1}{\sqrt{p_k}}. \quad (2.135)$$

Further, we have

$$\|\mathbf{r}_j - \mathbf{X}\mathbf{h}_j\|_F = \|\mathbf{r}_j - \ddot{\mathbf{X}}\ddot{\mathbf{h}}_j\|_F \quad (2.136)$$

where $\ddot{\mathbf{X}}$ and $\ddot{\mathbf{h}}_j$ are given in (2.57a) and (2.57b). By substituting (2.135) and (2.140) into (2.56), one obtains (2.59) and (2.60).

2.I Approximation of the Tuning Parameter λ_{ℓ_1} for the ℓ_1 norm DSA IoT Identification algorithm

By replacing $\|\mathbf{h}_{k,j}\|_0$ with $\|\mathbf{h}_{k,j}\|_1/|g_k|$, one can write

$$\lambda_{\ell_0} \|\mathbf{h}_j\|_0 = \lambda_{\ell_0} \sum_{k=0}^{K_u-1} \|\mathbf{h}_{k,j}\|_0 = \lambda_{\ell_0} \sum_{k=0}^{K_u-1} \frac{\|\mathbf{h}_{k,j}\|_1}{|g_k|} \quad (2.137)$$

In order to have the solution of (2.56) the same as the one of the original SSR problem in (2.26), we need to have

$$\lambda_{\ell_1} \|\mathbf{h}_j\|_1 = \lambda_{\ell_0} \|\mathbf{h}_j\|_0 = \lambda_{\ell_0} \sum_{k=0}^{K_u-1} \frac{\|\mathbf{h}_{k,j}\|_1}{|g_k|}. \quad (2.138)$$

Since $g_k \sim \mathcal{N}_c(0, p_k)$, $k \in \mathcal{X}_u$, $|g_k|$ follows Rayleigh distribution. By replacing $|g_k|$ with $\mathbb{E}\{|g_k|\} = \sqrt{\pi p_k}/2$ in (2.137), one can write

$$\lambda_{\ell_1} \|\mathbf{h}_j\|_1 \approx \frac{2\lambda_{\ell_0}}{\sqrt{\pi}} \sum_{k=0}^{K_u-1} \frac{\|\mathbf{h}_{k,j}\|_1}{\sqrt{p_k}}. \quad (2.139)$$

Further, we have

$$\|\mathbf{r}_j - \mathbf{X}\mathbf{h}_j\|_F = \|\mathbf{r}_j - \ddot{\mathbf{X}}\ddot{\mathbf{h}}_j\|_F \quad (2.140)$$

where $\ddot{\mathbf{X}}$ and $\ddot{\mathbf{h}}_j$ are given in (2.57a) and (2.57b). By substituting (2.139) and (2.140) into (2.56), one obtains (2.59) and (2.60).

2.J Proof of the KKT Optimality Conditions for the $\ell_1 - \ell_2$ Mixed-norm PDSA IoT Identification Algorithms

For a convex function f from a complex matrix space \mathbb{M} to \mathbb{R} , the subdifferential of f at matrix \mathbf{A} is defined as [98, 99]

$$\begin{aligned} \partial f(\mathbf{A}) \triangleq & \left\{ \mathbf{G} \in \mathbb{M} : f(\mathbf{Z}) \geq f(\mathbf{A}) \right. \\ & \left. + \operatorname{Re}\{ \langle \mathbf{Z} - \mathbf{A}, \mathbf{G} \rangle \} \forall \mathbf{G} \in \mathbb{M} \right\}. \end{aligned} \quad (2.141)$$

The elements of $\partial f(\mathbf{A})$ are called subgradients of the function f at \mathbf{A} . In [99], it is shown that for a convex and differentiable function f , $\partial f(\mathbf{A}) = \nabla f(\mathbf{A})$. If f is a convex function, then \mathbf{A} is the maximizer of f if and only if $\mathbf{0} \in \partial f(\mathbf{A})$.

By employing (2.141), the KKT optimality conditions of the optimization problem of PDSA IoT identification are obtained as

$$\mathbf{0} \in \frac{1}{2} \partial_{\check{\mathbf{h}}_{k,\bar{\alpha},l}} \left\| \mathbf{R} - \check{\mathbf{X}} \check{\mathbf{H}} \right\|_{\mathbb{F}}^2 + \partial_{\check{\mathbf{h}}_{k,\bar{\alpha},l}} \sum_{k=0}^{K_u-1} \left\| \check{\mathbf{h}}_{k,\bar{\alpha},1} \right\|_2. \quad (2.142)$$

Because $\left\| \mathbf{R} - \check{\mathbf{X}} \check{\mathbf{H}} \right\|_{\mathbb{F}}^2$ is convex and differentiable, one can write

$$\frac{1}{2} \partial_{\check{\mathbf{h}}_{k,\bar{\alpha},l}} \left\| \mathbf{R} - \check{\mathbf{X}} \check{\mathbf{H}} \right\|_{\mathbb{F}}^2 = \frac{1}{2} \nabla_{\check{\mathbf{h}}_{k,\bar{\alpha},l}} \left\| \bar{\mathbf{R}} - \check{\mathbf{X}} \check{\mathbf{H}} \right\|_{\mathbb{F}}^2 = \check{\Psi}_k, \quad (2.143)$$

where $\check{\Psi}_k \triangleq \left[\check{\Psi}_{k,0,0} \ \check{\Psi}_{k,0,1} \ \cdots \ \check{\Psi}_{k,l-1,0} \ \check{\Psi}_{k,l-1,1} \right]^\dagger$, $\left[\check{\Psi}_{k,0,f} \ \check{\Psi}_{k,1,f} \ \cdots \ \check{\Psi}_{k,l-1,f} \right] \triangleq \check{\mathbf{x}}_{k,f}^\dagger (\mathbf{R} - \check{\mathbf{X}} \check{\mathbf{H}})$, with $f \in \{0, 1\}$, $\check{\mathbf{x}}_{k,f} = p_k^{-1/2} \mathbf{x}_{k,f}$, and $\mathbf{x}_{k,1}$ and $\mathbf{x}_{k,0}$ are given in (2.11). Similar to (2.143), by using $\partial_{\check{\mathbf{h}}_{k,\bar{\alpha},l}} \left\| \check{\mathbf{h}}_{k,\bar{\alpha},1} \right\|_2 = \nabla_{\check{\mathbf{h}}_{k,\bar{\alpha},l}} \left\| \check{\mathbf{h}}_{k,\bar{\alpha},1} \right\|_2$ for $\check{\mathbf{h}}_{k,\bar{\alpha},l} \neq \mathbf{0}$ and by employing

the definition of the subdifferential in (2.141) for $\ddot{\mathbf{h}}_{k,\bar{\alpha},l} = \mathbf{0}$, one obtains

$$\partial_{\ddot{\mathbf{h}}_{k,\bar{\alpha},l}} \|\ddot{\mathbf{h}}_{k,\bar{\alpha},l}\|_2 = \begin{cases} \frac{\ddot{\mathbf{h}}_{k,\bar{\alpha},l}}{\|\ddot{\mathbf{h}}_{k,\bar{\alpha},l}\|_2} & \text{for } \ddot{\mathbf{h}}_{k,\bar{\alpha},l} \neq \mathbf{0} \\ \{\boldsymbol{\alpha} \in \mathbb{C}^{2l} \mid \|\boldsymbol{\alpha}\|_2 \leq 1\} & \text{for } \ddot{\mathbf{h}}_{k,\bar{\alpha},l} = \mathbf{0} \end{cases}, \quad (2.144)$$

Finally, by employing (2.143) and (2.144), and then (2.142), the KKT optimality conditions are obtained as

$$-\ddot{\Psi}_k + \lambda \frac{\ddot{\mathbf{h}}_{k,\bar{\alpha},l}}{\|\ddot{\mathbf{h}}_{k,\bar{\alpha},l}\|_2} = \mathbf{0} \quad \text{if } k \in \mathcal{X}_a \quad (2.145a)$$

$$\|\ddot{\Psi}_k\|_2 < \lambda \quad \text{if } k \notin \mathcal{X}_a. \quad (2.145b)$$

2.K Equivalence of $\ell_1 - \ell_2$ Mixed-norm Algorithm with Iterative Reweighed LS Estimator

Let us consider a random matrix \mathbf{A} with the same size as \mathbf{H} . It can be easily show that $\hat{\mathbf{H}}^{[t+1]}$ in (2.78) can be written as $\hat{\mathbf{H}}^{[t+1]} = \mathbf{F}^{[t]} \hat{\mathbf{A}}^{[t+1]}$ where $\hat{\mathbf{A}}^{[t+1]}$ is obtained by

$$\hat{\mathbf{A}}^{[t+1]} = \arg \min_{\mathbf{A}} \frac{1}{2} \|\bar{\mathbf{R}} - \mathbf{X} \mathbf{F}^{[t]} \mathbf{A}\|_{\mathbf{F}}^2 + \lambda_{\ell_2}^{\ell_1} \sum_{k=0}^{K_u-1} \|\mathbf{a}_{k,\bar{\alpha},l}\|_2^2. \quad (2.146)$$

Since $\hat{\mathbf{A}}^{[t+1]} = \left(\mathbf{F}^{[t]}\right)^{-1} \hat{\mathbf{H}}^{[t+1]}$, one can write

$$\begin{aligned} \hat{\mathbf{H}}^{[t+1]} = \underset{\mathbf{H}}{\operatorname{argmin}} \quad & \frac{1}{2} \|\bar{\mathbf{R}} - \mathbf{X} \mathbf{F}^{[t]} \left(\mathbf{F}^{[t]}\right)^{-1} \mathbf{H}\|_{\mathbf{F}}^2 \\ & + \lambda_{\ell_2}^{\ell_1} \sum_{k=0}^{K_u-1} \left\| \left(\mathbf{F}^{[t]}\right)^{-1} \mathbf{h}_{k,\bar{\alpha},l} \right\|_2^2. \end{aligned} \quad (2.147)$$

As seen, equation (2.147) represents iterative reweighed LS estimator.

Chapter 3

Doppler Spread Estimation in MIMO Frequency-selective Fading Channels

3.1 Introduction

One of the main challenges in high-speed mobile communications is the presence of large Doppler spreads. Thus, accurate estimation of maximum Doppler spread (MDS) plays an important role in such systems.

Maximum Doppler spread measures the coherence time, related to the rate of change, of wireless communication channels. Its knowledge is important to design efficient wireless communication systems for high-speed vehicles [20, 21, 100]. In particular, accurate estimation of the MDS is required for the design of adaptive transceivers, as well as in cellular and smart antenna systems [22, 23, 100–107]. For example, in the context of adaptive transceivers, system parameters such as coding, modulation, and power are adapted to the changes in the channel [101–104]. In cellular

systems, handoff is dictated by the velocity of the mobile station, which is also directly obtained from the Doppler information. Knowledge of the rate of the channel change is also employed to reduce unnecessary handoff; the handoff is initiated based on the received power at the mobile station, and the optimum window size for power estimation depends on the MDS [22, 23, 104, 105]. In the context of smart antenna systems, the MDS is used in the design of the maximum likelihood (ML) space-time transceivers [106, 107]. In addition, knowledge of MDS is required for channel tracking and equalization, as well as for the selection of the optimal interleaving length in wireless communication systems [108].

3.1.1 Literature Review

In general, parameter estimators can be categorized as: i) data-aided (DA), where the estimation relies on a pilot or preamble sequence [109–113] and ii) non-data-aided (NDA), where the estimation is performed with no *a priori* knowledge about the transmitted symbols [114–117].

With regard to the MDS estimation, the DA approach often provides accurate estimates for slowly-varying channels by employing a reduced number of pilot symbols, whereas this does not hold for fast-varying channels. In the latter case, the details of the channel variations cannot be captured accurately, and more pilots are required, which results in increased overhead and reduced system capacity.

There are five major classes of MDS estimators: ML-based, power spectral density (PSD)-based, level-crossing-rate (LCR)-based, covariance-based, and cyclostationarity-based estimators. The ML-based estimator maximizes the likelihood function, and, in general, is asymptotically unbiased, achieving the Cramer-Rao lower bound (CRLB) [118, 119]. However, maximum likelihood estimator (MLE) for MDS suffers from significant computational complexity. Hence, different modified low-complexity MLEs

for MDS in single-input single-output (SISO) flat-fading channel developed [1, 2]. With the PSD-based estimators, some unique features from the Doppler spectrum are obtained through the sample periodogram of the received signal [120]. Covariance-based estimators extract the Doppler information which exists in the sample auto-covariance of the received signal [3, 121, 122]. LCR-based estimators rely on the number of level crossings of the received signal statistics, which is proportional to the MDS [123]. The cyclostationarity-based estimators exploit the cyclostationarity of the received signal [4]. Comparing with other MDS estimators, the advantage of the cyclostationarity-based estimators is the robustness to stationary noise and interference.

While the problem of MDS estimation in SISO flat-fading channel has been extensively investigated in the literature [1–4, 118–124], the MDS estimation in multiple-input multiple-output (MIMO) frequency-selective or in MIMO flat-fading channel has not been considerably explored. Furthermore, DA-MDS estimation has mainly been studied in the literature. To the best of our knowledge, only a few works have addressed MDS estimation in conjunction with multiple antenna systems. In [119], the authors derived an asymptotic DA-MLE and DA-CRLB for joint MDS and noise variance estimation in MIMO flat-fading channel. In [4], the cyclic correlation (CC) of linearly modulated signals is exploited for the MDS estimation for single transmit antenna scenarios. While both DA and NDA estimators are studied in [4], only frequency-flat fading and single transmit antenna are considered.

3.1.2 Motivation

After reviewing the current MDS estimators in the literature, the followings were made:

- While the problem of DA-CRLB for MDS estimation in MIMO flat-fading

channel has been studied in the literature, the DA-CRLB in MIMO frequency-selective fading channel has not been investigated;

- The problem of NDA-CRLB for MDS estimation has not been investigated in the literature (neither in MIMO flat-fading nor frequency-selective channel);
- The DA-MLE for MDS in MIMO frequency-selective channel has not been investigated either;
- The NDA-MLEs for MDS in MIMO flat-fading and frequency-selective channel have not been explored;
- The existing DA and NDA MDS estimators suffer from huge computational complexity;
- The current MDS estimators require joint parameter estimation, such as carrier frequency offset (CFO), signal power, noise power, and channel delay profile estimation.

3.1.3 Problem Statement

The specific research problems which are studied in Chapter 3 of this thesis are presented as follows:

- The DA- and NDA-CRLBs for MDS estimation in MIMO frequency-selective fading channel are derived;
- The DA- and NDA-MLEs for MDS in MIMO frequency-selective fading channel are derived;
- A new low-complexity NDA-moment-based estimator (MBE) for MDS in MIMO frequency-selective channel is proposed;

- The optimal combining method for the proposed NDA-MBE in case of multiple receive antennas is derived through the bootstrap technique.

3.1.4 Methodology

The proposed DA-CRLB for MDS estimation and DA-MLE in MIMO frequency-selective channel are derived through the conditional probability density function (PDF) of the received vector given the known transmitted symbols. In this case, the Fisher information and the conditional log-likelihood function is obtained from the conditional PDF for the DA-CRLB and DA-MLE, respectively.

For the NDA-CRLB and NDA-MLE, it is assumed that the transmitted symbols are unknown, but the alphabet from which they are drawn is known. In this case, through applying the marginal PDF to the joint PDF of the received vector and constellation vector, the Fisher information and the log-likelihood function are obtained.

The proposed NDA-MBE relies on the statistical moment-based approach to estimate the normalized squared autocorrelation function of the fading channel. Then, the problem of MDS estimation is formulated as a non-linear regression problem, and the least-squares curve-fitting optimization technique is applied to determine the estimate of the MDS.

3.2 Maximum Doppler Spread (MDS) Estimation

3.2.1 System Model

Let us consider a MIMO wireless communication system with n_t transmit antennas and n_r receive antennas, where the received signals are affected by time-varying frequency-selective Rayleigh fading and are corrupted by additive white Gaussian

noise. The discrete-time complex-valued baseband signal at the n th receive antenna is expressed as [125]

$$\mathbf{r}_k^{(n)} = \sum_{m=1}^{n_t} \sum_{l=1}^L h_{k,l}^{(mn)} s_{k-l}^{(m)} + w_k^{(n)} \quad k = 1, \dots, N, \quad (3.1)$$

where N is the number of observation symbols, L is the length of the channel impulse response, $s_k^{(m)}$ is the symbol transmitted from the m th antenna at time k , satisfying $\mathbb{E}\{s_{k_1}^{(m_1)}(s_{k_2}^{(m_2)})^*\} = \sigma_{s_{m_1}}^2 \delta_{m_1, m_2} \delta_{k_1, k_2}$, with $\sigma_{s_{m_1}}^2$ being the transmit power of the m_1 th antenna, $w_k^{(n)}$ is the complex-valued additive white Gaussian noise at the n th receive antenna at time k , whose variance is $\sigma_{w_n}^2$, and $h_{k,l}^{(mn)}$ denotes the zero-mean complex-valued Gaussian fading process between the m th transmit and n th receive antennas for the l th tap of the fading channel and at time k . It is considered that the channels for different antennas are independent, with the cross-correlation of the l_1 and l_2 taps given by¹

$$\mathbb{E}\{h_{k,l_1}^{(mn)}(h_{k+u,l_2}^{(mn)})^*\} = \sigma_{h_{(mn),l_1}}^2 J_0(2\pi f_D T_s u) \delta_{l_1, l_2}, \quad (3.2)$$

where $J_0(\cdot)$ is the zero-order Bessel function of the first kind, $\sigma_{h_{(mn),l_1}}^2$ is the variance of the l_1 th tap between the m th transmit and n th receive antennas, T_s denotes the symbol period, and $f_D = v/\lambda = f_c v/c$ represents the MDS in Hz, with v as the relative speed between the transmitter and receiver, λ as the wavelength, f_c as the carrier frequency, and c as the speed of light.

3.2.2 CRLB for MDS Estimation

In this section, the DA- and NDA-CRLB for MDS estimation in MIMO frequency-selective fading channel are derived.

¹Here we consider the Jakes channel; it is worth noting that different parametric channel models can be also considered.

3.2.2.1 DA-CRLB

Let us consider $s_k^{(m)} = s_k^{(m)}$, $m = 1, 2, \dots, n_t$, $k = 1, 2, \dots, N - L + 1$, as employed pilots for DA-MDS estimation. The received signal at n th receive antenna in (3.1) can be written as

$$\begin{aligned} \mathbf{r}_k^{(n)} = \bar{r}_k^{(n)} + j\check{r}_k^{(n)} &= \sum_{m=1}^{n_t} \sum_{l=1}^L \bar{h}_{k,l}^{(mn)} \bar{s}_{k-l}^{(m)} - \check{h}_{k,l}^{(mn)} \check{s}_{k-l}^{(m)} + \bar{w}_k^{(n)} \\ &+ j \left(\sum_{m=1}^{n_t} \sum_{l=1}^L \bar{h}_{k,l}^{(mn)} \check{s}_{k-l}^{(m)} + \check{h}_{k,l}^{(mn)} \bar{s}_{k-l}^{(m)} + \check{w}_k^{(n)} \right), \end{aligned} \quad (3.3)$$

where $\bar{r}_k^{(n)} \triangleq \text{Re}\{r_k^{(n)}\}$, $\check{r}_k^{(n)} \triangleq \text{Im}\{r_k^{(n)}\}$, $\bar{h}_{k,l}^{(mn)} \triangleq \text{Re}\{h_{k,l}^{(mn)}\}$, $\check{h}_{k,l}^{(mn)} \triangleq \text{Im}\{h_{k,l}^{(mn)}\}$, $\bar{s}_{k-l}^{(m)} \triangleq \text{Re}\{s_{k-l}^{(m)}\}$, and $\check{s}_{k-l}^{(m)} \triangleq \text{Im}\{s_{k-l}^{(m)}\}$.

Let us define

$$\mathbf{r}^{(n)} \triangleq \left[\bar{r}_1^{(n)} \ \bar{r}_2^{(n)} \ \dots \ \bar{r}_N^{(n)} \ \check{r}_1^{(n)} \ \check{r}_2^{(n)} \ \dots \ \check{r}_N^{(n)} \right]^\dagger \quad (3.4)$$

and

$$\mathbf{r} \triangleq \left[\mathbf{r}^{(1)\dagger} \ \mathbf{r}^{(2)\dagger} \ \dots \ \mathbf{r}^{(n_t)\dagger} \right]^\dagger. \quad (3.5)$$

The elements of the vector $\mathbf{r}^{(n)}$, $n = 1, 2, \dots, n_t$, are linear combinations of the correlated Gaussian random variables as in (3.3). Thus, \mathbf{r} , is a Gaussian random vector with PDF given by

$$p(\mathbf{r}|\mathbf{s}; \boldsymbol{\theta}) = \frac{\exp\left(-\frac{1}{2}\mathbf{r}^\dagger \boldsymbol{\Sigma}^{-1}(\mathbf{s}, \boldsymbol{\theta})\mathbf{r}\right)}{(2\pi)^{Nn_r} \det^{\frac{1}{2}}\left(\boldsymbol{\Sigma}(\mathbf{s}, \boldsymbol{\theta})\right)}, \quad (3.6)$$

where $\boldsymbol{\Sigma}(\mathbf{s}, \boldsymbol{\theta}) \triangleq \mathbb{E}\{\mathbf{r}\mathbf{r}^\dagger\}$, $\mathbf{s} \triangleq \left[\mathbf{s}^{(1)\dagger} \ \mathbf{s}^{(2)\dagger} \ \dots \ \mathbf{s}^{(n_t)\dagger} \right]^\dagger$, $\mathbf{s}^{(m)} \triangleq \left[\bar{s}_1^{(m)} \ \bar{s}_2^{(m)} \ \dots \ \bar{s}_{N-L+1}^{(m)} \ \check{s}_1^{(m)} \ \check{s}_2^{(m)} \ \dots \ \check{s}_{N-L+1}^{(m)} \right]^\dagger$, and $\boldsymbol{\theta} \triangleq \left[\boldsymbol{\xi} \ \boldsymbol{\vartheta} \ f_D \right]^\dagger$ is the parameter vector, as

$$\boldsymbol{\xi} \triangleq [\sigma_{w_1}^2 \ \cdots \ \sigma_{w_{n_r}}^2]^\dagger \quad (3.7a)$$

$$\boldsymbol{\vartheta} \triangleq [\boldsymbol{\vartheta}_1^\dagger \ \boldsymbol{\vartheta}_2^\dagger \ \cdots \ \boldsymbol{\vartheta}_L^\dagger]^\dagger \quad (3.7b)$$

$$\boldsymbol{\vartheta}_l \triangleq \begin{bmatrix} \sigma_{h_{(11),l}}^2 & \cdots & \sigma_{h_{(1n_r),l}}^2 & \sigma_{h_{(21),l}}^2 & \cdots \\ \sigma_{h_{(2n_r),l}}^2 & \cdots & \sigma_{h_{(n_t 1),l}}^2 & \cdots & \sigma_{h_{(n_t n_r),l}}^2 \end{bmatrix}^\dagger. \quad (3.7c)$$

Since $\mathbf{r}^{(n_1)}$ and $\mathbf{r}^{(n_2)}$, $n_1 \neq n_2$, are uncorrelated random vectors, i.e. $\mathbb{E}\{\mathbf{r}^{(n_1)}\mathbf{r}^{(n_2)\dagger}\} = \mathbf{0}$, the covariance matrix of \mathbf{r} , $\boldsymbol{\Sigma}(\mathbf{s}, \boldsymbol{\theta})$, is block diagonal as

$$\boldsymbol{\Sigma}(\mathbf{s}, \boldsymbol{\theta}) \triangleq \mathbb{E}\{\mathbf{r}\mathbf{r}^\dagger\} = \begin{bmatrix} \boldsymbol{\Sigma}^{(1)} & & & & \\ & \boldsymbol{\Sigma}^{(2)} & & & \\ & & \ddots & & \\ & & & & \boldsymbol{\Sigma}^{(n_r)} \end{bmatrix}, \quad (3.8)$$

where $\boldsymbol{\Sigma}^{(n)} \triangleq \mathbb{E}\{\mathbf{r}^{(n)}\mathbf{r}^{(n)\dagger}\}$. By employing (3.2), (3.3), and (3.4), using the fact the real and imaginary part of the fading tap are independent random variables with $\mathbb{E}\{|\bar{h}_{k,l}^{(mn)}|^2\} = \{|\check{h}_{k,l}^{(mn)}|^2\} = \sigma_{h_{(mn),l}}^2/2$, and after some algebra, the elements of the covariance matrix $\boldsymbol{\Sigma}^{(n)}$, $n \in \{1, 2, \dots, n_r\}$, are obtained as

$$\begin{aligned} \mathbb{E}\left\{\bar{r}_k^{(n)}\bar{r}_{k+u}^{(n)}\right\} &= \mathbb{E}\left\{\check{r}_k^{(n)}\check{r}_{k+u}^{(n)}\right\} \\ &= \frac{1}{2} \sum_{m=1}^{n_t} \sum_{l=1}^L \sigma_{h_{(mn),l}}^2 \left(\bar{s}_{k-l}^{(m)}\bar{s}_{k+u-l}^{(m)} + \check{s}_{k-l}^{(m)}\check{s}_{k+u-l}^{(m)} \right) J_0(2\pi f_D T_s u) + \frac{\sigma_{w_n}^2}{2} \delta_{u,0} \end{aligned} \quad (3.9a)$$

$$\begin{aligned} \mathbb{E}\left\{\bar{r}_k^{(n)}\check{r}_{k+u}^{(n)}\right\} &= -\mathbb{E}\left\{\check{r}_k^{(n)}\bar{r}_{k+u}^{(n)}\right\} \\ &= \frac{1}{2} \sum_{m=1}^{n_t} \sum_{l=1}^L \sigma_{h_{(mn),l}}^2 \left(\bar{s}_{k-l}^{(m)}\check{s}_{k+u-l}^{(m)} - \check{s}_{k-l}^{(m)}\bar{s}_{k+u-l}^{(m)} \right) J_0(2\pi f_D T_s u). \end{aligned} \quad (3.9b)$$

The Fisher information matrix of the parameter vector $\boldsymbol{\theta}$, $\mathbf{I}(\boldsymbol{\theta})$, for the zero-mean Gaussian observation vector in (3.6) is obtained as

$$[\mathbf{I}(\boldsymbol{\theta})]_{ij} \triangleq -\mathbb{E}\left\{\frac{\partial^2 \ln p(\mathbf{r}|\mathbf{s}; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}_i \partial \boldsymbol{\theta}_j}\right\} = \frac{1}{2} \text{tr} \left[\boldsymbol{\Sigma}^{-1}(\mathbf{s}, \boldsymbol{\theta}) \frac{\partial \boldsymbol{\Sigma}(\mathbf{s}, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}_i} \boldsymbol{\Sigma}^{-1}(\mathbf{s}, \boldsymbol{\theta}) \frac{\partial \boldsymbol{\Sigma}(\mathbf{s}, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}_j} \right]. \quad (3.10)$$

For the MDS, f_D , $I(f_D) \triangleq [\mathbf{I}(\boldsymbol{\theta})]_{xx}$, $x = n_t n_r L + n_r + 1$, and one obtains

$$I(f_D) = -\mathbb{E}\left\{\frac{\partial^2 \ln p(\mathbf{r}|\mathbf{s}; \boldsymbol{\theta})}{\partial f_D^2}\right\} = \frac{1}{2} \text{tr} \left[\left(\boldsymbol{\Sigma}^{-1}(\mathbf{s}, \boldsymbol{\theta}) \frac{\partial \boldsymbol{\Sigma}(\mathbf{s}, \boldsymbol{\theta})}{\partial f_D} \right)^2 \right], \quad (3.11)$$

where $\frac{\partial \boldsymbol{\Sigma}(\mathbf{s}, \boldsymbol{\theta})}{\partial f_D}$ is obtained by replacing $J_0(2\pi f_D T_s u)$ with $-2\pi u T_s J_1(2\pi f_D T_s u)$ in $\boldsymbol{\Sigma}(\mathbf{s}, \boldsymbol{\theta})$, where $J_1(\cdot)$ is the Bessel function of the first kind.

Finally, by employing (3.11), the DA-CRLB for MDS estimation in MIMO frequency-selective fading channel is obtained as

$$\text{Var}(\hat{f}_D) \geq I^{-1}(f_D) = \frac{1}{\frac{1}{2} \text{tr} \left[\left(\boldsymbol{\Sigma}^{-1}(\mathbf{s}, \boldsymbol{\theta}) \frac{\partial \boldsymbol{\Sigma}(\mathbf{s}, \boldsymbol{\theta})}{\partial f_D} \right)^2 \right]}. \quad (3.12)$$

3.2.2.2 NDA-CRLB

Let us consider that the symbols transmitted by each antenna are selected from a constellation with elements $\{c_1 \ c_2 \ \dots \ c_{|M|}\}$, where $\frac{1}{|M|} \sum_{i=1}^{|M|} |c_i|^2 = 1$. The PDF of the received vector \mathbf{r} for NDA-MDS estimation is expressed as

$$p(\mathbf{r}; \boldsymbol{\varphi}) = \sum_{\mathbf{c}} p(\mathbf{r}, \mathbf{c}; \boldsymbol{\varphi}), \quad (3.13)$$

where \mathbf{c} is the constellation vector as $\mathbf{c} \triangleq [\mathbf{c}^{(1)\dagger} \ \mathbf{c}^{(2)\dagger} \ \dots \ \mathbf{c}^{(n_t)\dagger}]^\dagger$, $\mathbf{c}^{(m)} \triangleq [\bar{\mathbf{c}}_{1-L}^{(m)} \ \bar{\mathbf{c}}_{2-L}^{(m)} \ \dots \ \bar{\mathbf{c}}_{N-1}^{(m)} \ \check{\mathbf{c}}_{1-L}^{(m)} \ \check{\mathbf{c}}_{2-L}^{(m)} \ \dots \ \check{\mathbf{c}}_{N-1}^{(m)}]^\dagger$, $\mathbf{c}_k^{(m)} = \bar{\mathbf{c}}_k^{(m)} + j\check{\mathbf{c}}_k^{(m)}$ is the constellation point of the m th transmit

antenna at time k , and $\boldsymbol{\varphi} \triangleq [\boldsymbol{\beta}^\dagger \boldsymbol{\xi}^\dagger \boldsymbol{\vartheta}^\dagger f_D]^\dagger$ with $\boldsymbol{\beta} \triangleq [\sigma_{s_1}^2 \sigma_{s_2}^2 \cdots \sigma_{s_{n_t}}^2]^\dagger$, and $\boldsymbol{\xi}$ and $\boldsymbol{\vartheta}$ are given in (3.7).

By employing the chain rule of probability and using $p(\mathbf{c} = \mathbf{c}_{\langle i \rangle}) = |M|^{-N' n_t}$, $N' \triangleq N + L - 1$, one can write (3.13) as

$$\begin{aligned} p(\mathbf{r}; \boldsymbol{\varphi}) &= \sum_{\mathbf{c}} p(\mathbf{r}, \mathbf{c}; \boldsymbol{\varphi}) = \sum_{\mathbf{c}} p(\mathbf{c} = \mathbf{c}) p(\mathbf{r} | \mathbf{c} = \mathbf{c}; \boldsymbol{\varphi}) \\ &= \frac{1}{|M|^{N' n_t}} \sum_{i=1}^{|M|^{N' n_t}} p(\mathbf{r} | \mathbf{c} = \mathbf{c}_{\langle i \rangle}; \boldsymbol{\varphi}), \end{aligned} \quad (3.14)$$

where $\mathbf{c}_{\langle i \rangle}$ represents the i th possible constellation vector at the transmit-side.

Similar to the DA-CRLB, $p(\mathbf{r} | \mathbf{c} = \mathbf{c}_{\langle i \rangle}; \boldsymbol{\varphi})$ is Gaussian and

$$p(\mathbf{r} | \mathbf{c} = \mathbf{c}_{\langle i \rangle}; \boldsymbol{\varphi}) = \frac{\exp\left(-\frac{1}{2} \mathbf{r}^\dagger \boldsymbol{\Sigma}^{-1}(\mathbf{c}_{\langle i \rangle}, \boldsymbol{\varphi}) \mathbf{r}\right)}{(2\pi)^{N n_r} \det^{\frac{1}{2}}\left(\boldsymbol{\Sigma}(\mathbf{c}_{\langle i \rangle}, \boldsymbol{\varphi})\right)}, \quad (3.15)$$

where $\boldsymbol{\Sigma}(\mathbf{c}_{\langle i \rangle}, \boldsymbol{\varphi}) \triangleq \mathbb{E}\left\{\mathbf{r}_{\langle i \rangle} \mathbf{r}_{\langle i \rangle}^\dagger\right\}$ is the covariance matrix of the received vector $\mathbf{r}_{\langle i \rangle}$ given the constellation vector is $\mathbf{c} = \mathbf{c}_{\langle i \rangle}$, $i = 1, 2, \dots, |M|^{N' n_t}$. The $2N n_r \times 2N n_r$ covariance matrix $\boldsymbol{\Sigma}(\mathbf{c}_{\langle i \rangle}, \boldsymbol{\varphi})$ is block diagonal as in (3.8), where its diagonal elements, i.e., $\boldsymbol{\Sigma}_{\langle i \rangle}^{(n)} \triangleq \mathbb{E}\left\{\mathbf{r}_{\langle i \rangle}^{(n)} \mathbf{r}_{\langle i \rangle}^{(n)\dagger}\right\}$, $n \in \{1, 2, \dots, n_r\}$, are obtained as

$$\begin{aligned} \mathbb{E}\left\{\bar{r}_{k, \langle i \rangle}^{(n)} \bar{r}_{k+u, \langle i \rangle}^{(n)}\right\} &= \mathbb{E}\left\{\check{r}_{k, \langle i \rangle}^{(n)} \check{r}_{k+u, \langle i \rangle}^{(n)}\right\} \\ &= \frac{1}{2} \sum_{m=1}^{n_t} \sum_{l=1}^L \sigma_{h_{(mn), l}}^2 \sigma_{s_m}^2 \left(\bar{c}_{k-l, \langle i \rangle}^{(m)} \bar{c}_{k+u-l, \langle i \rangle}^{(m)} + \check{c}_{k-l, \langle i \rangle}^{(m)} \check{c}_{k+u-l, \langle i \rangle}^{(m)}\right) J_0(2\pi f_D T_s u) + \frac{\sigma_{w_n}^2}{2} \delta_{u,0} \end{aligned} \quad (3.16a)$$

$$\begin{aligned} \mathbb{E}\left\{\bar{r}_{k, \langle i \rangle}^{(n)} \check{r}_{k+u, \langle i \rangle}^{(n)}\right\} &= -\mathbb{E}\left\{\check{r}_{k, \langle i \rangle}^{(n)} \bar{r}_{k+u, \langle i \rangle}^{(n)}\right\} \\ &= \frac{1}{2} \sum_{m=1}^{n_t} \sum_{l=1}^L \sigma_{h_{(mn), l}}^2 \sigma_{s_m}^2 \left(\bar{c}_{k-l, \langle i \rangle}^{(m)} \check{c}_{k+u-l, \langle i \rangle}^{(m)} - \check{c}_{k-l, \langle i \rangle}^{(m)} \bar{c}_{k+u-l, \langle i \rangle}^{(m)}\right) J_0(2\pi f_D T_s u). \end{aligned} \quad (3.16b)$$

By substituting (3.15) into (3.14), one obtains

$$p(\mathbf{r}; \boldsymbol{\varphi}) = \frac{1}{|M|^{N'n_t}} \sum_{i=1}^{|M|^{N'n_t}} \frac{\exp\left(-\frac{1}{2}\mathbf{r}^\dagger \boldsymbol{\Sigma}^{-1}(\mathbf{c}_{\langle i \rangle}, \boldsymbol{\varphi}) \mathbf{r}\right)}{(2\pi)^{Nn_r} \det^{\frac{1}{2}}(\boldsymbol{\Sigma}(\mathbf{c}_{\langle i \rangle}, \boldsymbol{\varphi}))}. \quad (3.17)$$

Finally, by employing (3.17), the NDA-CRLB for MDS estimation in MIMO frequency-selective fading channel is expressed as

$$\text{Var}(\hat{f}_D) \geq I^{-1}(f_D) = \frac{1}{-\mathbb{E}\left\{\frac{\partial^2 \ln p(\mathbf{r}; \boldsymbol{\varphi})}{\partial f_D^2}\right\}}, \quad (3.18)$$

where $I(f_D)$ is given in (3.19), and $\int_{\mathbf{x}} \triangleq \int_{x_1} \int_{x_2} \cdots \int_{x_{(2Nn_r)}}$. As seen, there is no an explicit expression for (3.19), and thus, for the CRLB in (3.18). Therefore, numerical methods are used to solve (3.19) and (3.18).

$$I(f_D) = -\mathbb{E}\left\{\frac{\partial^2 \ln p(\mathbf{r}; \boldsymbol{\varphi})}{\partial f_D^2}\right\} = -\frac{1}{|M|^{N'n_t}} \int_{\mathbf{x}} \frac{\partial^2}{\partial f_D^2} \left(\ln \sum_{i=1}^{|M|^{N'n_t}} \frac{\exp\left(-\frac{1}{2}\mathbf{x}^\dagger \boldsymbol{\Sigma}^{-1}(\mathbf{c}_{\langle i \rangle}, \boldsymbol{\varphi}) \mathbf{x}\right)}{\det^{\frac{1}{2}}(\boldsymbol{\Sigma}(\mathbf{c}_{\langle i \rangle}, \boldsymbol{\varphi}))} \right) \sum_{q=1}^{|M|^{N'n_t}} \frac{\exp\left(-\frac{1}{2}\mathbf{x}^\dagger \boldsymbol{\Sigma}^{-1}(\mathbf{c}_{\langle q \rangle}, \boldsymbol{\varphi}) \mathbf{x}\right)}{(2\pi)^{Nn_r} \det^{\frac{1}{2}}(\boldsymbol{\Sigma}(\mathbf{c}_{\langle q \rangle}, \boldsymbol{\varphi}))} d\mathbf{x}. \quad (3.19)$$

3.2.3 MLE for MDS

In this section, we derive the DA- and NDA-MLEs for MDS in MIMO frequency-selective fading channel.

3.2.3.1 DA-MLE for MDS

The DA-MLE for f_D is obtained as

$$\hat{f}_D = \arg \max_{f_D} p(\mathbf{r}|\mathbf{s}; \boldsymbol{\theta}), \quad (3.20)$$

where $p(\mathbf{r}|\mathbf{s}; \boldsymbol{\theta})$ is given in (3.6). Since $p(\mathbf{r}|\mathbf{s}; \boldsymbol{\theta})$ is a differentiable function, the DA-MLE for f_D is obtained from

$$\frac{\partial \ln p(\mathbf{r}|\mathbf{s}; \boldsymbol{\theta})}{\partial f_D} = 0. \quad (3.21)$$

By substituting (3.6) into (3.21) and after some mathematical manipulations, one obtains

$$\frac{\partial \ln p(\mathbf{r}|\mathbf{s}; \boldsymbol{\theta})}{\partial f_D} = -\frac{1}{2} \text{tr} \left[\boldsymbol{\Sigma}^{-1}(\mathbf{s}, \boldsymbol{\theta}) \frac{\partial \boldsymbol{\Sigma}(\mathbf{s}, \boldsymbol{\theta})}{\partial f_D} \right] + \frac{1}{2} \mathbf{r}^\dagger \boldsymbol{\Sigma}^{-1}(\mathbf{s}, \boldsymbol{\theta}) \frac{\partial \boldsymbol{\Sigma}(\mathbf{s}, \boldsymbol{\theta})}{\partial f_D} \boldsymbol{\Sigma}^{-1}(\mathbf{s}, \boldsymbol{\theta}) \mathbf{r}. \quad (3.22)$$

As seen in (3.22), there is no closed-form solution for (3.21). Thus, numerical methods need to be used to obtain solution. By employing the scoring method [126],² the solution of (3.22) can be iteratively obtained as

$$\hat{f}_D^{[t+1]} = \hat{f}_D^{[t]} + I^{-1}(f_D) \frac{\partial \ln p(\mathbf{r}|\mathbf{s}; \boldsymbol{\theta})}{\partial f_D} \Bigg|_{f_D = \hat{f}_D^{[t]}}, \quad (3.23)$$

where $I(f_D)$ and $\frac{\partial \ln p(\mathbf{r}|\mathbf{s}; \boldsymbol{\theta})}{\partial f_D}$ are given in (3.11) and (3.22), respectively.

²The scoring method replaces the Hessian matrix in the Newtown-Raphson method with the negative of the Fisher information matrix.

3.2.3.2 NDA-MLE for MDS

Similar to the DA-MLE, the NDA-MLE for MDS is obtained from

$$\hat{f}_D = \arg \max_{f_D} p(\mathbf{r}; \boldsymbol{\varphi}), \quad (3.24)$$

where $p(\mathbf{r}; \boldsymbol{\varphi})$ is given in (3.17). Since $p(\mathbf{r}; \boldsymbol{\varphi})$ is a linear combination of differentiable functions, the NDA-MLE for f_D is obtained from

$$\frac{\partial \ln p(\mathbf{r}; \boldsymbol{\varphi})}{\partial f_D} = 0. \quad (3.25)$$

By substituting (3.17) into (3.25) and after some algebra, one obtains

$$\sum_{i=1}^{|M|^{N_t n_t}} \left\{ \frac{\mathbf{r}^\dagger \boldsymbol{\Sigma}^{-1}(\mathbf{c}^{(i)}, \boldsymbol{\varphi}) \frac{\partial \boldsymbol{\Sigma}(\mathbf{c}^{(i)}, \boldsymbol{\varphi})}{\partial f_D} \boldsymbol{\Sigma}^{-1}(\mathbf{c}^{(i)}, \boldsymbol{\varphi}) \mathbf{r}}{\det^{\frac{1}{2}} \boldsymbol{\Sigma}(\mathbf{c}^{(i)}, \boldsymbol{\varphi})} - \frac{\text{tr} \left[\boldsymbol{\Sigma}^{-1}(\mathbf{c}^{(i)}, \boldsymbol{\varphi}) \frac{\partial \boldsymbol{\Sigma}(\mathbf{c}^{(i)}, \boldsymbol{\varphi})}{\partial f_D} \right]}{\det^{\frac{1}{2}} \boldsymbol{\Sigma}(\mathbf{c}^{(i)}, \boldsymbol{\varphi})} \right\} = 0 \quad (3.26)$$

Similar to the DA-MLE, there is no closed-form solution for (3.26); thus, numerical methods are used to solve (3.26).

3.2.4 NDA-MB Estimation of MDS

In this section, we propose an NDA-moment-based (MB) MDS estimator for multiple-input single-output (MISO) systems under frequency-selective Rayleigh fading channel by employing the fourth-order moment of the received signal. Then, an extension of the proposed estimator to the MIMO systems is provided.

3.2.4.1 NDA-MBE for MDS in MISO Systems

Let us assume that the parameter vector $\boldsymbol{\varphi} = [\boldsymbol{\beta}^\dagger \boldsymbol{\xi}^\dagger \boldsymbol{\vartheta}^\dagger f_D]^\dagger$ is unknown at the receive-side. The statistical MB approach enables us to propose an NDA-MBE to estimate f_D without any priori knowledge of $\boldsymbol{\beta}$, $\boldsymbol{\xi}$, and $\boldsymbol{\vartheta}$. Let us consider the fourth-order two-conjugate moment of the received signal at the n th receive antenna, defined as

$$\kappa_u^{(n)} \triangleq \mathbb{E} \left\{ \left| r_k^{(n)} \right|^2 \left| r_{k+u}^{(n)} \right|^2 \right\}. \quad (3.27)$$

With the transmitted symbols, $s_k^{(m)}$, $m = 1, \dots, n_t$ being independent, drawn from symmetric complex-valued constellation points,³ and with $u \geq L$, $\kappa_u^{(n)}$ is expressed as (see Appendix 3.A for proof).

$$\begin{aligned} \kappa_u^{(n)} &= \mathbb{E} \left\{ \left| r_k^{(n)} \right|^2 \left| r_{k+u}^{(n)} \right|^2 \right\} \\ &= \sum_{m=1}^{n_t} \sum_{l=1}^L \mathbb{E} \left\{ \left| h_{k,l}^{(mn)} \right|^2 \left| h_{k+u,l}^{(mn)} \right|^2 \right\} \sigma_{s_m}^4 + \sum_{m_1=1}^{n_t} \sum_{m_2 \neq m_1}^{n_t} \sum_{l=1}^L \mathbb{E} \left\{ \left| h_{k,l}^{(m_1 n)} \right|^2 \left| h_{k+u,l}^{(m_2 n)} \right|^2 \right\} \sigma_{s_{m_1}}^2 \sigma_{s_{m_2}}^2 \\ &\quad + \sum_{m=1}^{n_t} \sum_{l_1=1}^L \sum_{l_2 \neq l_1}^L \mathbb{E} \left\{ \left| h_{k,l_1}^{(mn)} \right|^2 \left| h_{k+u,l_2}^{(mn)} \right|^2 \right\} \sigma_{s_m}^4 + 2\sigma_{w_n}^2 \sum_{m=1}^{n_t} \sum_{l=1}^L \mathbb{E} \left\{ \left| h_{k,l}^{(mn)} \right|^2 \right\} \sigma_{s_m}^2 + \sigma_{w_n}^4 \\ &\quad + \sum_{m_1=1}^{n_t} \sum_{m_2 \neq m_1}^{n_t} \sum_{l_1=1}^L \sum_{l_2 \neq l_1}^L \mathbb{E} \left\{ \left| h_{k,l_1}^{(m_1 n)} \right|^2 \left| h_{k+u,l_2}^{(m_2 n)} \right|^2 \right\} \sigma_{s_{m_1}}^2 \sigma_{s_{m_2}}^2, \quad u \geq L. \end{aligned} \quad (3.28)$$

By employing the first-order autoregressive model of the Rayleigh fading channel, one can write [128, 129]

$$h_{k,l}^{(mn)} = \Psi_u h_{k+u,l}^{(mn)} + v_{k,l}^{(mn)}, \quad (3.29)$$

where $\Psi_u \triangleq J_0(2\pi f_D T_s u)$ and $v_{k,l}^{(mn)}$ is a zero-mean complex-valued Gaussian white process with variance $\mathbb{E}\{|v_{k,l}^{(mn)}|^2\} = (1 - |\Psi_u|^2) \sigma_{h_{(mn),l}}^2$, which is independent of $h_{k,l}^{(mn)}$.

³ $\mathbb{E}\{(s_k^{(m)})^2\} = 0$ for M -ary phase-shift-keying (PSK) and quadrature amplitude modulation (QAM), $M > 2$ [127].

By using (3.29) and exploiting the property of a complex-valued Gaussian random variable $z \sim \mathcal{N}_c(0, \sigma_z^2)$ that $\mathbb{E}\{|z|^{2n}\} = n!\sigma_z^{2n}$ [130], one obtains

$$\begin{aligned}
\mathbb{E}\left\{|h_{k,l}^{(mn)}|^2 |h_{k+u,l}^{(mn)}|^2\right\} &= \mathbb{E}\left\{\left|(J_0(2\pi f_d T_s u)h_{k+u,l}^{(mn)} + v_{k,l}^{(mn)})\right|^2 |h_{k+u,l}^{(mn)}|^2\right\} \\
&= J_0^2(2\pi f_d T_s u)\mathbb{E}\left\{|h_{k+u,l}^{(mn)}|^4\right\} + J_0(2\pi f_d T_s u)\mathbb{E}\left\{(h_{k+u,l}^{(mn)})^* |h_{k+u,l}^{(mn)}|^2 (v_{k,l}^{(mn)})\right\} \\
&\quad + \mathbb{E}\left\{|v_{k,l}^{(mn)}|^2 |h_{k+u,l}^{(mn)}|^2\right\} + J_0(2\pi f_d T_s u)\mathbb{E}\left\{h_{k+u,l}^{(mn)} |h_{k+u,l}^{(mn)}|^2 (v_{k,l}^{(mn)})^*\right\} \\
&= 2J_0^2(2\pi f_d T_s u)\sigma_{h_{(mn),l}}^4 + \left(1 - J_0^2(2\pi f_d T_s u)\right)\sigma_{h_{(mn),l}}^4 \\
&= \left(1 + J_0^2(2\pi f_d T_s u)\right)\sigma_{h_{(mn),l}}^4 \quad m = 1, \dots, n_t, \quad l = 1, \dots, L.
\end{aligned} \tag{3.30}$$

With the channel taps l_1 and l_2 being uncorrelated for each transmit antenna, i.e., $\mathbb{E}\{h_{k,l_1}^{(mn)}(h_{k,l_2}^{(mn)})^*\} = \sigma_{h_{(mn),l_1}}^2 \delta_{l_1,l_2}$ and employing

$$\begin{aligned}
\mathbb{E}\left\{|h_{k,l_1}^{(m_1 n)}|^2 |h_{k+u,l_2}^{(m_2 n)}|^2\right\} &= \sigma_{h_{(m_1 n),l_1}}^2 \sigma_{h_{(m_2 n),l_2}}^2 \left[(1 - \delta_{l_1,l_2})(1 - \delta_{m_1,m_2}) + \delta_{l_1,l_2}(1 - \delta_{m_1,m_2}) \right. \\
&\quad \left. + (1 - \delta_{l_1,l_2})\delta_{m_1,m_2} \right] + \sigma_{h_{(m_1 n),l_1}}^4 \left(1 + J_0^2(2\pi f_d T_s u)\right) \delta_{m_1,m_2} \delta_{l_1,l_2},
\end{aligned} \tag{3.31}$$

one can write (3.28) as

$$\begin{aligned}
\kappa_u^{(n)} &= \sum_{m=1}^{n_t} \sum_{l=1}^L \sigma_{h_{(mn),l}}^4 \sigma_{s_m}^4 \left(1 + J_0^2(2\pi f_d T_s u)\right) + \sum_{m_1=1}^{n_t} \sum_{m_2 \neq m_1}^{n_t} \sum_{l=1}^L \sigma_{h_{(m_1 n),l}}^2 \sigma_{h_{(m_2 n),l}}^2 \sigma_{s_{m_1}}^2 \sigma_{s_{m_2}}^2 \\
&\quad + \sum_{m=1}^{n_t} \sum_{l_1=1}^L \sum_{l_2 \neq l_1}^L \sigma_{h_{(mn),l_1}}^2 \sigma_{h_{(mn),l_2}}^2 \sigma_{s_m}^4 + \sum_{m_1=1}^{n_t} \sum_{m_2 \neq m_1}^{n_t} \sum_{l_1=1}^L \sum_{l_2 \neq l_1}^L \sigma_{h_{(m_1 n),l_1}}^2 \sigma_{h_{(m_2 n),l_2}}^2 \sigma_{s_{m_1}}^2 \sigma_{s_{m_2}}^2 \\
&\quad + 2\sigma_{w_n}^2 \sum_{m=1}^{n_t} \sum_{l=1}^L \sigma_{h_{(mn),l}}^2 \sigma_{s_m}^2 + \sigma_{w_n}^4.
\end{aligned} \tag{3.32}$$

Further, let us consider the second-order moment of the received signal, i.e., $\mu_2^{(n)} \triangleq \mathbb{E}\{|r_k^{(n)}|^2\}$. By using (3.1), it can be easily shown that

$$\mu_2^{(n)} = \sum_{m=1}^{n_t} \sum_{l=1}^L \sigma_{h_{(mn),l}}^2 \sigma_{s_m}^2 + \sigma_{w_n}^2. \quad (3.33)$$

By employing (3.32) and (3.33), one obtains the normalized squared autocorrelation function (AF) of the fading channel as (see Appendix 3.B for proof)

$$\Psi_u \triangleq J_0^2(2\pi f_d T_s u) = \eta^{(n)} \left(\kappa_u^{(n)} - \left(\mu_2^{(n)} \right)^2 \right), \quad (3.34)$$

where $\eta^{(n)} = 1 / \sum_{m=1}^{n_t} \sum_{l=1}^L \sigma_{h_{(mn),l}}^4 \sigma_{s_m}^4$.

For non-constant modulus constellations, $\eta^{(n)}$ is expressed in terms of $\mu_4^{(n)} \triangleq \mathbb{E}\{|r_k^{(n)}|^4\}$ and $\mu_2^{(n)}$ as (see Appendix 3.C for proof)

$$\eta^{(n)} = \frac{2(\Omega_s - 1)}{\mu_4^{(n)} - 2\left(\mu_2^{(n)}\right)^2}, \quad (3.35)$$

where $\Omega_s = \frac{1}{|M|} \sum_{i=1}^{|M|} |c_i|^4$ is a constant, and $1 < \Omega_s \leq 2$.⁴

Finally, substituting (3.35) into (3.34) yields

$$\Psi_u \triangleq 2(\Omega_s - 1) \frac{\kappa_u^{(n)} - \left(\mu_2^{(n)}\right)^2}{\mu_4^{(n)} - 2\left(\mu_2^{(n)}\right)^2}. \quad (3.36)$$

As seen, the normalized squared AF of the fading channel is expressed as a non-linear function of the $\mu_2^{(n)}$, $\mu_4^{(n)}$, and $\kappa_u^{(n)}$. In practice, statistical moments are estimated by time averages of the received signal. For (3.36), the following estimators of the moments are employed

⁴For 16-QAM, 64-QAM, and complex-valued zero-mean Gaussian signals, Ω_s is 1.32, 1.38, and 2, respectively [127].

$$\begin{aligned}
\hat{\mu}_2^{(n)} &= \frac{1}{N} \sum_{k=1}^N |r_k^{(n)}|^2 \\
\hat{\mu}_4^{(n)} &= \frac{1}{N} \sum_{k=1}^N |r_k^{(n)}|^4 \\
\hat{\kappa}_u^{(n)} &= \frac{1}{N-u} \sum_{k=1}^{N-u} |r_k^{(n)}|^2 |r_{k+u}^{(n)}|^2,
\end{aligned} \tag{3.37}$$

where $u \geq L > 0$.

By substituting the corresponding estimators in (3.36), the estimate of the normalized squared AF is given as

$$\hat{\Psi}_u^{(n)} \triangleq 2(\Omega_s - 1) \frac{\hat{\kappa}_u^{(n)} - \left(\hat{\mu}_2^{(n)}\right)^2}{\hat{\mu}_4^{(n)} - 2\left(\hat{\mu}_2^{(n)}\right)^2}. \tag{3.38}$$

Now, based on (3.34) and (3.38), the problem of MDS estimation can be formulated as a non-linear regression problem. Given the estimated normalized squared AF, $\hat{\Psi}_u^{(n)}$, the non-linear regression model assumes that the relationship between $\hat{\Psi}_u^{(n)}$ and Ψ_u is modeled through a disturbance term or error variable $\epsilon_u^{(n)}$ as [131, 132]

$$\hat{\Psi}_u^{(n)} = \Psi_u + \epsilon_u^{(n)} = J_0^2(2\pi f_D T_s u) + \epsilon_u^{(n)}, \quad u = U_{\min}, \dots, U_{\max}, \tag{3.39}$$

where U_{\min} and U_{\max} are the maximum and minimum delay lags, respectively.

To solve the non-linear regression problem in (3.39), the LS curve-fitting optimization technique is employed. Based on the LS curve-fitting optimization, the estimate of f_D , i.e., \hat{f}_D , is obtained through minimizing the sum of the squared residuals (SSR) as [132]

$$\underset{f_D}{\text{minimize}} \quad \sum_{u=U_{\min}}^{U_{\max}} \left(\hat{\Psi}_u^{(n)} - J_0^2(2\pi f_D T_s u) \right)^2 \quad (3.40)$$

$$\text{subject to } f_l \leq f_D \leq f_h,$$

where f_l and f_h are the minimum and maximum possible MDSs, respectively. To obtain \hat{f}_D , we consider the derivative of the SSR with respect to f_D and set it equal to zero as follows

$$\sum_{u=U_{\min}}^{M_{\max}} 8\pi T_s u \left(\hat{\Psi}_u^{(n)} - J_0^2(2\pi f_D T_s u) \right) \quad (3.41)$$

$$J_0(2\pi f_D T_s u) J_1(2\pi f_D T_s u) = 0.$$

As seen, for the non-linear regression, the derivative in (3.41) is a function of f_D . Thus, an explicit solution for \hat{f}_D cannot be obtained. However, numerical methods [133] can be employed to solve the LS curve-fitting optimization problem in (3.40).

By employing the Newton-Raphson method, \hat{f}_D can be iteratively obtained as in (3.42). The main problem with the Newton-Raphson method is that it suffers from the

$$\hat{f}_D^{[t+1]} = \hat{f}_D^{[t]} - \frac{\sum_{u=U_{\min}}^{M_{\max}} 8\pi T_s u \left(\hat{\Psi}_u^{(n)} - J_0^2(2\pi f_D^{[t]} T_s u) \right) J_0(2\pi f_D^{[t]} T_s u) J_1(2\pi f_D^{[t]} T_s u)}{\frac{\partial^2}{\partial f_D^2} \sum_{u=U_{\min}}^{U_{\max}} \left(\hat{\Psi}_u^{(n)} - J_0^2(2\pi f_D T_s u) \right)^2 \Big|_{f_D = \hat{f}_D^{[t]}}} \quad (3.42)$$

$$\frac{\partial^2}{\partial f_D^2} \sum_{u=U_{\min}}^{U_{\max}} \left(\hat{\Psi}_u - J_0^2(2\pi f_D T_s u) \right)^2 \Big|_{f_D = \hat{f}_D^{[t]}} = \sum_{u=U_{\min}}^{M_{\max}} \left\{ 32\pi^2 T_s^2 u^2 J_0^2(2\pi f_D^{[t]} T_s u) J_1^2(2\pi f_D^{[t]} T_s u) \right.$$

$$+ 8\pi T_s u \left(2\pi T_s u \left(J_0^2(2\pi f_D^{[t]} T_s u) - J_1^2(2\pi f_D^{[t]} T_s u) \right) - \frac{J_0(2\pi f_D^{[t]} T_s u) J_1(2\pi f_D^{[t]} T_s u)}{f_D^{[t]}} \right)$$

$$\left. \left(\hat{\Psi}_u^{(n)} - J_0^2(2\pi f_D^{[t]} T_s u) \right) \right\}.$$

Algorithm 7 : NDA-MBE for MDS in MISO systems

- 1: Set f_l , f_h , Δ , and δ
 - 2: Acquire the measurements $\{r_k^{(n)}\}_{k=1}^N$
 - 3: Estimate the statistics $\hat{\mu}_2^{(n)}$, $\hat{\mu}_4^{(n)}$, and $\hat{\kappa}_u^{(n)}$, by employing (3.37)
 - 4: Compute $\hat{\Psi}_u^{(n)}$, $\forall u \in \{U_{\min}, \dots, U_{\max}\}$ by using (3.38)
 - 5: Obtain $\hat{f}_D^{(r)}$ by solving the minimization problem in (3.40) through the grid search method with grid step size Δ
 - 6: Obtain $\hat{f}_D^{(s)}$ by solving the minimization problem in (3.40) through the grid search method over $[\hat{f}_D^{(r)} - \Delta, \hat{f}_D^{(r)} + \Delta]$ with grid step size δ
 - 7: $\hat{f}_D = \hat{f}_D^{(s)}$
-

convergence problem [134]. Since the parameter space for the MDS estimation is one-dimensional, the grid search method can be employed, which ensures the convergence. With the grid search method, the parameter space, i.e., $[f_l, f_h]$ is discretized as a grid with step size δ , and the value which minimizes SSR is considered as the estimated f_D . This procedure can be performed in two steps, including a rough estimate of the MDS, $\hat{f}_D^{(r)}$, by choosing a larger step size Δ followed by a fine estimate, $\hat{f}_D^{(s)}$, through small grid step size δ around the rough estimate, i.e., $[\hat{f}_D^{(r)} - \Delta, \hat{f}_D^{(r)} + \Delta]$. A formal description of the proposed NDA-MBE for MDS in MISO frequency-selective channel is presented in Algorithm 7.

It is worth noting that f_D can be estimated by using a downsampled version of $\hat{\Psi}_u^{(n)}$. For the case of uniform downsampling, i.e., $u = \ell u_s$, the SSR is given as

$$\sum_{\ell=0}^{N_{\text{la}}-1} \left(\hat{\Psi}_{U_{\min} + \ell u_s}^{(n)} - \Psi_{U_{\min} + \ell u_s} \right)^2, \quad (3.43)$$

where u_s is the downsampling period expressed in delay lags, N_{la} is the number of delay lag,

$$\Psi_{\ell u_s} = J_0^2 \left(2\pi f_D T_s (U_{\min} + \ell u_s) \right), \quad (3.44)$$

and

$$\hat{\Psi}_{U_{\min+\ell u_s}}^{(n)} = 2(\Omega_s - 1) \frac{\hat{\kappa}_{U_{\min+\ell u_s}}^{(n)} - \left(\hat{\mu}_2^{(n)}\right)^2}{\hat{\mu}_4^{(n)} - 2\left(\hat{\mu}_2^{(n)}\right)^2}. \quad (3.45)$$

The downsampled version of $\hat{\Psi}_u^{(n)}$ is usually employed for the rough MDS estimation, where Δ is a large value.

3.2.4.2 NDA-MBE for MDS in MIMO Systems

The performance of the proposed NDA-MBE for MDS in MISO system can be improved when employing multiple receive antennas due to the spatial diversity, by combining the estimated normalized squared AFs, $\hat{\Psi}_u^{(n)}$, $n = 1, \dots, n_r$ as

$$\tilde{\Psi}_u = \sum_{n=1}^{n_r} \lambda_u^{(n)} \hat{\Psi}_u^{(n)}, \quad (3.46)$$

where $\mathbf{\Lambda}_u \triangleq [\lambda_u^{(1)} \lambda_u^{(2)} \dots \lambda_u^{(n_r)}]^\dagger$, with $\sum_{n=1}^{n_r} \lambda_u^{(n)} = 1$, is the weighting vector. Let us define $\hat{\Psi}_u \triangleq [\hat{\Psi}_u^{(1)} \hat{\Psi}_u^{(2)} \dots \hat{\Psi}_u^{(n_r)}]^\dagger$. The mean square error (MSE) of the combined normalized squared AF in (3.46) is expressed as

$$\mathbb{E}\left\{\left(\tilde{\Psi}_u - \Psi_u\right)^2\right\} = \mathbf{\Lambda}_u^\dagger \mathbf{C}_u \mathbf{\Lambda}_u + \left(\mathbf{\Lambda}_u^\dagger \boldsymbol{\mu}_u - \Psi_u\right)^2, \quad (3.47)$$

where $\mathbf{C}_u \triangleq \mathbb{E}\left\{\left(\hat{\Psi}_u - \boldsymbol{\mu}_u\right)\left(\hat{\Psi}_u - \boldsymbol{\mu}_u\right)^\dagger\right\}$ and $\boldsymbol{\mu}_u \triangleq \mathbb{E}\left\{\hat{\Psi}_u\right\}$.

By employing the method of Lagrange multipliers, the optimal weighting vector $\mathbf{\Lambda}_u^{\text{op}}$ in (3.47) in terms of minimum MSE is obtained as

$$\mathbf{\Lambda}_u^{\text{op}} = \left(\mathbf{1}^\dagger \mathbf{y}_u\right)^{-1} \mathbf{y}_u, \quad (3.48)$$

where $\mathbf{y}_u \triangleq \left(\mathbf{C}_u + (\boldsymbol{\mu}_u - \Psi_u \mathbf{1})(\boldsymbol{\mu}_u - \Psi_u \mathbf{1})^\dagger\right)^{-1} \mathbf{1}$ and $\mathbf{1} \triangleq [1 \ 1 \ \dots \ 1]^\dagger$ is an n_r -dimensional

Algorithm 8 : Bootstrap Algorithm for Optimal Combining

- 1: Set N_B
- 2: **for** $t = 1, 2, \dots, N_B$ **do**
- 3: Draw a random sample of size N , with replacement, from $\mathcal{X} \triangleq \{1, 2, \dots, N\}$ and name it \mathcal{X}^*
- 4: **for** $n = 1, 2, \dots, n_r$ **do**

$$\hat{\Psi}_u^{(n)^*}[t] = \frac{\frac{1}{N-u} \sum_{k \in \mathcal{X}^*} |r_k^{(n)}|^2 |r_{k+u}^{(n)}|^2 - \left(\frac{1}{N} \sum_{k \in \mathcal{X}^*} |r_k^{(n)}|^2 \right)^2}{\frac{1}{N} \sum_{k \in \mathcal{X}^*} |r_k^{(n)}|^4 - 2 \left(\frac{1}{N} \sum_{k \in \mathcal{X}^*} |r_k^{(n)}|^2 \right)^2}$$

- 5: **end for**
 - 6: $\hat{\Psi}_u^*[t] \triangleq 2(\Omega_s - 1) [\hat{\Psi}_u^{(1)^*} \hat{\Psi}_u^{(2)^*} \dots \hat{\Psi}_u^{(n_r)^*}]^\dagger$
 - 7: **end for**
 - 8: $\Gamma_u = [\hat{\Psi}_u^*[1] \hat{\Psi}_u^*[2] \dots \hat{\Psi}_u^*[N_B]]$
 - 9: $\hat{\boldsymbol{\mu}}_u = \frac{1}{N_B} \sum_{t=1}^{N_B} \hat{\Psi}_u^*[t]$
 - 10: $\hat{\mathbf{C}}_u = \frac{1}{N_B - 1} (\Gamma_u - \hat{\boldsymbol{\mu}}_u \mathbf{1}^\dagger) (\Gamma_u - \hat{\boldsymbol{\mu}}_u \mathbf{1}^\dagger)^\dagger$
-

vector of ones.

As seen, the optimal weighting vector, $\boldsymbol{\Lambda}_u^{\text{op}}$, in (3.48) depends on the true value of MDS, i.e., f_D , through the true normalized squared AF, Ψ_u , in \mathbf{y}_u . To obtain the optimal weighting vector, the mean vector $\boldsymbol{\mu}_u$ and covariance matrix \mathbf{C}_u are required to be estimated from the received symbols. One approach is bootstrapping [135–137]. The bootstrap method suggests to re-sample the empirical joint cumulative distribution function (CDF) of $\hat{\Psi}_u$ to estimate $\boldsymbol{\mu}_u$ and \mathbf{C}_u as summarized in Algorithm 8.⁵

As seen in Algorithm 8, the optimal weighting vector for each delay lag u is derived at the expense of higher computational complexity. In order to avoid this computational complexity, the suboptimal equal weight combining method can be

⁵Since $\hat{\Psi}_u^{(n)}$, $n = 1, \dots, n_r$ are uncorrelated random variables, \mathbf{C}_u is a diagonal matrix. Thus, only the diagonal elements of $\hat{\mathbf{C}}_u$ are employed to obtain the optimal weighting vector.

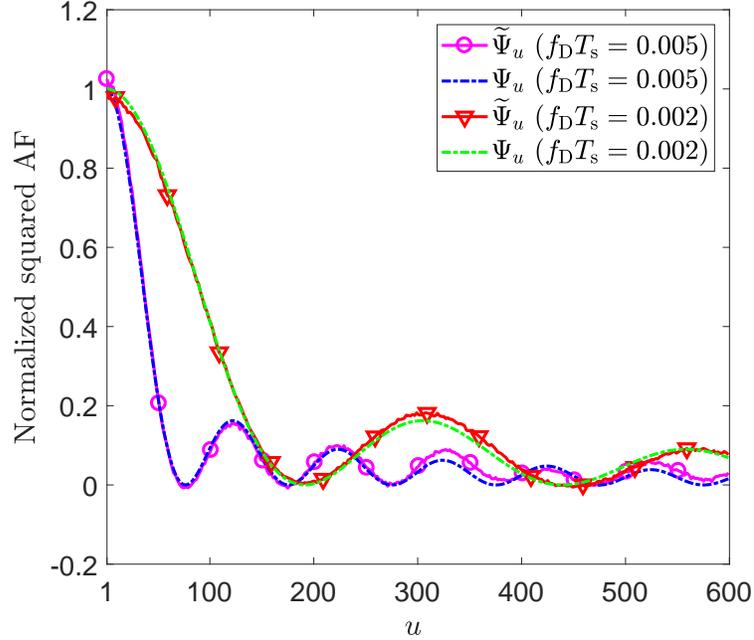


Fig. 3.1: Illustration of the non-linear LS regression for the uniformly sampled normalized squared AF for $f_D T_s = 0.02$ and $f_D T_s = 0.005$, with $n_t = 1$, $n_r = 2$, $L = 1$, $u_s = 2$, and at $\gamma = 10$ dB.

employed as

$$\tilde{\Psi}_u = \frac{1}{n_r} \sum_{n=1}^{n_r} \hat{\Psi}_u^{(n)}. \quad (3.49)$$

Fig. 3.1 shows how $\tilde{\Psi}_u$ fits Ψ_u through the equal weight combining in (3.49) for $f_d T_s = 0.02$ and $f_d T_s = 0.005$ with $n_t = 1$, $n_r = 2$, $L = 1$, $u_s = 2$, and at $\gamma = 10$ dB.

Finally, similar to the MISO scenario, the problem of MDS estimation for multiple receive antennas is formulated as non-linear regression problem in (3.39) for $\tilde{\Psi}_u$. A formal description of the proposed NDA-MBE for MDS in MIMO frequency-selective channel is presented in Algorithm 9.

Algorithm 9 : NDA-MBE for MDS in MIMO systems

- 1: Set f_1 , f_h , Δ , and δ
 - 2: Acquire the measurements $\{r_k^{(n)}\}_{k=1}^N$, $\forall n \in \{1, \dots, n_r\}$
 - 3: Estimate the statistics $\hat{\mu}_2^{(n)}$, $\hat{\mu}_4^{(n)}$, and $\hat{\kappa}_u^{(n)}$, by employing (3.37) for $\{r_k^{(n)}\}_{k=1}^N$, $\forall n \in \{1, \dots, n_r\}$
 - 4: Compute $\hat{\Psi}_u^{(n)}$, $\forall u \in \{U_{\min}, \dots, U_{\max}\}$, $\forall n \in \{1, \dots, n_r\}$, by using (3.38)
 - 5: Compute $\tilde{\Psi}_u$, $\forall u \in \{U_{\min}, \dots, U_{\max}\}$, by using (3.49)
 - 6: Obtain $\hat{f}_D^{(r)}$ by solving the minimization problem in (3.40) for $\tilde{\Psi}_u$ through the grid search method with step size Δ
 - 7: Obtain $\hat{f}_D^{(s)}$ by solving the minimization problem in (3.40) for $\tilde{\Psi}_u$ via the grid search method over $[\hat{f}_D^{(r)} - \Delta, \hat{f}_D^{(r)} + \Delta]$ with step size δ
 - 8: $\hat{f}_D = \hat{f}_D^{(s)}$
-

3.2.4.3 Semi-blind NDA-MBE

The proposed NDA-MBE for MISO and MIMO systems do not require knowledge of the parameter vector $\varphi = [\beta^\dagger \ \xi^\dagger \vartheta^\dagger \ f_D]^\dagger$. In other words, the proposed NDA-MBE in section 3.2.4.1 and 3.2.4.2 are blind. For the scenarios in which the variance of the additive noise can be accurately estimated at the receive antennas, i.e., ξ is known, a semi-blind NDA-MBE for the case of SISO transmission and flat-fading channel, i.e., $n_t = 1$ and $L = 1$, can be proposed. In this case, for the n th receive antennas, one can easily obtain⁶

$$\mu_2^{(n)} = \sigma_{h_n}^2 \sigma_s^2 + \sigma_{w_n}^2 \quad (3.50)$$

and

$$\eta^{(n)} = (\sigma_{h_n}^4 \sigma_s^4)^{-1} = \frac{1}{\left(\mu_2^{(n)} - \sigma_{w_n}^2\right)^2}. \quad (3.51)$$

⁶The index of transmit antenna, i.e., $m = 1$ and the index of channel tap, i.e., $l = 1$ is dropped.

By using (3.34), (3.50) and (3.51), and by replacing the statistical moments and the noise variance with their corresponding estimates, one obtains

$$\hat{\Psi}_u^{(n)} = \frac{\hat{\kappa}_u^{(n)} - \left(\hat{\mu}_2^{(n)}\right)^2}{\left(\hat{\mu}_2^{(n)} - \hat{\sigma}_w^2\right)^2}, \quad (3.52)$$

where $\hat{\sigma}_{w_n}^2$ is the estimate of the noise variance, and $\hat{\kappa}_u^{(n)}$ and $\hat{\mu}_2^{(n)}$ are given in (3.37). Clearly, similar to the SISO transmission, the optimal and suboptimal combining methods for the multiple receive antennas can be employed, as well.

3.3 Complexity Analysis

By employing the two steps grid search method to solve the optimization problem in (3.40), the number of real additions and multiplications employed in the proposed NDA-MBE is shown in Table 3.1, where N_{la} is the number of delay lag, $N_g \triangleq (N_{g_1} + N_{g_2})$, and N_{g_1} and N_{g_2} are the number of grid points used for the rough and fine estimation, respectively. As seen, the proposed NDA-MBE exhibits a complexity order of $\mathcal{O}(N)$. It should be mentioned that the complexity order of the derived DA-MLE and NDA-MLE are $\mathcal{O}(N^3)$ and $\mathcal{O}(|M|^{N'n_t})$, respectively.

Table 3.1: Number of real additions, real multiplications, and complexity order of the proposed NDA-MBE.

Algorithm	Real additions	Real multiplications	Order
MISO	$\left(N + 2N_g - \frac{(U_{\max} + U_{\min})}{2}\right)N_{\text{la}} + 3N - N_g - 1$	$\left(N + N_g - \frac{(U_{\max} + U_{\min})}{2} + 2\right)N_{\text{la}} + 3N + 4$	$\mathcal{O}(N)$
MIMO	$n_r \left(\left(N - \frac{(U_{\max} + U_{\min})}{2}\right)N_{\text{la}} + 3N - 1 \right) + (2N_{\text{la}} - 1)N_g$	$n_r \left(\left(N - \frac{(U_{\max} + U_{\min})}{2} + 2\right)N_{\text{la}} + 3N + 4 \right) + N_g N_{\text{la}} + 1$	$\mathcal{O}(N)$

Fig. 3.2 compares the total number of operations used by the proposed NDA-MBE with the low-complexity DA-MLE in [1, 2] and the DA-COMAT estimator

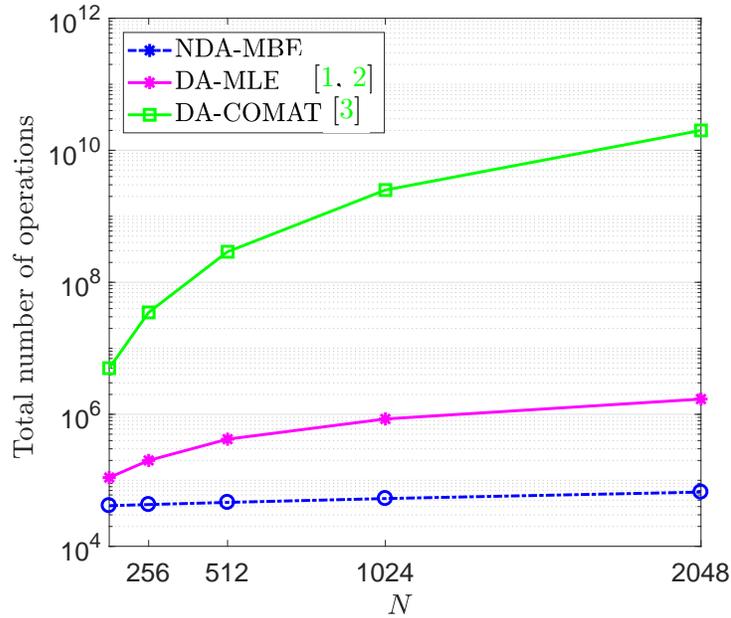


Fig. 3.2: Computational complexity comparison of the proposed NDA-MBE, the low-complexity DA-MLE in [1, 2], and the DA-COMAT estimator in [3].

in [3]. As seen, the proposed NDA-MBE exhibits significantly lower computational complexity compared to the DA-COMAT [3] and the low-complexity DA-MLE in [1, 2]. This substantial reduced-complexity enables the proposed MBE to exhibit good performance in the NDA scenarios, where the observation window can be selected large enough.

3.4 Simulations

In this section, we examine the performance of the proposed NDA-MBE, as well as the derived DA-MLE and DA-CRLB for MDS in MIMO frequency-selective fading channel through several simulation experiments.

3.4.1 Simulation Setup

We consider a MIMO system employing spatial multiplexing, with carrier frequency $f_c = 2.4$ GHz. Unless otherwise mentioned, $n_t = 2$, $n_r = 2$, $T_s = 10 \mu s$, $N = 10^5$, and the modulation is 64-QAM. The delay profile of the Rayleigh fading channel is $\sigma_{h_{(mn),l}}^2 = \beta \exp(-l_{\text{rms}}l/L)$, where β is a normalization factor, i.e., $\beta \sum_l (-l_{\text{rms}}l/L) = 1$, with $L = 5$ and $l_{\text{rms}} = L/4$ as the maximum and RMS delay spread of the channel, respectively. The parameters for the downsampled LS curve-fitting optimization are $U_{\min} = L$, $U_{\max} = \lfloor \frac{N}{4} \rfloor$, and $u_s = 10$. The additive white noise was modeled as a complex-valued Gaussian random variable with zero-mean and variance σ_w^2 for each receive antennas. Without loss of generality, it was assumed that $\sigma_{s_m}^2 = 1/(n_t n_r)$, $m = 1, 2, \dots, n_t$, and thus, the average SNR was defined as $\gamma = 10 \log(1/n_r \sigma_w^2)$. The performance of the MDS estimators was presented in terms of normalized root-mean-square error (NRMSE), i.e., $\mathbb{E}\{(\hat{f}_D T_s - f_D T_s)^2\}^{1/2}/f_D T_s$, obtained from 1000 Monte Carlo trials for each $f_D T_s \in [10^{-3}, 18 \times 10^{-3}]$ with the search step size $\Delta = 10$ Hz and $\delta = 0.5$ Hz, respectively.

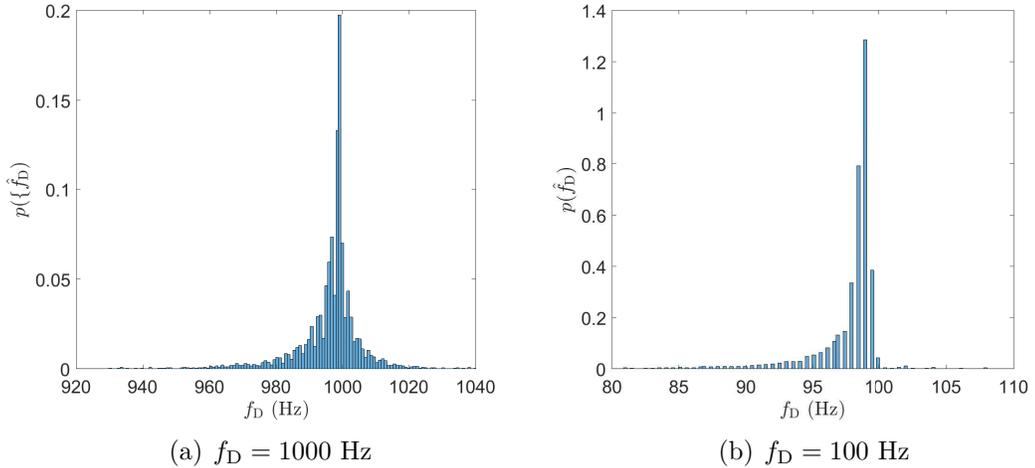


Fig. 3.3: Distribution of the estimated \hat{f}_D for the proposed NDA-MBE for $n_t = 2$, $n_r = 2$, and at $\gamma = 10$ dB.

3.4.2 Simulation Results

Fig. 3.3 shows the distributions of the estimated \hat{f}_D by the proposed NDA-MBE for different MDSs, $f_D = 1000$ Hz and $f_D = 100$ Hz, with $n_t = 2$, $n_r = 2$, and at $\gamma = 10$ dB. As seen, the distributions are not symmetric around their mean values; hence, this leads to bias in MDS estimation. Furthermore, Fig. 3.4 illustrates $\mathbb{E}\{\hat{f}_D/f_D\}$ versus f_D for $\gamma = 10$ dB and $\gamma = 20$ dB. As seen, the proposed NDA-MBE is nearly unbiased, i.e., $\mathbb{E}\{\hat{f}_D\} \approx f_D$ over a wide range of MDS. This can be explained, as while the distribution of the estimated f_D is not symmetric, the estimated values are accumulated around their mean value. It should be mentioned that by increasing the length of the observation window, N , the bias of the proposed estimator approaches zero.

In Fig. 3.5, the NRMSE of the NDA-MBE versus $f_d T_s$ is illustrated for $\gamma = 0$ dB, $\gamma = 10$ dB, and $\gamma = 20$ dB. As seen, the proposed estimator exhibits a good performance over a wide range of Doppler rates, $f_d T_s$. As observed, the NRMSE decreases as $f_d T_s$ increases. This performance improvement can be explained, as for lower Doppler rates, a larger observation window is required to capture the variation of the fading channel. Also, as expected, the NRMSE decreases as γ increases. This can be easily explained, as an increase in γ leads to more accurate estimates of the statistics in (3.38).

Fig. 3.6 presents the NRMSE of the proposed NDA-MBE versus $f_d T_s$ for different numbers of transmit antennas, n_t , for $n_r = 2$ and at $\gamma = 10$ dB. As expected, the NRMSE increases as the number of transmit antennas increases. This increase can be explained, as the variance of the statistics employed in (3.38) increases with the number of transmit antennas, thus, leading to higher estimation error in curve-fitting.

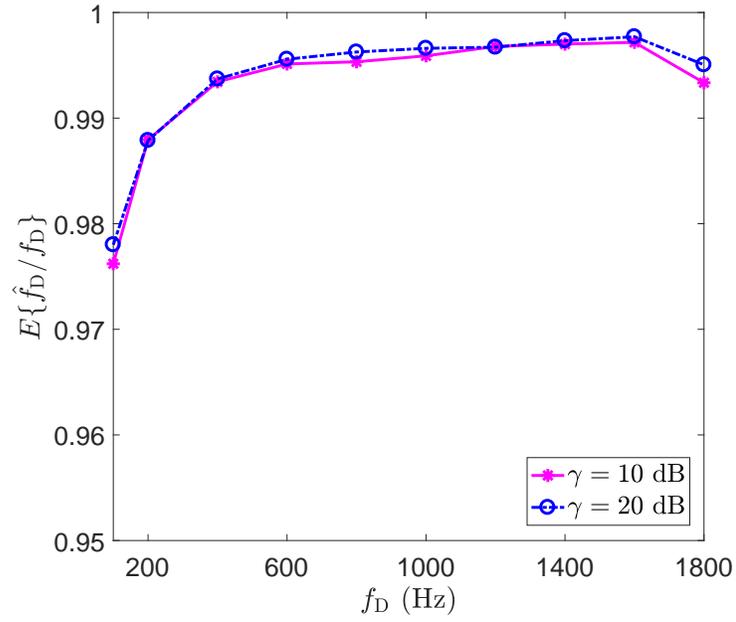


Fig. 3.4: The mean value of the estimated MDS by NDA-MBE for various SNR values for $n_t = 2$ and $n_r = 2$.

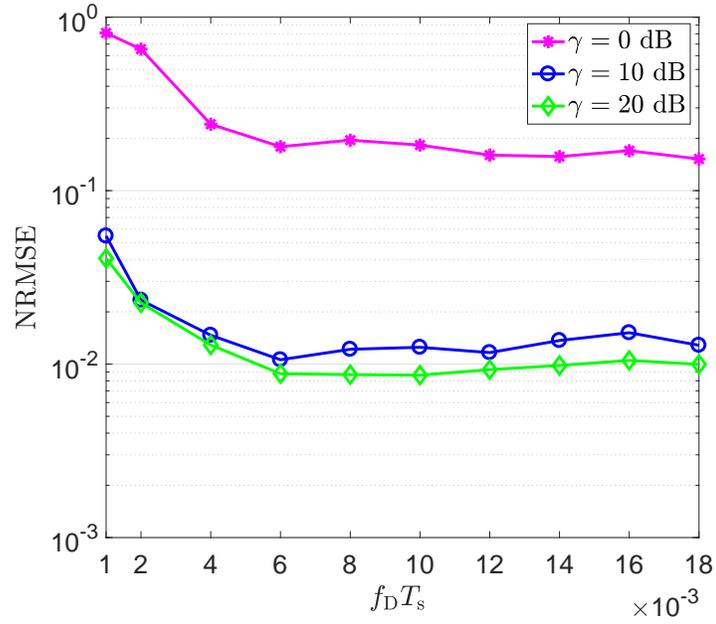


Fig. 3.5: The NRMSE of the proposed NDA-MBE versus $f_D T_s$ for different SNR values.

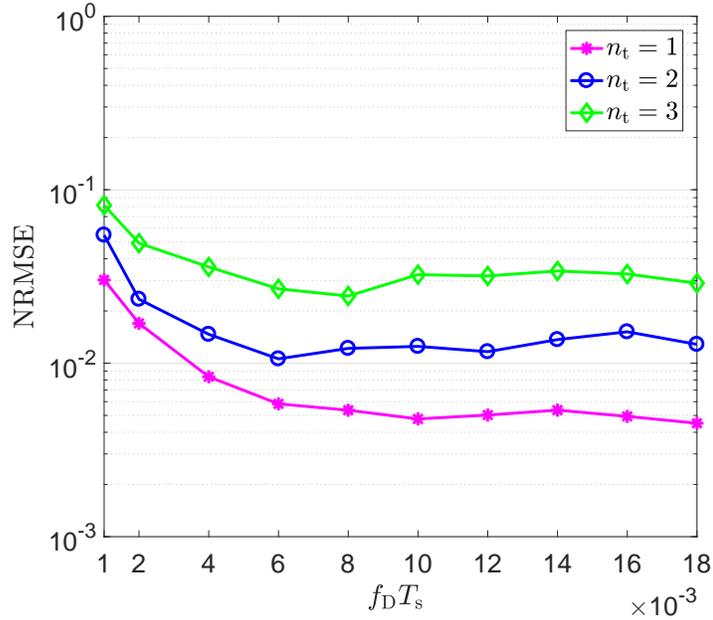


Fig. 3.6: The effect of n_t on the performance of the proposed NDA-MBE for $n_r = 2$ and at $\gamma = 10$ dB.

In Fig. 3.7, the NRMSE of the proposed NDA-MBE is shown versus $f_D T_s$ for different numbers of receive antennas, n_r , for $n_t = 2$, and at $\gamma = 10$ dB. It can be seen that an increment in n_r leads to a reduced NRMSE. This decrease can be easily explained, as averaging at the receive-side yields more accurate estimation of Ψ_u , thus, leading to a more accurate result in the LS curve-fitting.

In Fig. 3.8, the effect of the parameter U_{\min} on the performance of the proposed NDA-MBE is illustrated for $L = 5$ and $u_s = 10$. As observed, the proposed estimator exhibits a low sensitivity to the value of U_{\min} . This can be explained, as a large number of lags, $U_{\min} \leq u \leq U_{\max}$, are employed for fitting $\tilde{\Psi}_u$ to $J_0^2(2\pi f_D T_s u)$ in the LS estimation; thus, the estimator is nearly robust to a few missing delay lags, $L \leq u < U_{\min}$, or nuisance delay lags, $1 \leq u < L$. As such, basically the estimator

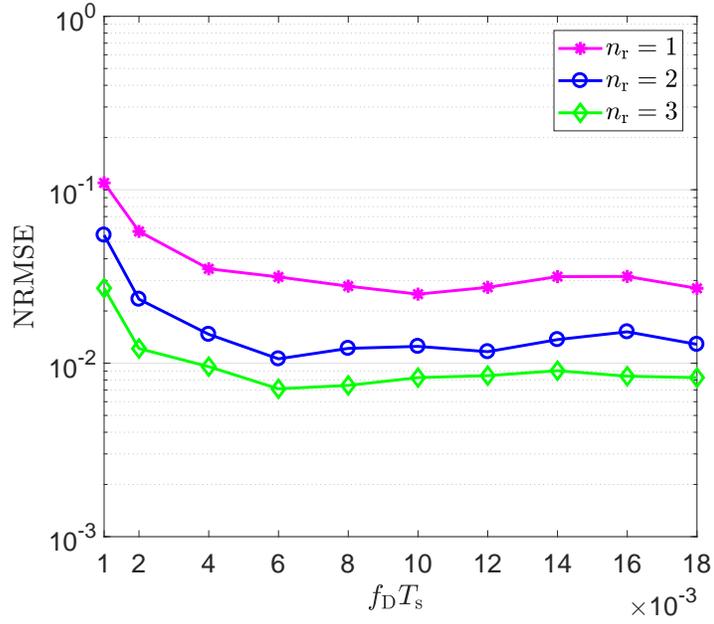


Fig. 3.7: The effect of n_r on the performance of the proposed NDA-MBE for $n_t = 2$ and at $\gamma = 10$ dB.

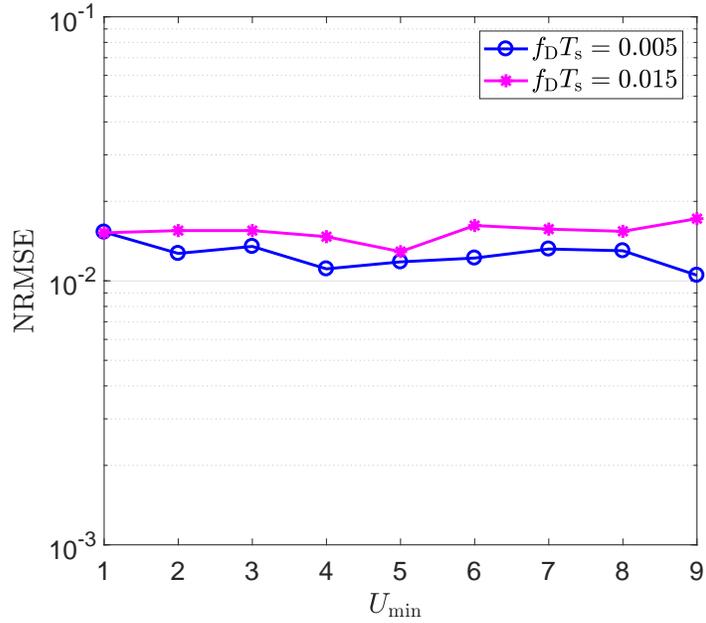


Fig. 3.8: The effect of the parameter U_{\min} on the performance of the proposed NDA-MBE for different values of $f_D T_s$ and at $\gamma = 10$ dB.

does not require an accurate estimate of L .

Fig. 3.9 shows the effect of the observation window size, N , on the performance of the proposed NDA-MBE. As expected, the performance of the proposed estimator improves as the length of the observation window increases. This performance improvement can be explained, as the variance of the estimated statistics employed in (3.38) decreases when N increases.

In Fig. 3.10, the NRMSE is plotted versus $f_D T_s$ for the proposed NDA-MBE, the low-complexity DA-MLE in [1, 2], the NDA-CC estimator (CCE) in [4], and the DA-CRLB for MDS estimation in SISO frequency-flat fading channel for $N = 1000$ and at $\gamma = 10$ dB. As seen, the proposed NDA-MBE outperforms the NDA-CCE, and provides a similar performance as the DA-MLE for $f_D T_s \geq 0.012$.

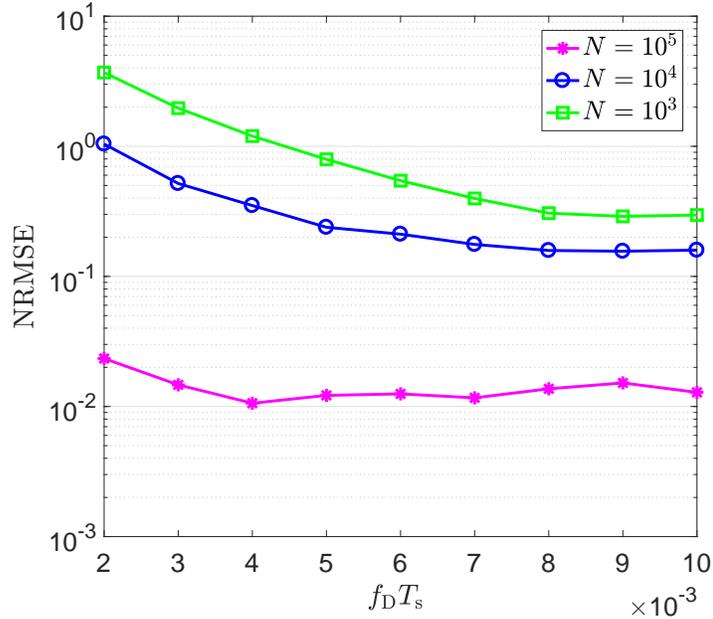


Fig. 3.9: The effect of the observation window size, N , on the performance of the proposed NDA-MBE for $n_t = 2$ and $n_r = 2$ in frequency-selective channel, and at $\gamma = 10$ dB.

Fig. 3.11 illustrates the NRMSE versus $f_D T_s$ for the proposed NDA-MBE, the

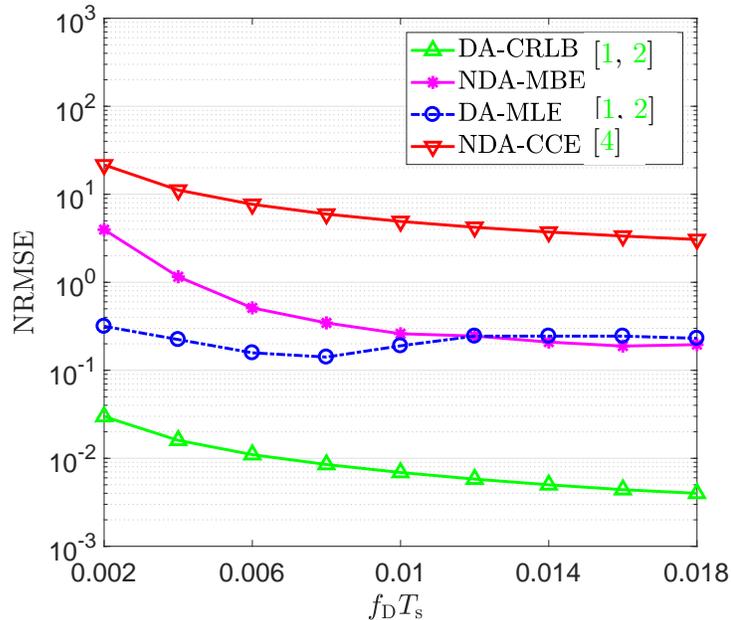


Fig. 3.10: Performance comparison of the proposed NDA-MBE, DA-CRLB [1, 2], the low-complexity DA-MLE in [1, 2], and the NDA-CCE in [4] in SISO frequency-flat Rayleigh fading channel for $N = 10^3$ at $\gamma = 10$ dB.

derived DA-MLE, and the derived DA-CRLB in MIMO frequency-selective fading channel for $N = 1000$ at $\gamma = 10$ dB, As seen, the performance of the derived DA-MLE is close to the DA-CRLB. This high performance is obtained at the expense of significant computational complexity in the order of $\mathcal{O}(N^3)$. On the other hand, the proposed NDA-MBE cannot reach the DA-CRLB. This behaviour can be explained, as the NDA-MBE requires a larger number of observation symbols to accurately estimate the second- and fourth-order statistics in time-varying channel.

In Fig. 3.12, the NRMSE is shown versus $f_D T_s$ for the proposed NDA-MBE, the derived NDA-MLE, and the derived NDA-CRLB in SISO flat-fading channel for BPSK signal, $N = 10$, and at $\gamma = 20$ dB.⁷ As expected, the NDA-MBE does not

⁷The complexity order of the derived NDA-MLE and NDA-CRLB are in the order of $\mathcal{O}(|M|^{N' n_t})$; for large values of N' ($N' = N + L - 1$) and $n_t > 1$, the corresponding curves are not obtainable. Hence, $N = 10$ and SISO flat-fading channel considered.

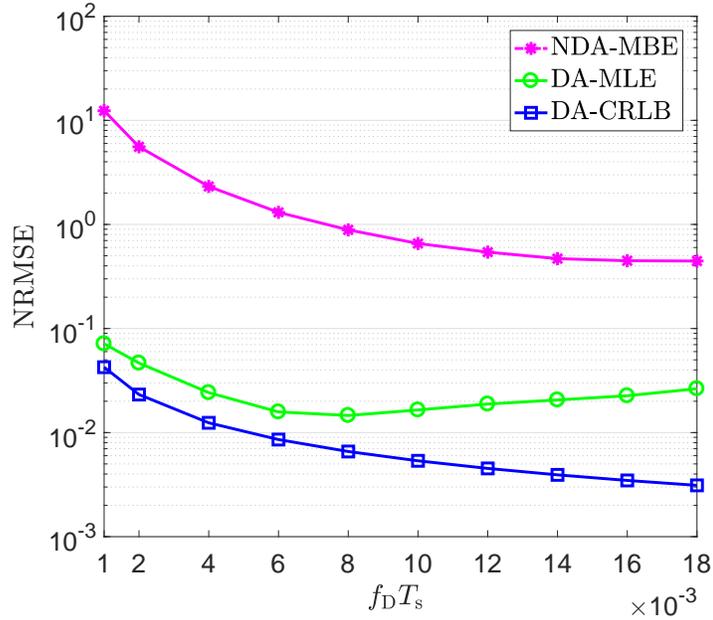


Fig. 3.11: Performance comparison of the NDA-MBE, the derived DA-CRLB, and the derived DA-MLE in MIMO frequency-selective Rayleigh fading channel for $n_t = 2$, $n_r = 2$, $L = 5$, $N = 10^3$, and at $\gamma = 10$ dB.

exhibit good performance for a short observation window size because the second- and fourth-order statistics employed in (3.38) are not accurately estimated. On the other hand, the derived NDA-MLE exhibits low NRMSE even for a short observation window. Moreover, there is no significant performance gap between the derived NDA-MLE and NDA-CRLB.

3.5 Conclusions and Directions for Future Research

In Chapter 3 of this thesis, MDS estimation in MIMO frequency-selective fading channel was studied. In this section, a summary of the main results in Chapter 3 is provided, and possible directions for future research are pointed out.

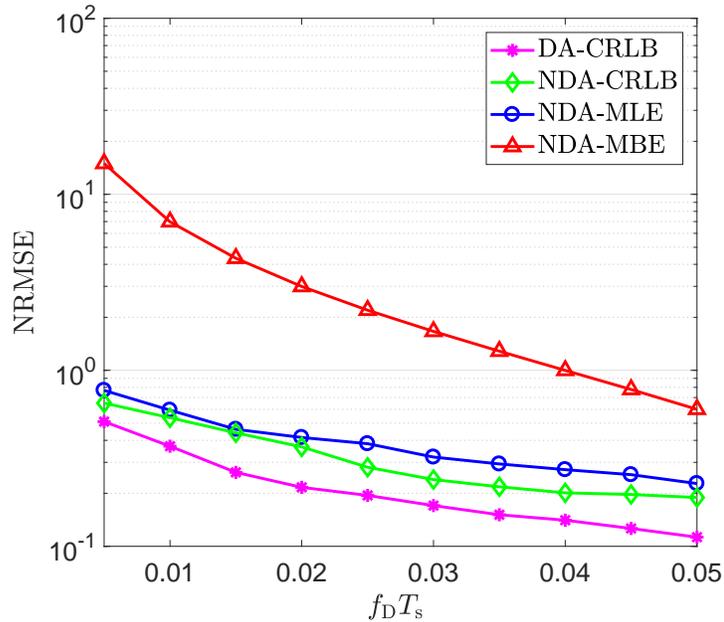


Fig. 3.12: Performance comparison of the proposed NDA-MBE, derived NDA-MLE, derived NDA-CRLB, and DA-CRLB in SISO flat-fading channel for $N = 10$ and at $\gamma = 20$ dB.

3.5.1 Summary

In Chapter 3, the DA and NDA CRLBs and MLEs for the MDS in MIMO frequency-selective fading channel were derived. Moreover, a low-complexity NDA-MBE was proposed. The NDA-MBE employs the statistical moment-based approach and relies on the second- and fourth-order statistics of the received signal, as well as the LS curve-fitting optimization technique to estimate the normalized squared AF and MDS of the fading channel. Compared to the existing DA estimators, the proposed NDA-MBE provides higher system capacity due to the absence of pilots. Also, the substantial reduced-complexity enables the proposed MBE to exhibit good performance in the NDA scenarios, where the observation window can be selected large enough. Simulation results show that there is no significant performance gap between the derived NDA-MLE and NDA-CRLB when the observation window is relatively

small. Furthermore, the significant reduced-complexity in the NDA-MBE leads to low NRMSE over a wide range of MDSs when the observation window is selected large enough. The proposed NDA-MBE exhibits the following advantages:

- It exhibits lower computational complexity compared to the MLEs;
- It does not require time synchronization;
- It is robust to the CFO;
- It increases system capacity;
- It does not require *a priori* knowledge of noise power, signal power, and channel delay profile;
- It does not require *a priori* knowledge of the number of transmit antennas;
- It removes the need of joint parameter estimation, such as CFO, signal power, noise power, and channel delay profile estimation.

3.5.2 Future research

The results in Chapter 3 of this thesis open interesting directions for a number of future research topics. Next, we outline two of them, as follows:

- The first one is an extension of the proposed NDA-MBE for MDS to an NDA joint channel delay profile and MDS estimator in MIMO frequency-selective channel;
- The second one pertains to developing an NDA-MDS estimator for multiple users.

Appendix

3.A Proof of (3.28)

To obtain an explicit closed-form expression for $\kappa_u^{(n)} \triangleq \mathbb{E}\{|r_k^{(n)}|^2 |r_{k+u}^{(n)}|^2\}$, we first write $|r_k^{(n)}|^2 = r_k^{(n)} (r_k^{(n)})^*$ by employing (3.1) as

$$\begin{aligned}
|r_k^{(n)}|^2 &= \sum_{m=1}^{n_t} \sum_{l=1}^L |h_{k,l}^{(mn)}|^2 |s_{k-l}^{(m)}|^2 + \sum_{m_1=1}^{n_t} \sum_{m_2 \neq m_1}^{n_t} \sum_{l=1}^L h_{k,l}^{(m_1 n)} (h_{k,l}^{(m_2 n)})^* s_{k-l}^{(m_1)} (s_{k-l}^{(m_2)})^* \quad (3.53) \\
&+ \sum_{m=1}^{n_t} \sum_{l_1=1}^L \sum_{l_2 \neq l_1}^L h_{k,l_1}^{(mn)} (h_{k,l_2}^{(mn)})^* s_{k-l_1}^{(m)} (s_{k-l_2}^{(m)})^* \\
&+ \sum_{m_1=1}^{n_t} \sum_{m_2 \neq m_1}^{n_t} \sum_{l_1=1}^L \sum_{l_2 \neq l_1}^L h_{k,l_1}^{(m_1 n)} (h_{k,l_2}^{(m_2 n)})^* s_{k-l_1}^{(m_1)} (s_{k-l_2}^{(m_2)})^* \\
&+ (w_k^{(n)})^* \sum_{m=1}^{n_t} \sum_{l=1}^L h_{k,l}^{(mn)} s_{k-l}^{(m)} + (w_k^{(n)}) \sum_{m=1}^{n_t} \sum_{l=1}^L (h_{k,l}^{(mn)})^* (s_{k-l}^{(m)})^* + |w_k^{(n)}|^2.
\end{aligned}$$

Then, $|r_{k+u}^{(n)}|^2$ is straightforwardly calculated by replacing k with $k+u$ in (3.53), and $|r_k^{(n)}|^2 |r_{k+u}^{(n)}|^2$ can be easily expressed in a summation form, which is omitted due to space constraints.

As fading is independent of the signal and noise, the statistical expectation in $\kappa_u^{(n)}$ can be decomposed into statistical expectations over the signal, fading, and noise distributions, respectively. With independent and identically distributed (i.i.d.) symbols, $s_k^{(m)}$, $k = 1, \dots, N$, and $u \geq L$, the symbols from the m th antenna contributed in $r_k^{(n)}$,

i.e., $\{s_{k-l}^{(m)}\}_{l=1}^L$, are different from those contributed in $r_{k+u}^{(n)}$, i.e., $\{s_{k+u-l}^{(m)}\}_{l=1}^L$; furthermore, by using that $\mathbb{E}\{(s_k^{(m)})^2\} = 0$, $\mathbb{E}\{s_k^{(m)}\} = 0$, and the linearity property of the statistical expectation, one obtains $\kappa_u^{(n)}$ as in (3.54). Finally, with $\mathbb{E}\{|s_k^{(m)}|^2\} = \sigma_{s_m}^2$ and $\mathbb{E}\{|w_k^{(n)}|^2\} = \sigma_{w_n}^2$, (3.28) is obtained.

$$\begin{aligned}
\kappa_u^{(n)} &= \sum_{m=1}^{n_t} \sum_{l=1}^L \mathbb{E} \left\{ |h_{k,l}^{(mn)}|^2 |h_{k+u,l}^{(mn)}|^2 \right\} \mathbb{E} \left\{ |s_{k-l}^{(m)}|^2 |s_{k+u-l}^{(m)}|^2 \right\} \\
&+ \sum_{m_1=1}^{n_t} \sum_{m_2 \neq m_1}^{n_t} \sum_{l=1}^L \mathbb{E} \left\{ |h_{k,l}^{(m_1 n)}|^2 |h_{k+u,l}^{(m_2 n)}|^2 \right\} \mathbb{E} \left\{ |s_{k-l}^{(m_1)}|^2 |s_{k+u-l}^{(m_2)}|^2 \right\} \\
&+ \sum_{m=1}^{n_t} \sum_{l_1=1}^L \sum_{l_2 \neq l_1}^L \mathbb{E} \left\{ |h_{k,l_1}^{(mn)}|^2 |h_{k+u,l_2}^{(mn)}|^2 \right\} \mathbb{E} \left\{ |s_{k-l_1}^{(m)}|^2 |s_{k+u-l_2}^{(m)}|^2 \right\} \\
&+ \sum_{m_1=1}^{n_t} \sum_{m_2 \neq m_1}^{n_t} \sum_{l_1=1}^L \sum_{l_2 \neq l_1}^L \mathbb{E} \left\{ |h_{k,l_1}^{(m_1 n)}|^2 |h_{k+u,l_2}^{(m_2 n)}|^2 \right\} \mathbb{E} \left\{ |s_{k-l_1}^{(m_1)}|^2 |s_{k+u-l_2}^{(m_2)}|^2 \right\} \\
&+ \mathbb{E} \left\{ |w_{k+u}^{(n)}|^2 \right\} \sum_{m=1}^{n_t} \sum_{l=1}^L \mathbb{E} \left\{ |h_{k,l}^{(mn)}|^2 \right\} \mathbb{E} \left\{ |s_{k-l}^{(m)}|^2 \right\} \\
&+ \mathbb{E} \left\{ |w_k^{(n)}|^2 \right\} \sum_{m=1}^{n_t} \sum_{l=1}^L \mathbb{E} \left\{ |h_{k+u,l}^{(mn)}|^2 \right\} \mathbb{E} \left\{ |s_{k+u-l}^{(m)}|^2 \right\} + \mathbb{E} \left\{ |w_k^{(n)}|^2 \right\} \mathbb{E} \left\{ |w_{k+u}^{(n)}|^2 \right\}.
\end{aligned} \tag{3.54}$$

3.B Proof of (3.34)

By employing (3.33), one can write

$$\begin{aligned}
\left(\mu_2^{(n)}\right)^2 &= \left(\sum_{m=1}^{n_t} \sum_{l=1}^L \sigma_{h_{(mn),l}}^2 \sigma_{s_m}^2 + \sigma_{w_n}^2 \right)^2 = \sum_{m=1}^{n_t} \sum_{l=1}^L \sigma_{h_{(mn),l}}^4 \sigma_{s_m}^4 \\
&+ \sum_{m_1=1}^{n_t} \sum_{m_2 \neq m_1}^{n_t} \sum_{l=1}^L \sigma_{h_{(m_1 n),l}}^2 \sigma_{h_{(m_2 n),l}}^2 \sigma_{s_{m_1}}^2 \sigma_{s_{m_2}}^2 + \sum_{m=1}^{n_t} \sum_{l_1=1}^L \sum_{l_2 \neq l_1}^L \sigma_{h_{(mn),l_1}}^2 \sigma_{h_{(mn),l_2}}^2 \sigma_{s_m}^4 \\
&+ \sum_{m_1=1}^{n_t} \sum_{m_2 \neq m_1}^{n_t} \sum_{l_1=1}^L \sum_{l_2 \neq l_1}^L \sigma_{h_{(m_1 n),l_1}}^2 \sigma_{h_{(m_2 n),l_2}}^2 \sigma_{s_{m_1}}^2 \sigma_{s_{m_2}}^2 + 2\sigma_{w_n}^2 \sum_{m=1}^{n_t} \sum_{l=1}^L \sigma_{h_{(mn),l}}^2 \sigma_{s_m}^2 + \sigma_{w_n}^4.
\end{aligned} \tag{3.55}$$

Then, by subtracting $(\mu_2^{(n)})^2$ in (3.55) from $\kappa_u^{(n)}$ in (3.32), one obtains

$$\kappa_u^{(n)} - (\mu_2^{(n)})^2 = \frac{J_0^2(2\pi f_D T_s u)}{\eta^{(n)}}, \quad (3.56)$$

where $\eta^{(n)} = 1 / \sum_{m=1}^{n_t} \sum_{l=1}^L \sigma_{h_{(mn),l}}^4 \sigma_{s_m}^4$.

3.C Proof of (3.35)

With independent fading, noise, and signal processes, by using (3.53) and then (3.31) and that $\mathbb{E}\{(s_k^{(m)})^2\} = \mathbb{E}\{s_k^{(m)}\} = \mathbb{E}\{(h_{k,l}^{(mn)})^2\} = \mathbb{E}\{w_k^{(n)}\} = 0$, similar to Appendix 3.A, one obtains

$$\begin{aligned} \mu_4^{(n)} = \mathbb{E}\{|r_k^{(n)}|^4\} &= 2 \sum_{m=1}^{n_t} \sum_{l=1}^L \sigma_{h_{(mn),l}}^4 \sigma_{s_m}^4 + 2 \sum_{m=1}^{n_t} \sum_{l=1}^L \sigma_{h_{(mn),l}}^4 (\Omega_s - 1) \sigma_{s_m}^4 \\ &+ 2 \sum_{m_1=1}^{n_t} \sum_{m_2 \neq m_1}^{n_t} \sum_{l=1}^L \sigma_{h_{(m_1 n),l}}^2 \sigma_{h_{(m_2 n),l}}^2 \sigma_{s_{m_1}}^2 \sigma_{s_{m_2}}^2 + 2 \sum_{m=1}^{n_t} \sum_{l_1=1}^L \sum_{l_2 \neq l_1}^L \sigma_{h_{(mn),l_1}}^2 \sigma_{h_{(mn),l_2}}^2 \sigma_{s_m}^4 \\ &+ 2 \sum_{m_1=1}^{n_t} \sum_{m_2 \neq m_1}^{n_t} \sum_{l_1=1}^L \sum_{l_2 \neq l_1}^L \sigma_{h_{(m_1 n),l_1}}^2 \sigma_{h_{(m_2 n),l_2}}^2 \sigma_{s_{m_1}}^2 \sigma_{s_{m_2}}^2 + 4\sigma_{w_n}^2 \sum_{m=1}^{n_t} \sum_{l=1}^L \sigma_{h_{(mn),l}}^2 \sigma_{s_m}^2 + 2\sigma_{w_n}^4, \end{aligned} \quad (3.57)$$

where Ω_s is the fourth-order two-conjugate statistic for unit variance signal, which represents the effect of the modulation format. Finally, by employing (3.33) and (3.57), (3.35) is easily obtained.

Chapter 4

Antenna Enumeration in Time-varying Fading Channels

4.1 Introduction

In many signal processing problems, the vector of observations can be modeled as the superposition of a finite number of signals embedded in noise, such as in multiuser detection [138, 139], blind source separation [140], smart antenna [141], massive MIMO 5G [142, 143], radar and sonar systems, flowing sensor networks [144], sensor array processing [13, 15, 145, 146], seismology and tomography. In these problems, determining the number of signal sources might be a critical first step. For example, the performance of the high-resolution methods such as estimation of signal parameters via rotational invariance techniques (ESPRIT) [147] and multiple signal classifier (MUSIC)[148] in conjunction with direction-of-arrival (DOA) estimation in array signal processing essentially relies on the prior knowledge of the number of signals. Multiuser detectors require the knowledge about the number of signals to exhibit an acceptable performance in terms of bit-error-rate (BER) [139]. A key and primary

issue in blind source separation is the estimation of the number of unknown sources from the mixed signals before an effective source separation.

The process of determining the number of signal sources is called source enumeration [145, 146, 149]. In the context of multiple-antenna systems, in which synchronized transmission is considered, it is referred to the process of counting the number of transmit antennas, and it is called antenna enumeration/number of transmit antennas detection in the literature [6, 150]. Antenna enumeration has important applications in adaptive wireless systems employing multiple antennas, such as cognitive radio, software-defined radio, and smart antennas [6, 150–156]. In the context of cognitive radio systems, the knowledge of the secondary users (SUs) about the number of primary users' (PUs) transmit antennas permits them to adjust their transmit power and beamforming techniques to avoid interference to the PUs; thus, this knowledge ameliorates the coexistence of the SUs and PUs equipped with multiple antennas [152]. Moreover, since multiple radio frequency (RF) chains associated with multiple antennas are costly, antenna selection techniques have recently been considered as a low-cost low-complexity practical alternative [157–160]. Accordingly, detecting the number of transmit antennas is of interest for receivers to eliminate the need for additional signaling, which introduces overhead and transmission latency [154–156].

4.1.1 Literature Review

The problem of detecting the number of transmit antennas employing multiple receive antennas has been explored in the literature [6, 150–156]. Two commonly used approaches are information-theoretic [6, 150, 151] and feature-based (FB) [152–156]. With the former approach, the problem of the number of transmit antennas detection is formulated as the rank estimation of the received signal correlation matrix. In [6], the minimum description length (MDL) and Akaike information criterion (AIC) algo-

rithms are employed. While MDL provides an asymptotically consistent estimate, it underestimates the number of transmit antennas for small sample sizes. On the other hand, AIC does not provide an asymptotically consistent estimate and overestimates the number of transmit antennas for large sample sizes. Both algorithms require the eigen-decomposition of the sample covariance matrix. Additionally, they are sensitive to both timing offset (TO) and carrier frequency offset (CFO) [6]. In [150], the authors consider the number of transmit antennas detection as an adaptive rank estimation problem, employing the *Schur* complement test. The algorithm detects the number of transmit antennas without tracking the eigenvalues of the covariance matrix and thus has a lower computational complexity. It is to be noted that the information-theoretic approaches require a number of receive antennas larger than the number of transmit antennas, which is not always the case in practice.

On the other hand, the FB approach relies on features extracted from the received signal, based on which a decision is made. Hence, the FB approach requires *a priori* information about the transmitted signals. In [152] and [153], the orthogonality of the LTE pilot patterns is exploited as a feature, whereas the orthogonality of the LTE preamble sequences related to each antenna is used in [154] and [155]. The algorithms in [154] and [155] require *a priori* knowledge about the pilot patterns and preamble sequences, respectively. It is worth noting that the algorithms in [152] and [153] are sensitive to both TO and CFO, the one in [154] is robust to CFO and sensitive to TO, and the one in [155] is robust to both transmission impairments. When compared with the information-theoretic algorithms, the FB algorithms are usually simpler to implement, and they may detect the number of transmit antennas by employing a reduced number of receive antennas when compared with the number of transmit antennas.

4.1.2 Motivation

As mentioned in the previous section, antenna enumeration has been addressed by employing both ICT and FB approaches in the literature. After reviewing the current antenna enumeration algorithms, the following were made:

- Antenna enumeration for time-varying fading channel has not been investigated;
- New FB algorithms should be developed to fill in the gap between ICT and FB approaches, which does not depend on the signal type and can detect the number of transmit antennas with reduced number of receive antennas, as well as reduced complexity;
- There is no analytical performance analysis for the ICT and FB approaches;
- The existing antenna enumeration algorithms are severely affected by transmission imperfections, such as TO and CFO;
- Both ICT and FB approaches require knowledge of the maximum number of transmit antennas.

4.1.3 Problem Statement

The specific research problems which are studied in Chapter 4 of this thesis are presented as follows:

- The problem of antenna enumeration in time-varying fading channel is investigated, and two low-complexity blind and semi-blind antenna enumeration algorithms are proposed. The proposed blind algorithm does not depend on the signal type, pilot and preamble patterns, and can detect the number of transmit antennas with a single receive antennas. However, the semi-blind algorithm requires the modulation format for antenna enumeration.

- The analytical performance analysis of the proposed blind algorithm, i.e., a closed-form expression for the probability of correct detection is derived;
- Finally, the performance of the proposed algorithms is studied under various transceiver and channel impairments.

4.1.4 Methodology

The statistical moment-based (MB) approach is employed in Chapter 4 of this thesis. The proposed algorithms rely on the second- and fourth-order statistics of the received signal and use the time-varying nature of the fading channel to detect the number of transmit antennas.

4.2 Antenna Enumeration

4.2.1 System model

We consider a multiple-input single-output (MISO) block fading channel [161, 162] with n_t transmit antennas, where n_t is unknown at the receive side. It is assumed that N_b observation blocks, $b = 1, 2, \dots, N_b$, each with length N_c symbols, are affected by independent and identically distributed (i.i.d.) fading characterized by the $(n_t \times 1)$ vector \mathbf{h}_b and corrupted by additive white Gaussian noise. A typical value for the block length, N_c , in the case of the Clarke-Jakes Doppler spectrum is $N_c = \lfloor 9/16\pi f_d T_s \rfloor$ [163], where f_d and T_s are the maximum Doppler frequency and the symbol period, respectively. Thus, the received complex-valued signal can be expressed as

$$r_{k,b} = \mathbf{h}_b^\dagger \mathbf{s}_{k,b} + w_{k,b}, \quad k = 1, 2, \dots, N_c, \quad b = 1, 2, \dots, N_b \quad (4.1)$$

where $r_{k,b}$ is the k th received symbol in the b th observation block; $\mathbf{s}_{k,b} = [s_{k,b}^{(1)}, s_{k,b}^{(2)}, \dots, s_{k,b}^{(n_t)}]^\dagger$ represents the zero-mean uncorrelated transmitted signals from the n_t transmit antennas; whose variance $\mathbb{E}\{|s_{k,b}^{(m)}|^2\} = \sigma_s^2$, $m = 1, 2, \dots, n_t$ is unknown at the receive side; $w_{k,b}$ is circular complex additive white Gaussian noise with variance σ_w^2 , which is assumed to be known at the receive side¹; and $\mathbf{h}_b = [h_b^{(1)}, h_b^{(2)}, \dots, h_b^{(n_t)}]^\dagger$ denotes the channel coefficients, with $h_b^{(m)}$, $m = 1, 2, \dots, n_t$ as the channel coefficient between the m th transmit antenna and the receive antenna for the b th observation block. It is assumed that the channel coefficients in each block are uncorrelated circular complex Gaussian random variables with $\mathbb{E}\{h_b^{(m_1)}(h_b^{(m_2)})^*\} = \sigma_h^2 \delta(m_1 - m_2)$, where σ_h^2 is unknown at the receive side.

4.2.2 Blind Antenna Enumeration Algorithm

Let us first consider the second-order moment of the received signal in an observation block, which is easily obtained as

$$\boldsymbol{\mu}_b \triangleq \mathbb{E}\left\{|r_{k,b}|^2 \mid \mathbf{h}_b\right\} = \sigma_s^2 \sum_{m=1}^{n_t} |h_b^{(m)}|^2 + \sigma_w^2. \quad (4.2)$$

Further, let us define

$$\mu^{(d)} \triangleq \mathbb{E}\left\{(\boldsymbol{\mu}_b)^d\right\} = \mathbb{E}\left\{\left(\mathbb{E}\left\{|r_{k,b}|^2 \mid \mathbf{h}_b\right\}\right)^d\right\} \quad (4.3)$$

where d is a positive integer. With the channel coefficients corresponding to different transmit antennas being independent circular complex Gaussian random variables with variance σ_h^2 , i.e., $\mathbb{E}\{\mathbf{h}_b \mathbf{h}_b^H\} = \sigma_h^2 \mathbf{I}$, and by using (4.2), when $d = 1$ (second-order

¹Note that estimators of the noise variance which exist in the literature, e.g., [5] can be employed.

statistics), one can easily write (4.3) as

$$\mu^{(1)} \triangleq \mathbb{E} \{ \mu_b \} = \sigma_s^2 \sum_{m=1}^{n_t} \mathbb{E} \left\{ \left| h_b^{(m)} \right|^2 \right\} + \sigma_w^2 = n_t \sigma_s^2 \sigma_h^2 + \sigma_w^2. \quad (4.4)$$

Furthermore, following the proof in Appendix 4.A, we obtain

$$\mu^{(2)} \triangleq \mathbb{E} \{ (\mu_b)^2 \} = (n_t^2 + n_t) \sigma_s^4 \sigma_h^4 + 2n_t \sigma_s^2 \sigma_h^2 \sigma_w^2 + \sigma_w^4. \quad (4.5)$$

By substituting (4.4) into (4.5), n_t can be expressed as

$$n_t = \frac{(\mu^{(1)} - \sigma_w^2)^2}{\mu^{(2)} - (\mu^{(1)})^2}. \quad (4.6)$$

It should be noted that unbiased estimates of the statistics $\mu^{(1)}$, $\zeta \triangleq (\mu^{(1)})^2$, and $\mu^{(2)}$, are used in practice to obtain n_t from (4.6). To estimate these statistics, the following unbiased estimators are employed²

$$\hat{\mu}^{(1)} = \frac{1}{N_b N_c} \sum_b^{N_b} \sum_k^{N_c} |r_{k,b}|^2 \quad (4.7)$$

$$\hat{\zeta} = \frac{1}{N_b(N_b - 1)N_c(N_c - 1)} \sum_{b_1:b_2}^{N_b} \sum_{k_1:k_2}^{N_c} |r_{k_1,b_1}|^2 |r_{k_2,b_2}|^2 \quad (4.8)$$

and

$$\hat{\mu}^{(2)} = \frac{1}{N_b N_c (N_c - 1)} \sum_b^{N_b} \sum_{k_1:k_2}^{N_c} |r_{k_1,b}|^2 |r_{k_2,b}|^2, \quad (4.9)$$

where

$$\sum_{f_1:f_i}^F \triangleq \sum_{f_1=1}^F \sum_{\substack{f_2=1 \\ f_2 \neq f_1}}^F \cdots \sum_{\substack{f_i=1 \\ f_i \neq \{f_1, \dots, f_{i-1}\}}}^F. \quad (4.10)$$

By using the linearity property of the statistical expectation, one can easily show

²It is worth noting that $(\hat{\mu}^{(1)})^2$ cannot be employed for the estimation of ζ , as it results in a biased estimator.

that the employed estimators are unbiased, i.e., $\mathbb{E}\{\hat{\mu}^{(1)}\} = \mu^{(1)}$, $\mathbb{E}\{\hat{\zeta}\} = (\mu^{(1)})^2$, and $\mathbb{E}\{\hat{\mu}^{(2)}\} = \mu^{(2)}$ (see Appendix 4.B for proof).

Substituting (4.7), (4.8), and (4.9) in (4.6) yields the following decision statistics for the number of transmit antennas,

$$\Upsilon = \frac{\hat{\zeta} - 2\hat{\mu}^{(1)}\sigma_w^2 + \sigma_w^4}{\hat{\mu}^{(2)} - \hat{\zeta}}. \quad (4.11)$$

Although Υ can be seen as providing an estimate of n_t , it is actually a continuous-valued random variable. In order to determine the number of transmit antennas, which has discrete values, we formulate the problem as a multiple-hypothesis testing problem, i.e., under hypothesis H_i the number of transmit antenna equals $i = 1, 2, \dots$, and compare the test statistic Υ against thresholds to make a decision.

By employing first-order Taylor expansion of Υ , the mean of Υ can be expressed according to Appendix 4.C as

$$\mathbb{E}\{\Upsilon\} \approx \frac{\mathbb{E}\{\hat{\zeta} - 2\hat{\mu}^{(1)}\sigma_w^2 + \sigma_w^4\}}{\mathbb{E}\{\hat{\mu}^{(2)} - \hat{\zeta}\}} = \frac{(\mu^{(1)} - \sigma_w^2)^2}{\mu^{(2)} - (\mu^{(1)})^2} = n_t. \quad (4.12)$$

Hence, to detect the number of transmit antennas, decision regions for the hypothesis testing can be set as

$$\hat{n}_t = i \quad \text{when} \quad \Gamma_{i-1} < \Upsilon \leq \Gamma_i, \quad i = 1, 2, \dots \quad (4.13)$$

with $\Gamma_0 = -\infty$ and $\Gamma_i \in (i, i + 1)$, $i = 1, 2, \dots$. Details on the thresholds setting are provided in Section 4.3.2. A formal description of the proposed blind antenna enumeration algorithm is presented in Algorithm 10.

Algorithm 10 Blind Antenna Enumeration

Input: $r_{k,b}$, $k = 1, 2, \dots, N_c$, $b = 1, 2, \dots, N_b$, σ_w^2 , $\Gamma_0 = -\infty$, $\Gamma_i \in (i, i + 1)$, $i = 1, 2, \dots$

Output: Number of transmit antennas \hat{n}_t

Initialization: $i = 1$

- 1: Obtain $\hat{\mu}^{(1)}$, $\hat{\zeta}$, and $\hat{\mu}^{(2)}$ by employing (4.7), (4.8), and (4.9)
 - 2: Compute the decision statistic Υ according to (4.11)
 - 3: Set the threshold value Γ_i
 - 4: **if** $\Gamma_{i-1} < \Upsilon \leq \Gamma_i$ **then**
 - 5: $\hat{n}_t = i$
 - 6: **else**
 - 7: Increment $i = i + 1$ and go to step 3
 - 8: **end if**
-

4.2.3 Semi-blind Antenna Enumeration Algorithm

Let us consider

$$\phi_b \triangleq \varphi_b - 2\mu_b^2, \quad (4.14)$$

where

$$\varphi_b \triangleq \mathbb{E} \left\{ |r_{k,b}|^4 | \mathbf{h}_b \right\}, \quad (4.15)$$

and μ_b is given in (4.2). Analogous to the proposed blind antenna enumeration algorithm, one can easily obtain

$$\Phi_b = \Omega_s \sigma_s^4 \sum_{m=1}^{n_t} |h_b^{(m)}|^4, \quad (4.16)$$

and

$$\phi^{(1)} \triangleq \mathbb{E} \{ \Phi_b \} = \Omega_s \sigma_s^4 \sum_{m=1}^{n_t} \mathbb{E} \{ |h_b^{(m)}|^4 \} = 2\Omega_s \sigma_s^4 \sigma_h^4 n_t, \quad (4.17)$$

where $1 \leq \Omega_s \leq 2$ denotes the fourth-order/ two conjugate statistic for unit variance constellations.³

With the assumption that Ω_s , i.e., modulation format is known at the receive-side, by employing (4.4) and (4.17), n_t can be straightforwardly expressed as

$$n_t = \frac{2\Omega_s(\mu^{(1)} - \sigma_w^2)^2}{\phi^{(1)}} = 2\Omega_s \frac{\zeta + \sigma_w^4 - 2\mu^{(1)}\sigma_w^2}{\phi^{(1)}}. \quad (4.18)$$

By replacing ζ and $\mu^{(1)}$ in (4.18) with $\hat{\zeta}$ and $\hat{\mu}^{(1)}$ in (4.7) and (4.8), and $\phi^{(1)}$ with

$$\hat{\phi}^{(1)} = \frac{1}{N_b N_c} \sum_b^{N_b} \sum_k^{N_c} |r_{k,b}|^4 - \frac{2}{N_b N_c (N_c - 1)} \sum_{b_1}^{N_b} \sum_{k_1:k_2}^{N_c} |r_{k_1,b_1}|^2 |r_{k_2,b_1}|^2, \quad (4.19)$$

the following decision metric for the number of transmit antennas is obtained.

$$\Theta = 2\Omega_s \frac{\hat{\zeta} + \sigma_w^4 - 2\hat{\mu}^{(1)}\sigma_w^2}{\hat{\phi}^{(1)}}. \quad (4.20)$$

Similar to the blind algorithm, we formulate the antenna enumeration problem as a multiple-hypothesis testing problem, and compare the decision statistics Θ against thresholds to determine the number of transmit antennas.

By employing first-order Taylor expansion, similar to that in Appendix 4.C, the mean of Θ is expressed as

$$\mathbb{E}\{\Theta\} \approx 2\Omega_s \frac{\mathbb{E}\{\hat{\zeta} + \sigma_w^4 - 2\hat{\mu}^{(1)}\sigma_w^2\}}{\mathbb{E}\{\hat{\phi}^{(1)}\}} = n_t. \quad (4.21)$$

Based on (4.21), decision regions for the hypothesis testing problem can be set as

$$\hat{n}_t = i \quad \text{when} \quad \Gamma'_{i-1} < \Theta \leq \Gamma'_i, \quad i = 1, 2, \dots, \quad (4.22)$$

³For 16-QAM, 64-QAM, and complex-valued zero-mean Gaussian signals, Ω_s is 1.32, 1.38, and 2, respectively [127].

with $\Gamma'_0 = -\infty$ and $\Gamma'_i \in (i, i + 1)$, $i = 1, 2, \dots$. A formal description of the proposed semi-blind antenna enumeration algorithm is presented in Algorithm 11.

Algorithm 11 Semi-blind Antenna Enumeration

- Input:** $r_{k,b}$, $k = 1, 2, \dots, N_c$, $b = 1, 2, \dots, N_b$, σ_w^2 , Ω_s , $\Gamma'_0 = -\infty$, $\Gamma'_i \in (i, i + 1)$, $i = 1, 2, \dots$
- Output:** Number of transmit antennas \hat{n}_t
- Initialization:** $i = 1$
- 1: Obtain $\hat{\mu}^{(1)}$, $\hat{\zeta}$, and $\hat{\phi}^{(1)}$ by employing (4.7), (4.8), and (4.19)
 - 2: Compute the decision statistic Θ according to (4.20)
 - 3: Set the threshold value Γ'_i
 - 4: **if** $\Gamma'_{i-1} < \Theta \leq \Gamma'_i$ **then**
 - 5: $\hat{n}_t = i$
 - 6: **else**
 - 7: Increment $i = i + 1$ and go to step 3
 - 8: **end if**
-

4.3 Performance Analysis and Threshold setting

This section derives a closed-form expression for the conditional probability of correctly detecting the number of transmit antennas, $\mathbb{P}\{\hat{n}_t = n_t | H_{n_t}\}$, $n_t = 1, 2, \dots$ for the proposed blind antenna enumeration algorithm in 10. Also, the optimal thresholds that maximize the probability of correct detection are obtained.

4.3.1 Performance Analysis

From (4.11), the conditional probability of correct detection can be expressed as

$$\mathbb{P}\{\hat{n}_t = n_t | H_{n_t}\} = \mathbb{P}\{\Gamma_{n_t-1} < \Upsilon \leq \Gamma_{n_t} | H_{n_t}\} = F_{\Upsilon|H_{n_t}}(\Gamma_{n_t}) - F_{\Upsilon|H_{n_t}}(\Gamma_{n_t-1}) \quad (4.23)$$

where $F_{\Upsilon|H_{n_t}}(\cdot)$ is the cumulative distribution function (CDF) of the random variable Υ when the number of transmit antennas is n_t . As can be seen in (4.11), the numerator

and denominator of Υ are correlated, and the probability density function (PDF) of Υ cannot be expressed straightforwardly; thus, obtaining the conditional probability of correct detection via the CDF of Υ is difficult. However, the denominator in the right-hand side of (4.11), i.e., $\hat{\mu}^{(2)} - \hat{\zeta}$, is positive since it represents the sample variance of $\hat{\mu}_b$, and hence one can write (4.13) as an equivalent two-dimensional test, i.e.,

$$\hat{n}_t = n_t \quad \text{when} \quad (x_{n_t-1} < 0 \cap x_{n_t} \geq 0), \quad n_t = 1, 2, \dots \quad (4.24)$$

where \cap is the intersection operator and

$$x_\ell = \Gamma_\ell \hat{\mu}^{(2)} - (\Gamma_\ell + 1) \hat{\zeta} + 2\sigma_w^2 \hat{\mu}^{(1)} - \sigma_w^4, \quad \ell = n_t, n_t - 1. \quad (4.25)$$

As an example, the decision regions of the equivalent two-dimensional test for x_2 and x_3 under hypothesis H_3 and its corresponding correct and incorrect decisions at SNR per transmit antenna $\gamma \triangleq \sigma_h^2 \sigma_s^2 / \sigma_w^2 = -10$ dB are shown in Fig. 4.1. As can be seen, x_2 and x_3 are highly correlated and exhibit positive correlation such that their values increase or decrease together. The conditional probability of correct detection is thus given as

$$\mathbb{P}\{\hat{n}_t = n_t | H_{n_t}\} = \mathbb{P}\{x_{n_t-1} < 0 \cap x_{n_t} \geq 0 | H_{n_t}\}, \quad n_t = 1, 2, \dots \quad (4.26)$$

For hypothesis H_1 , since $\Gamma_0 = -\infty$, (4.26) yields

$$\mathbb{P}\{\hat{n}_t = 1 | H_1\} = \mathbb{P}\{x_1 \geq 0 | H_1\}. \quad (4.27)$$

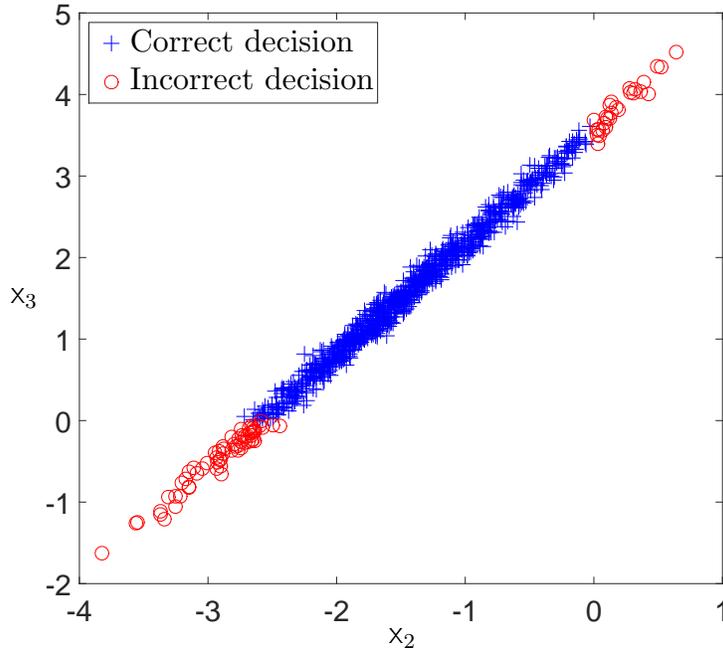


Fig. 4.1: The equivalent two-dimensional test under hypothesis H_3 for $\Gamma_2 = 2.5$ and $\Gamma_3 = 3.5$ at SNR $\gamma = -10$ dB.

Let us consider $y_\ell \triangleq \frac{x_\ell}{\sigma_s^4 \sigma_h^4}$, $\ell = n_t, n_t - 1$ ⁴ and re-write (4.26) as

$$\mathbb{P}\{\hat{n}_t = n_t | H_{n_t}\} = \mathbb{P}\{y_{n_t-1} < 0 \cap y_{n_t} \geq 0 | H_{n_t}\}, n_t = 1, 2, \dots \quad (4.28)$$

With the assumption that $N_c \gg 1$ and $N_b \gg 1$, from the central limit theorem [164],⁵ the joint distribution of the random variables y_{n_t} (x_{n_t}) and y_{n_t-1} (x_{n_t-1}) under hypothesis H_{n_t} can be approximated as a joint Gaussian distribution (see Fig. 4.2)

$$[y_{n_t}, y_{n_t-1} | H_{n_t}] \sim \mathcal{N}(\boldsymbol{\omega}_{n_t, n_t-1}, \mathbf{C}_{n_t, n_t-1}) \quad (4.29)$$

⁴ y_ℓ is used instead of x_ℓ in order to obtain expressions of the variance and covariance as a function of the SNR.

⁵Note that simulation results show that this assumption is valid even for relatively small N_c and N_b values, e.g., $N_c = 10$ and $N_b = 20$.

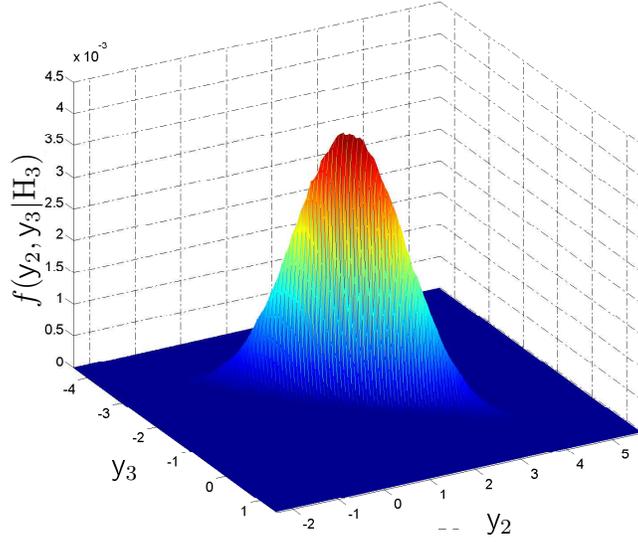


Fig. 4.2: Illustration of the estimated joint PDF of y_2 and y_3 under hypothesis H_3 , $f(y_2, y_3|H_3)$.

with the mean vector, $\boldsymbol{\omega}_{n_t, n_t-1} \triangleq [\omega_{y_{n_t}}, \omega_{y_{n_t-1}}]^T$ given by (see Appendix 4.D for proof)

$$\omega_{y_\ell} = \Gamma_\ell n_t - n_t^2, \quad \ell = n_t, n_t - 1 \quad (4.30)$$

and the covariance matrix \mathbf{C}_{n_t, n_t-1} as

$$\mathbf{C}_{n_t, n_t-1} = \begin{bmatrix} \sigma_{y_{n_t}}^2 & \rho_{y_{n_t} y_{n_t-1}} \sigma_{y_{n_t}} \sigma_{y_{n_t-1}} \\ \rho_{y_{n_t} y_{n_t-1}} \sigma_{y_{n_t}} \sigma_{y_{n_t-1}} & \sigma_{y_{n_t-1}}^2 \end{bmatrix}, \quad (4.31)$$

where

$$\rho_{y_{n_t} y_{n_t-1}} = \frac{\text{Cov}\{y_{n_t}, y_{n_t-1}\}}{\sigma_{y_{n_t}} \sigma_{y_{n_t-1}}} \quad (4.32)$$

and $\sigma_{y_\ell}^2$, $\ell = n_t, n_t - 1$ and $\text{Cov}\{y_{n_t}, y_{n_t-1}\}$ are given in (4.33) and (4.34), respectively

(see Appendix 4.E for proof).

$$\begin{aligned}
\sigma_{y_\ell}^2 = & \Gamma_\ell^2 \left[\beta_{20}^1 [2\alpha_{22}^2 \bar{\vartheta}^{(2)} + 4\alpha_{22}^3 \bar{\kappa}^{(3)} + \alpha_{22}^4 \bar{\mu}^{(4)}] + [\beta_{20}^2 (2\alpha_{22}^2 + 4\alpha_{22}^3 + \alpha_{22}^4) - 1] (\bar{\mu}^{(2)})^2 \right] \\
& + (\Gamma_\ell + 1)^2 \left[\beta_{22}^2 [2\alpha_{22}^2 ((\bar{\vartheta}^{(1)})^2 + (\bar{\mu}^{(2)})^2) + 4\alpha_{22}^3 \bar{\vartheta}^{(1)} \bar{\mu}^{(2)} + (4\alpha_{22}^3 + 2\alpha_{22}^4) (\bar{\mu}^{(2)})^2] \right. \\
& + 4\beta_{22}^3 [\alpha_{22}^2 (\bar{\vartheta}^{(1)} + \bar{\mu}^{(2)}) + \alpha_{22}^3 \bar{\vartheta}^{(1)} + (3\alpha_{22}^3 + \alpha_{22}^4) \bar{\mu}^{(2)}] (\bar{\mu}^{(1)})^2 \\
& + [\beta_{22}^4 (2\alpha_{22}^2 + 4\alpha_{22}^3 + \alpha_{22}^4) - 1] (\bar{\mu}^{(1)})^4 \left. + 4\gamma^{-2} [\beta_{20}^1 [\alpha_{20}^1 \bar{\vartheta}^{(1)} + \alpha_{20}^2 \bar{\mu}^{(2)}] + [\beta_{20}^2 (\alpha_{20}^1 + \alpha_{20}^2) - 1] (\bar{\mu}^{(1)})^2] \right. \\
& - 2\Gamma_\ell (\Gamma_\ell + 1) \left[\beta_{21}^2 [4(\alpha_{22}^2 + \alpha_{22}^3) \bar{\kappa}^{(2)} + (4\alpha_{22}^3 + 2\alpha_{22}^4) \bar{\mu}^{(3)}] \bar{\mu}^{(1)} + [\beta_{21}^3 (2\alpha_{22}^2 + 4\alpha_{22}^3 + \alpha_{22}^4) - 1] \bar{\mu}^{(2)} (\bar{\mu}^{(1)})^2 \right] \\
& + 4\Gamma_\ell \gamma^{-1} \left[\beta_{20}^1 (2\alpha_{21}^2 \bar{\kappa}^{(2)} + \alpha_{21}^3 \bar{\mu}^{(3)}) + [\beta_{20}^2 (2\alpha_{21}^2 + \alpha_{21}^3) - 1] \bar{\mu}^{(2)} \bar{\mu}^{(1)} \right] \\
& - 4(\Gamma_\ell + 1) \gamma^{-1} \left[2\beta_{21}^2 [(\alpha_{21}^2 (\bar{\vartheta}^{(1)} + \bar{\mu}^{(2)}) + \alpha_{21}^3 \bar{\mu}^{(2)}) \bar{\mu}^{(1)}] + [\beta_{21}^3 (2\alpha_{21}^2 + \alpha_{21}^3) - 1] (\bar{\mu}^{(1)})^3 \right] \quad \ell = n_t, n_t - 1
\end{aligned} \tag{4.33}$$

$$\text{Cov}\{y_{n_t}, y_{n_t-1}\} = \tag{4.34}$$

$$\begin{aligned}
& \Gamma_{n_t} \Gamma_{n_t-1} \left[\beta_{20}^1 [2\alpha_{22}^2 \bar{\vartheta}^{(2)} + 4\alpha_{22}^3 \bar{\kappa}^{(3)} + \alpha_{22}^4 \bar{\mu}^{(4)}] + [\beta_{20}^2 (2\alpha_{22}^2 + 4\alpha_{22}^3 + \alpha_{22}^4) - 1] (\bar{\mu}^{(2)})^2 \right] \\
& + (\Gamma_{n_t} \Gamma_{n_t-1} + \Gamma_{n_t} + \Gamma_{n_t-1} + 1) \left[\beta_{22}^2 [2\alpha_{22}^2 ((\bar{\vartheta}^{(1)})^2 + (\bar{\mu}^{(2)})^2) + 4\alpha_{22}^3 \bar{\vartheta}^{(1)} \bar{\mu}^{(2)} + (4\alpha_{22}^3 + 2\alpha_{22}^4) (\bar{\mu}^{(2)})^2] \right. \\
& + 4\beta_{22}^3 [\alpha_{22}^2 (\bar{\vartheta}^{(1)} + \bar{\mu}^{(2)}) + \alpha_{22}^3 \bar{\vartheta}^{(1)} + (3\alpha_{22}^3 + \alpha_{22}^4) \bar{\mu}^{(2)}] (\bar{\mu}^{(1)})^2 + [\beta_{22}^4 (2\alpha_{22}^2 + 4\alpha_{22}^3 + \alpha_{22}^4) - 1] (\bar{\mu}^{(1)})^4 \left. + 4\gamma^{-2} [\beta_{20}^1 [\alpha_{20}^1 \bar{\vartheta}^{(1)} + \alpha_{20}^2 \bar{\mu}^{(2)}] + [\beta_{20}^2 (\alpha_{20}^1 + \alpha_{20}^2) - 1] (\bar{\mu}^{(1)})^2] \right. \\
& - (2\Gamma_{n_t} \Gamma_{n_t-1} + \Gamma_{n_t} + \Gamma_{n_t-1}) \left[\beta_{21}^2 [4(\alpha_{22}^2 + \alpha_{22}^3) \bar{\kappa}^{(2)} + (4\alpha_{22}^3 + 2\alpha_{22}^4) \bar{\mu}^{(3)}] \bar{\mu}^{(1)} + [\beta_{21}^3 (2\alpha_{22}^2 + 4\alpha_{22}^3 + \alpha_{22}^4) - 1] \right. \\
& \left. \bar{\mu}^{(2)} (\bar{\mu}^{(1)})^2 \right] + 2(\Gamma_{n_t} + \Gamma_{n_t-1}) \gamma^{-1} \left[\beta_{20}^1 [2\alpha_{21}^2 \bar{\kappa}^{(2)} + \alpha_{21}^3 \bar{\mu}^{(3)}] + [\beta_{20}^2 (2\alpha_{21}^2 + \alpha_{21}^3) - 1] \bar{\mu}^{(2)} \bar{\mu}^{(1)} \right] \\
& - 2(\Gamma_{n_t} + \Gamma_{n_t-1} + 2) \gamma^{-1} \left[2\beta_{21}^2 [\alpha_{21}^2 (\bar{\vartheta}^{(2)} + \bar{\mu}^{(2)}) + \alpha_{21}^3 \bar{\mu}^{(2)}] \bar{\mu}^{(1)} + [\beta_{21}^3 (2\alpha_{21}^2 + \alpha_{21}^3) - 1] (\bar{\mu}^{(1)})^3 \right].
\end{aligned}$$

In (4.33) and (4.34), $\alpha_{ij}^L \triangleq \frac{\prod_{l=1}^L (N_c - l + 1)}{N_c^i (N_c - 1)^j}$, $\beta_{ij}^L \triangleq \frac{\prod_{l=1}^L (N_b - l + 1)}{N_b^i (N_b - 1)^j}$, $L = 1, 2, 3, 4$, and Ω_s

is the fourth-order/ two-conjugate moment of the unit variance constellations; the

expressions for other symbols are provided in (4.35).

$$\begin{aligned}
\bar{\mu}^{(1)} &= n_t + \gamma^{-1} \\
\bar{\mu}^{(2)} &= (n_t^2 + n_t) + 2n_t\gamma^{-1} + \gamma^{-2} \\
\bar{\mu}^{(3)} &= (n_t^3 + 3n_t^2 + 2n_t) + 3(n_t^2 + n_t)\gamma^{-1} + 3n_t\gamma^{-2} + \gamma^{-3} \\
\bar{\mu}^{(4)} &= (n_t^4 + 6n_t^3 + 11n_t^2 + 6n_t) + 4(n_t^3 + 3n_t^2 + 2n_t)\gamma^{-1} \\
&\quad + 6(n_t^2 + n_t)\gamma^{-2} + 4n_t\gamma^{-3} + \gamma^{-4} \\
\bar{\vartheta}^{(1)} &= 2(n_t^2 + (\Omega_s - 1)n_t) + 4n_t\gamma^{-1} + 2\gamma^{-2} \\
\bar{\vartheta}^{(2)} &= 4n_t^4 + 8(\Omega_s + 1)n_t^3 + (4\Omega_s^2 + 24\Omega_s - 20)n_t^2 \\
&\quad + (20\Omega_s^2 - 32\Omega_s + 8)n_t + [16n_t^3 + 16(\Omega_s + 1)n_t^2 \\
&\quad + 32(\Omega_s - 1)n_t]\gamma^{-1} + [24n_t^2 + 8(\Omega_s + 1)n_t]\gamma^{-2} \\
&\quad + 16n_t\gamma^{-3} + 4\gamma^{-4} \\
\bar{\kappa}^{(2)} &= 2n_t^3 + 2(\Omega_s + 1)n_t^2 + (4\Omega_s - 4)n_t \\
&\quad + [6n_t^2 + 2(\Omega_s + 1)n_t]\gamma^{-1} + 6n_t\gamma^{-2} + 2\gamma^{-3} \\
\bar{\kappa}^{(3)} &= 2n_t^4 + (8 + 2\Omega_s)n_t^3 + (2 + 10\Omega_s)n_t^2 + 12(\Omega_s - 1)n_t \\
&\quad + [8n_t^3 + 4(4 + \Omega_s)n_t^2 + 8\Omega_s n_t]\gamma^{-1} \\
&\quad + [12n_t^2 + (8 + 2\Omega_s)n_t]\gamma^{-2} + 8n_t\gamma^{-3} + 2\gamma^{-4}. \tag{4.35}
\end{aligned}$$

Using (4.29)–(4.35), the conditional probability of correct detection in (4.28) is obtained as

$$\mathbb{P}\{\hat{n}_t = n_t | \mathbf{H}_{n_t}\} = Q\left(\frac{\omega_{y_{n_t-1}}}{\sigma_{y_{n_t-1}}}, \frac{-\omega_{y_{n_t}}}{\sigma_{y_{n_t}}}, -\rho_{y_{n_t}y_{n_t-1}}\right), \tag{4.36}$$

where

$$Q(x, y, \rho) = \frac{1}{2\pi\sqrt{1-\rho^2}} \int_x^\infty \int_y^\infty \exp\left(-\frac{x_1^2 + y_1^2 - 2\rho x_1 y_1}{2(1-\rho^2)}\right) dx_1 dy_1 \tag{4.37}$$

represents the two-dimensional Gaussian Q -function [165]. Under H_1 , the two-dimensional

Gaussian Q -function is replaced with the one-dimensional Q -function, yielding

$$\mathbb{P}\{\hat{n}_t = 1 | \mathbf{H}_1\} = Q\left(\frac{-\omega_{y_1}}{\sigma_{y_1}}\right) \quad (4.38)$$

where $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp\left(-\frac{z^2}{2}\right) dz$ is the one-dimensional Gaussian Q -function.

4.3.2 Threshold setting

In this subsection, the optimal thresholds, in the sense of maximizing the probability of correct detection, are obtained as follows.

With the assumption that the maximum number of transmit antennas is N_t ,⁶ the probability of correct detection is defined as

$$\mathbb{P}_D = \sum_{n_t=1}^{N_t} \mathbb{P}(n_t) \mathbb{P}\{\hat{n}_t = n_t | \mathbf{H}_{n_t}\} \quad (4.39)$$

where $\mathbb{P}(n_t)$ is the probability of transmission with n_t antennas, and $\mathbb{P}\{\hat{n}_t = n_t | \mathbf{H}_{n_t}\}$ is the conditional probability of correct detection under hypothesis \mathbf{H}_{n_t} , which is given in (4.36) and (4.38) for $1 < n_t < N_t$ and $n_t = 1$, respectively. Since the maximum number of transmit antennas is N_t , $\Gamma_{N_t} = +\infty$ for \mathbf{H}_{N_t} , which results in

$$\mathbb{P}\{\hat{n}_t = N_t | \mathbf{H}_{N_t}\} = \mathbb{P}\{y_{N_t-1} < 0 | \mathbf{H}_{N_t}\} = Q\left(\frac{\omega_{y_{N_t-1}}}{\sigma_{y_{N_t-1}}}\right). \quad (4.40)$$

Given (4.4), the average SNR, $\bar{\gamma}$, at the receive antenna is expressed as

$$\bar{\gamma} = \frac{\mu^{(1)} - \sigma_w^2}{\sigma_w^2} = \frac{n_t \sigma_h^2 \sigma_s^2}{\sigma_w^2} = n_t \gamma. \quad (4.41)$$

After substitution of $\gamma = \bar{\gamma}/n_t$ into (4.33), (4.34), and (4.35), and with use of (4.36),

⁶Note that while N_t is required to obtain the optimal threshold values in the sense of maximizing the probability of correct detection, the proposed algorithm detects the number of transmit antennas without knowledge of N_t by employing in-the-middle thresholds.

(4.38), and (4.40), the conditional probability of correct detection and the probability of correct detection are straightforwardly expressed in terms of $\bar{\gamma}$. The optimal thresholds are obtained by maximizing the probability of correct detection, i.e.,

$$\begin{aligned} & \underset{\Gamma_1, \Gamma_2, \dots, \Gamma_{N_t-1}}{\text{maximize}} && \mathbb{P}_D \\ & \text{subject to} && \bar{\gamma} = \hat{\gamma}, \end{aligned} \tag{4.42}$$

where $\hat{\gamma}$ is the estimated average SNR at the receive antenna. Note that \mathbb{P}_D is a function of the threshold values through the mean, variance, and covariance given in (4.30), (4.33), and (4.34), respectively.

The optimization problem in (4.42) can be numerically solved to compute the optimal thresholds. These thresholds are obtained at the expense of a very large computational complexity to numerically solve the multi-dimensional optimization problem in (4.42) given the estimated average received SNR, $\hat{\gamma}$, and *a priori* knowledge about the maximum number of transmit antennas N_t .

For $N_c \gg 1$, $N_b \gg 1$, and over a wide range of SNRs, numerical results shows that the optimization problem yields the thresholds $\Gamma_{n_t} \approx n_t + 1/2$, $n_t = 1, \dots, N_t - 1$ (see results in Fig. 4.5). These are used with the proposed algorithm for simplicity, as they avoid solving the optimization problem and are referred to as suboptimal. Furthermore, the assumption on N_t is dropped with the proposed algorithm, by employing $\Gamma_0 = -\infty$ and $\Gamma_{n_t} = n_t + 1/2$ for $n_t = 1, 2, \dots$ and sequentially checking if the decision statistic belongs to the interval $(\Gamma_{n_t-1}, \Gamma_{n_t}]$; if it does, the decision that $\hat{n}_t = n_t$ is made (see Algorithm 10).

4.4 Simulations

In this section, we examine the performance of the proposed antenna enumeration algorithms through several simulation results.

4.4.1 Simulation Setup

Unless otherwise mentioned, the modulation was quadrature phase-shift-keying (QPSK) with spatial multiplexing transmission scheme, and $N_c = 100$ and $N_b = 1000$.⁷ The channel coefficients were modeled as independent circular complex Gaussian random variables with variance σ_h^2 . The additive white noise was modeled as circular complex Gaussian random variables with variance σ_w^2 , and the average SNR per transmit antennas in dB was defined as $\gamma = 10 \log(\sigma_h^2 \sigma_s^2 / \sigma_w^2)$ dB. Without loss of generality, it was assumed that $\sigma_h^2 \sigma_s^2 = 1$. $\mathbb{P}\{\hat{n}_t = n_t | H_{n_t}\}$ and $\mathbb{P}_D = \frac{1}{4} \sum_{n_t=1}^4 P\{\hat{n}_t = n_t | H_{n_t}\}$ were used as performance measures and were obtained based on 1000 Monte Carlo trials for each hypothesis. Different threshold settings were considered:

- 1) Optimal thresholds $\Gamma_1, \Gamma_2, \Gamma_3, \Gamma_0 = -\infty, \Gamma_4 = +\infty$ (Scenario 1).
- 2) $\Gamma_{n_t} = n_t + 1/2, n_t = 1, 2, 3, \Gamma_0 = -\infty, \Gamma_4 = +\infty$ (Scenario 2).
- 3) $\Gamma_{n_t} = n_t + 1/2, n_t = 1, 2, 3, 4, \dots, \Gamma_0 = -\infty$ (Scenario 3).⁸

The next section presents results for the proposed antenna enumeration algorithms under scenario 3, unless otherwise mentioned. Furthermore, the algorithm performance was evaluated under various transceiver and channel impairments, with

⁷Note that $N_c = \lfloor 9/16\pi f_d T_s \rfloor$ [163] and $N_b = \lfloor N/N_c \rfloor$, where N represents the number of observed symbols.

⁸For Scenarios 1 and 2, \mathbb{P}_D represents the probability of correct detection in Bayesian framework. On the other hand, Scenario 3 corresponds to the proposed algorithm that requires no knowledge of $N_t = 4$, and in this case \mathbb{P}_D does not represent the probability of correct detection in Bayesian framework. We use it as an evaluation metric in simulation results for the comparison reason and because showing separate curves for $\mathbb{P}\{\hat{n}_t = n_t | H_{n_t}\}, n_t = 1, 2, 3, 4$, generates crowded figures.

the received signal expressed as

$$r_{k,b} = \sum_{m=1}^{N_b} \sum_{n=1}^{N_c} [\tilde{\mathbf{h}}_m \mathbf{s}_{n,m} e^{j(2\pi((b-1)N_c+k)\Delta f)} p((k-n+(b-m)N_c+\tau)T_s)] + w_{k,b} \quad (4.43)$$

where Δf is the CFO normalized by data rate, $p(t)$ is the equivalent impulse response of the transmit and receive filters, τ is the normalized TO with $\tau \in [0, 1/2]$, and $\tilde{\mathbf{h}}_b$ is the Kronecker spatially correlated fading given by [166]

$$\tilde{\mathbf{h}}_b = \mathbf{R}_T^{1/2} \mathbf{h}_b. \quad (4.44)$$

Here \mathbf{h}_b is a vector with independent and identically distributed (i.i.d.) elements, drawn from the circular complex Gaussian distribution with unit variance, and \mathbf{R}_T is the transmitter correlation matrix,

$$\mathbf{R}_{T(ij)} = \begin{cases} \rho_h^{j-i}, & i \leq j \\ \mathbf{R}_{T(ij)}^*, & i > j \end{cases}, \quad |\rho_h| < 1, \quad (4.45)$$

with ρ_h as the correlation coefficient for the adjacent antennas. The effect of the TO was studied when a root-raised-cosine filter with roll-off factor $\beta = 0.35$ was used at the transmitter, while either a root-raised-cosine filter with $\beta = 0.35$ or a low-pass 10th-order Butterworth filter with cutoff frequency $w_c = 0.2\pi$ was employed at the receiver to remove the out-of-band noise. Unless otherwise mentioned, the former case was considered, along with no TO. Also unless otherwise mentioned, $\Delta f = 0$ and $\rho_h^{j-i} = 1$ if $i = j$ and 0 otherwise.

4.4.2 Simulation Results

Fig. 4.3 shows the conditional probability of correct detection, $\mathbb{P}\{\hat{n}_t = n_t | H_{n_t}\}$, of the proposed blind antenna enumeration algorithm versus SNR for different numbers of transmit antennas. As expected, the conditional probability of correct detection increases as SNR increases. This increase can be explained, as an increase in the SNR leads to a decrease in $\sigma_{y_{n_t}}^2$ and $\sigma_{y_{n_t-1}}^2$ in (4.33), which in turn yields an improvement in the conditional probability of correct detection. Also, as can be seen, the conditional probability of correct detection decreases as the number of transmit antennas increases. This decrease can be also explained, as $\omega_{y_{n_t-1}}/\sigma_{y_{n_t-1}}$ and $\omega_{y_{n_t}}/\sigma_{y_{n_t}}$ in (4.36) and (4.38) decrease as the number of transmit antennas increases. It is worth noting that there is a very close agreement between the theoretical results given by (4.36) and (4.38) and simulation results.

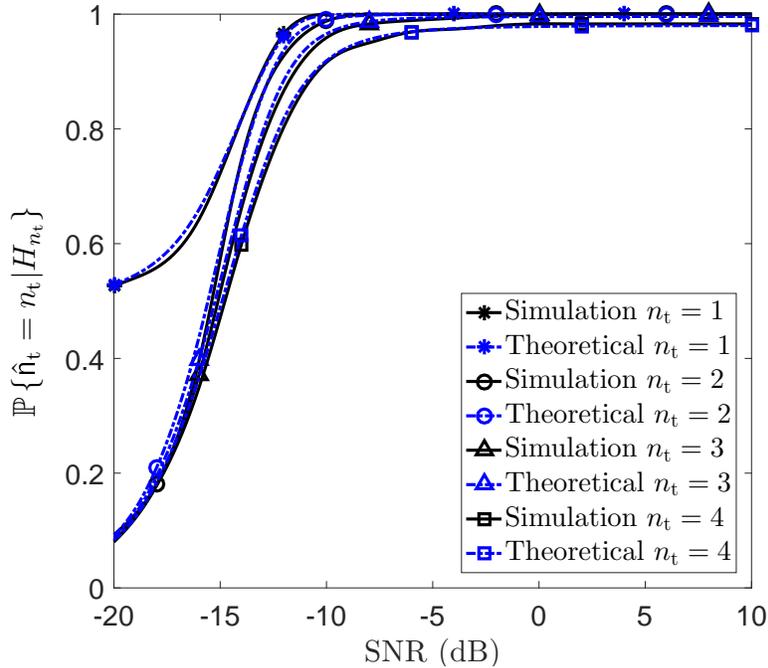


Fig. 4.3: The conditional probability of correct detection, $\mathbb{P}\{\hat{n}_t = n_t | H_{n_t}\}$ versus SNR (γ) for different numbers of transmit antennas, n_t . Results are presented for the proposed blind antenna enumeration algorithm.

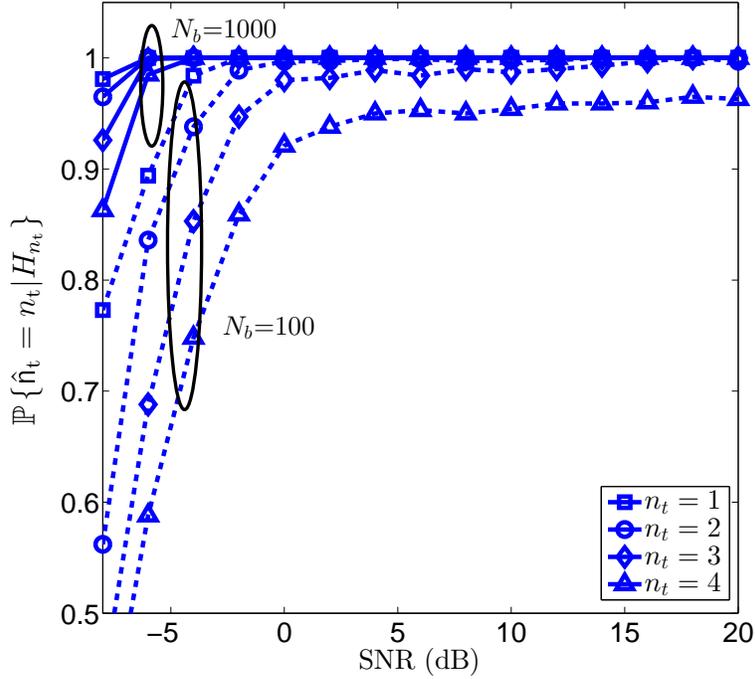


Fig. 4.4: The conditional probability of correct detection, $\mathbb{P}\{\hat{n}_t = n_t | H_{n_t}\}$, of the proposed semi-blind antenna enumeration algorithm versus SNR for different n_t and N_b values.

Fig. 4.4 shows the conditional probability of correct detection, $\mathbb{P}\{\hat{n}_t = n_t | H_{n_t}\}$, of the proposed semi-blind antenna enumeration algorithm versus SNR for different number of transmit antennas, n_t , $n_t = 1, \dots, 4$, and different N_b values. As seen, the proposed algorithm exhibits a good performance over a wide range of SNRs for $N_b = 100$ and 1000 , and the conditional probability of correct detection goes to one even at negative SNRs for $N_b = 1000$. The performance improves as either N_b or SNR increases. Additionally, the conditional probability of correct detection decreases as the number of transmit antenna increases; this is because the variance of the decision statistic Θ in (4.20) increases with n_t , as confirmed by simulation experiments.

In Fig. 4.5, \mathbb{P}_D for the proposed blind antenna enumeration algorithm is shown versus SNR for Scenarios 1, 2, and 3. As expected, \mathbb{P}_D is higher under Scenario 1 when compared with Scenarios 2 and 3. Results for Scenario 2 almost coincide with

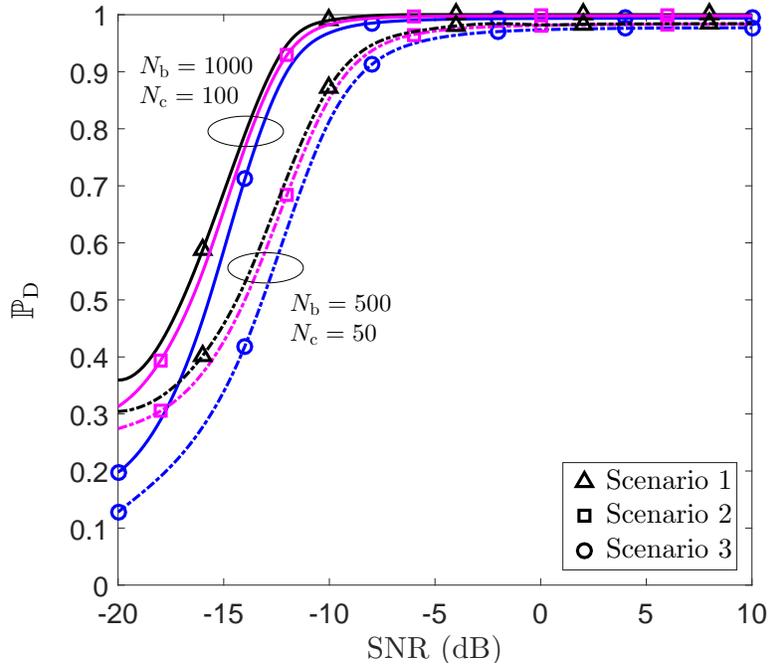


Fig. 4.5: \mathbb{P}_D versus SNR (γ) for different threshold settings. Scenario 1: optimal thresholds $\Gamma_1, \Gamma_2, \Gamma_3$, $\Gamma_0 = -\infty$ and $\Gamma_4 = +\infty$; Scenario 2: $\Gamma_{n_t} = n_t + 1/2$, $n_t = 1, 2, 3$, $\Gamma_0 = -\infty$, $\Gamma_4 = +\infty$; and Scenario 3: $\Gamma_{n_t} = n_t + 1/2$, $n_t = 1, 2, 3, 4, \dots$, $\Gamma_0 = -\infty$. Results are presented for the proposed blind antenna enumeration algorithm.

those for Scenario 1 for a wide range of SNRs; additionally, results for Scenario 3 follow closely as the SNR increases.

Fig. 4.6 illustrates the effect of an incorrect N_t value for Scenario 1, i.e., when the blind antenna enumeration algorithm employs the optimal thresholds. The conditional probability of correct detection, $\mathbb{P}\{\hat{n}_t = n_t | H_{n_t}\}$, $n_t = 1, 2, 3$ is shown versus SNR when it is incorrectly assumed that $N_t = 5$ and the true value is $N_t = 3$. As can be seen, the performance is not significantly affected under H_1 and H_2 . On the other hand, under H_3 , knowledge of the true value of the maximum number of transmit antennas ($N_t = 3$) leads to a significant performance improvement when compared with the overestimated value ($N_t = 5$). This improvement can be easily explained, as $\Gamma_3 = +\infty$ with the true value, whereas $3 < \Gamma_3 < 4$ with the overestimated value.

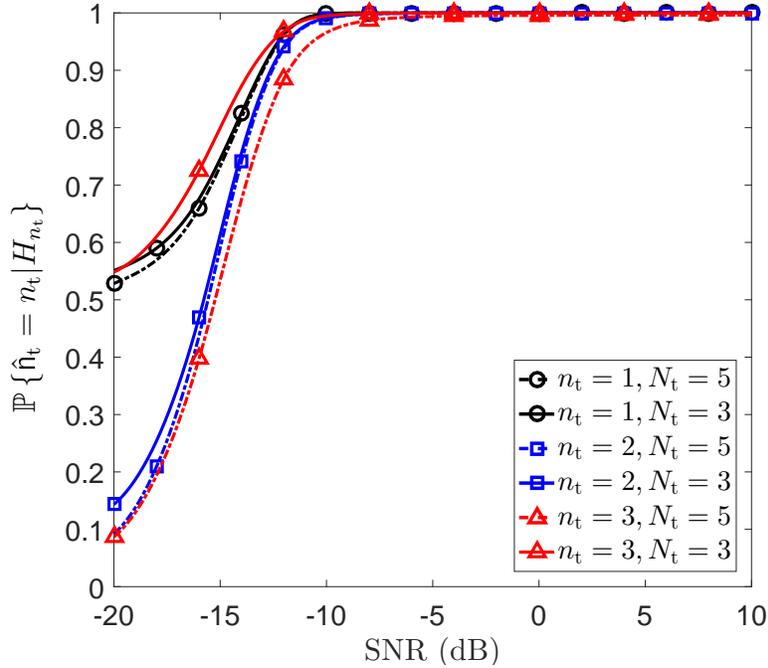


Fig. 4.6: The effect of an incorrect N_t value ($N_t=5$) on the conditional probability of correct detection of the proposed blind antenna enumeration algorithm with optimal threshold setting (Scenario 1), $\mathbb{P}\{\hat{n}_t = n_t|H_{n_t}\}$, $n_t = 1, 2, 3$, when the true value of N_t is 3.

In Fig. 4.7, \mathbb{P}_D is presented over the CFO normalized to the data rate, Δf . As seen, the proposed blind antenna enumeration algorithm is basically insensitive to the CFO; this occurs because of the absolute value operator of the second-order moment in (4.2), which eliminates the effect of CFO.

Fig. 4.8 illustrates the effect of the modulation type on \mathbb{P}_D . As observed, the modulation type does not affect the performance of the proposed blind antenna enumeration algorithm. This phenomenon can be explained, as different values of Ω_s in (4.35), representing different modulation types, cause minor changes in $\sigma_{y_{n_t}}^2$ and $\sigma_{y_{n_t-1}}^2$ in (4.36) and (4.38). It is worth noting that for all PSK modulations $\Omega_s = 1$, and for 16-QAM and 64-QAM, Ω_s is equal to 1.32 and 1.38, respectively [127].

Fig. 4.9 shows the effect of the spatially correlated fading on \mathbb{P}_D versus SNR for a

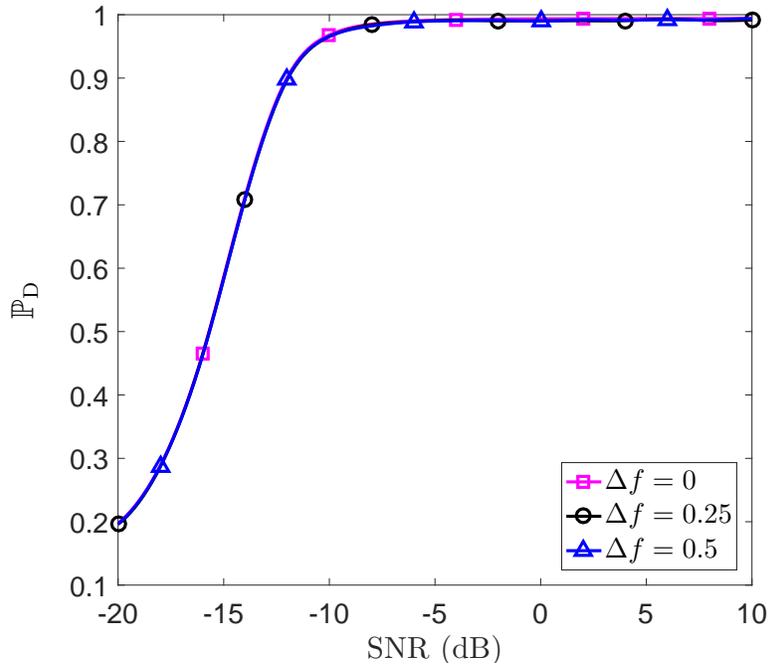


Fig. 4.7: The effect of the CFO on the performance of the proposed blind antenna enumeration algorithm.

correlation coefficient $\rho_h = 0, 0.1, \text{ and } 0.2$. As can be observed, the performance of the proposed blind antenna enumeration algorithm is robust to the spatial correlation for $\rho_h < 0.1$. This phenomenon can be explained, as while both ω_{y_ℓ} and σ_{y_ℓ} , $\ell = n_t, n_t - 1$ decrease, $\omega_{y_\ell}/\sigma_{y_\ell}$ remains almost constant for $\rho_h < 0.1$.

Fig. 4.10 demonstrates the performance of the proposed blind antenna enumeration algorithm in the presence of the TO for the root-raised-cosine and Butterworth filter at the receive side, respectively. As expected, the proposed algorithm exhibits a better performance for the former, as it is the matched filter. A normalized TO, $\tau = 0, 0.2, 0.4$, is considered; as can be seen, the proposed algorithm is robust to TO.

Fig. 4.11 shows the effect of the Doppler frequency on the performance of the

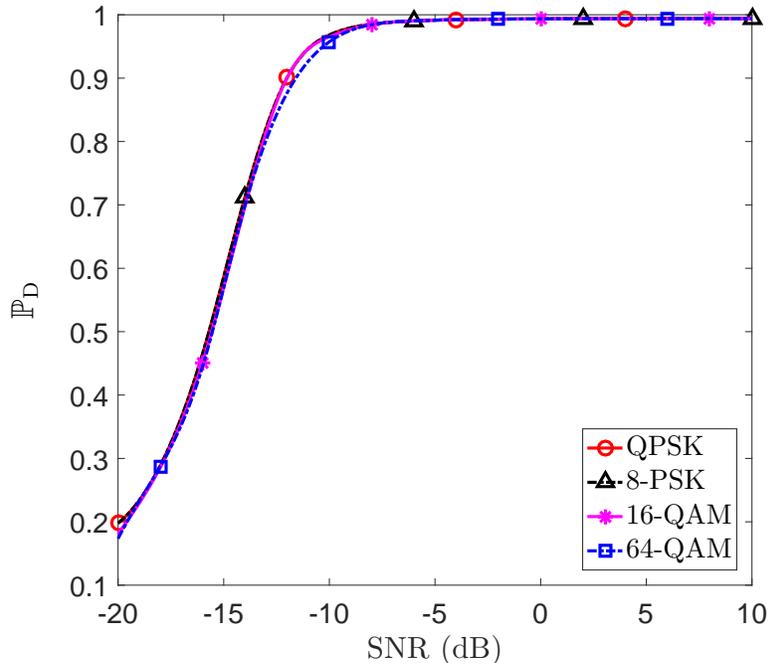


Fig. 4.8: The effect of the modulation type on the performance of the proposed blind antenna enumeration algorithm.

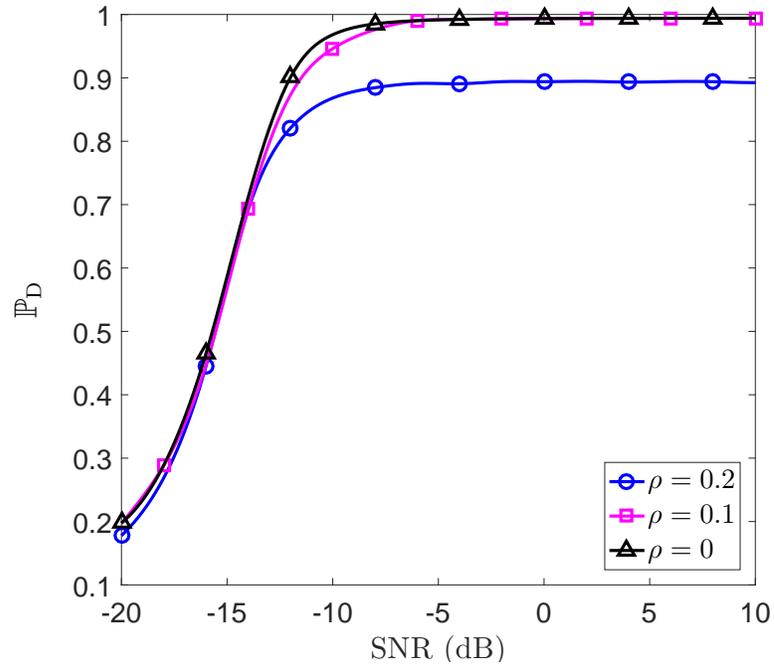


Fig. 4.9: The effect of the spatially correlated fading on the performance of the proposed blind antenna enumeration algorithm.

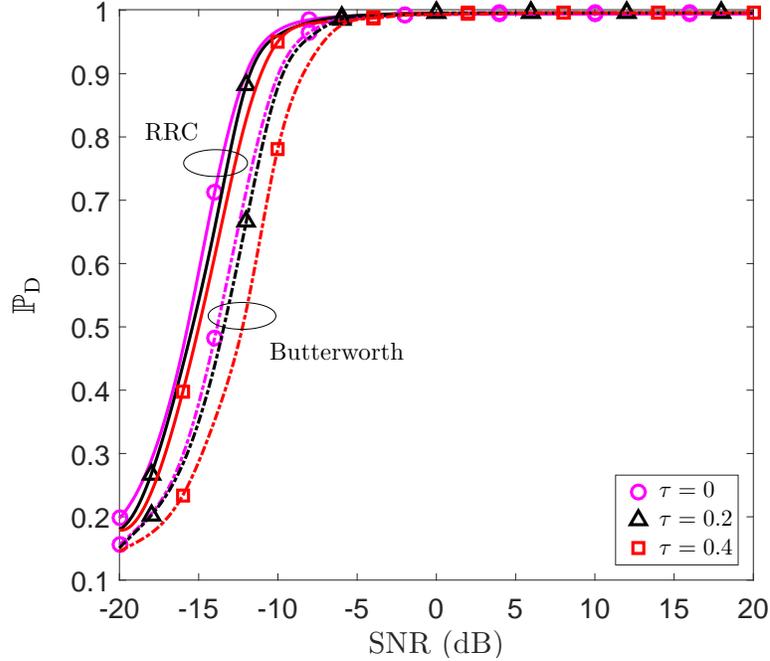


Fig. 4.10: The effect of TO for root-raised-cosine (RRC) filter with $\beta = 0.35$ and low-pass 10th-order Butterworth filter with cutoff frequency $w_c = 0.2\pi$ at receive side. Results are presented for the proposed blind antenna enumeration algorithm.

proposed blind antenna enumeration algorithm for the block length $N_c = \lfloor 9/16\pi f_d T_s \rfloor$ and $NT_s = 1$ sec at SNR = 0. As can be seen, the conditional probability of correct detection increases with f_d . The reason is that for lower f_d values, while N_c is large, N_b is small, which results in a large estimation error of the expectation over the channel distribution. As f_d increases, both N_c and N_b are reasonably large, and the estimation error of the expectations is small. It is worth noting that for very high values of f_d ,⁹ the conditional probability of correct detection decreases, as N_c significantly decreases, and thus, there is a large estimation error of the expectation over the signal and noise distributions.

In Fig. 4.12, \mathbb{P}_D for the proposed blind antenna enumeration algorithm is presented versus SNR when the estimate of σ_w^2 [5] is employed in (4.11). As expected,

⁹This decrease is not shown in Fig. 10, as it appears out of the practical f_d range.

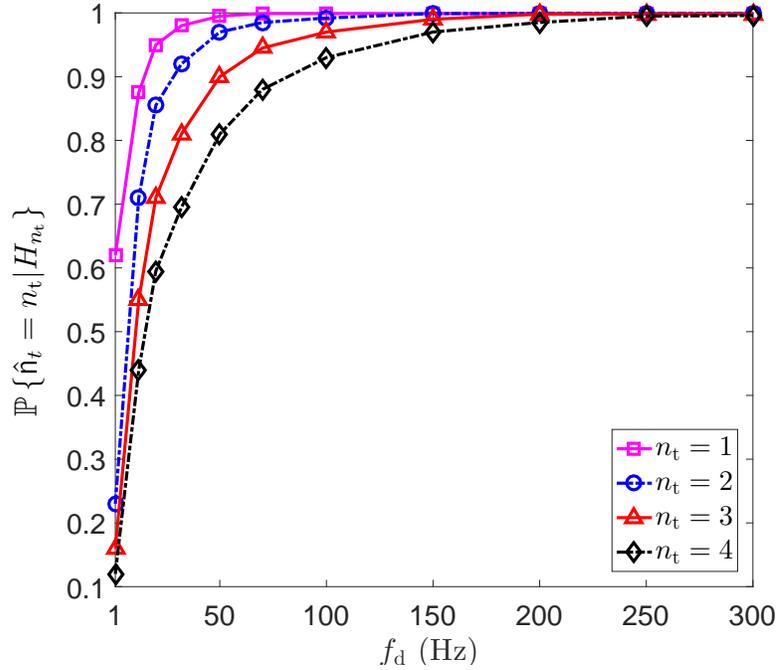


Fig. 4.11: The effect of the Doppler frequency, f_d , on the conditional probability of correct detection, $\mathbb{P}\{\hat{n}_t = n_t | H_{n_t}\}$, for $N_c = \lfloor 9/16\pi f_d T_s \rfloor$, $NT_s = 1$ sec at SNR=0 dB. Results are presented for the proposed blind antenna enumeration algorithm.

\mathbb{P}_D reduces when the estimate of σ_w^2 is used. Also, as shown, the performance of the proposed algorithm improves with a more accurate estimate of σ_w^2 .

Fig. 4.13 shows the conditional probability of correct detection of the number of transmit antennas under hypothesis H_2 , $P\{\hat{n}_t = 2 | H_2\}$, versus SNR for the proposed blind antenna enumeration algorithm with ($n_r = 1$), and the soft MDL and AIC approaches with ($n_r = 5$). As can be seen, even under the unfair comparison condition, the proposed algorithm outperforms the soft AIC and MDL approaches at low SNR values. In contrast, the soft AIC and MDL exhibit better performance at high SNR values for a reduced number of observation blocks. This can be explained, as the proposed algorithm relies on the variability of the channel and requires a larger number of blocks. It should be mentioned that increasing the number of blocks leads to performance improvement in the soft MDL and AIC only for N_b below 100 and SNRs

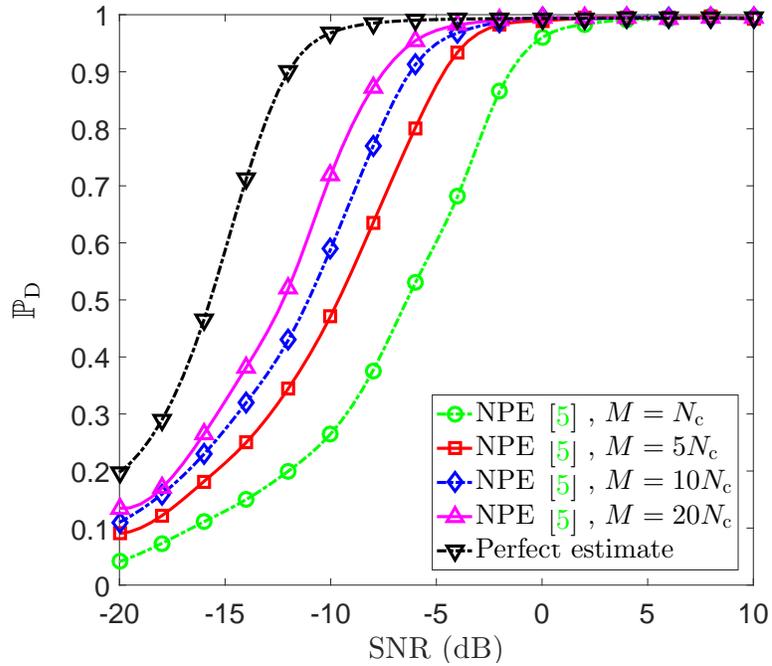


Fig. 4.12: The performance of the proposed blind antenna enumeration algorithm with noise power estimate (NPE) [5].

above -4 dB and -8 dB, respectively.

4.5 Conclusions and Directions for Future Research

Chapter 4 of this thesis dealt with antenna enumeration in time-varying fading channels. In this section, a summary of the main results in Chapter 4 are provided, and possible directions of future research are point out.

4.5.1 Summary

Motivated by the need for and importance of antenna enumeration in adaptive multi-antenna wireless communication systems, two new antenna enumeration algorithms for time-varying fading channel were developed in Chapter 4 of this thesis. The pro-

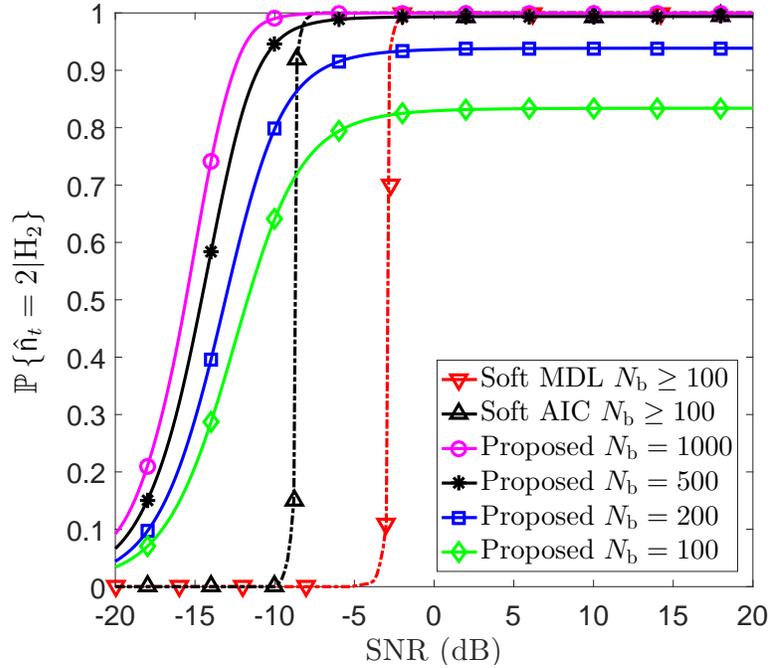


Fig. 4.13: Performance comparison of the proposed blind antenna enumeration algorithm with $n_r = 1$ and the soft MDL and AIC approaches in [6] with $n_r = 5$, for $n_t = 2$.

posed blind and semi-blind antenna enumeration algorithms exploit the second- and fourth-order statistics of the received signal and the time-varying nature of the fading channel, while the receiver is equipped with a single receive antenna. For both algorithms, the antenna enumeration problem was formulated as a multiple hypothesis testing problem, and a one-dimensional decision statistic was proposed. The proposed blind algorithm does not depend on the signal type. However, the semi-blind algorithm requires the modulation format for antenna enumeration. Furthermore, unlike the existing algorithms, the proposed algorithms can detect a large number of transmit antennas.

In order to evaluate the performance of the proposed blind algorithm analytically, an equivalent two-dimensional decision statistic was employed, and an a closed-form expression for the probability of correct detection was derived.

The effectiveness of the proposed algorithms was confirmed through simulations,

and it was shown that they are robust to the CFO and TO. Moreover, simulation results show that the proposed blind antenna enumeration algorithm outperforms the MDL and AIC algorithms for low SNRs while it employs only a single receive antenna.

4.5.2 Future research

The results in Chapter 4 of this thesis open interesting directions for a number of future research topics. Here, we outline two of them, as follows

- The proposed antenna enumeration algorithms can be extended to multiple receive antennas;
- The optimal combining method for the multiple receive antennas can be derived through the bootstrap technique.

Appendix

4.A Proof of the Second-order Statistics

With the channel coefficients corresponding to different transmit antennas being independent circular complex Gaussian random variables with variance σ_h^2 , i.e., $\mathbb{E}\{\mathbf{h}_b \mathbf{h}_b^H\} = \sigma_h^2 \mathbf{I}$, and by employing the property of a circular complex Gaussian random variable $x \sim \mathcal{N}_c(0, \sigma_x^2)$ from [167] that proposes

$$\mathbb{E}\{|x|^{2n}\} = n! \sigma_x^{2n} \quad (4.46)$$

where n is a positive integer value and $n!$ is the factorial of n , (4.3) for $d = 2$ is expressed as

$$\begin{aligned} \mu^{(2)} &= \mathbb{E} \left\{ \left(\sigma_s^2 \sum_{m_1}^{n_t} |h_b^{(m_1)}|^2 + \sigma_w^2 \right)^2 \right\} = \sigma_s^4 \sum_{m_1}^{n_t} \mathbb{E} \left\{ |h_b^{(m_1)}|^4 \right\} + 2\sigma_s^2 \sigma_w^2 \sum_{m_1}^{n_t} \mathbb{E} \left\{ |h_b^{(m_1)}|^2 \right\} \\ &\quad + \sigma_s^4 \sum_{m_1: m_2}^{n_t} \mathbb{E} \left\{ |h_b^{(m_1)}|^2 \right\} \mathbb{E} \left\{ |h_b^{(m_2)}|^2 \right\} + \sigma_w^4 \\ &= (n_t^2 + n_t) \sigma_s^4 \sigma_h^4 + 2n_t \sigma_s^2 \sigma_h^2 \sigma_w^2 + \sigma_w^4. \end{aligned} \quad (4.47)$$

4.B Proof of the Unbiased Estimators in (4.8)-(4.9)

As is known, $\hat{\theta}$ is an unbiased estimate of a parameter θ if $\mathbb{E}\{\hat{\theta}\} = \theta$ [168]. Accordingly, we show that $\mathbb{E}\{\hat{\zeta}\} = (\mu^{(1)})^2$ and $\mathbb{E}\{\hat{\mu}^{(2)}\} = \mu^{(2)}$.

By applying the statistical expectation operator on both sides of (4.8), utilizing the linearity property of the expectation, and employing a property of the conditional expectation,¹⁰ one obtains

$$\mathbb{E}\{\hat{\zeta}\} = \frac{1}{N_1} \sum_{b_1:b_2}^{N_b} \sum_{k_1:k_2}^{N_c} \mathbb{E} \left\{ \mathbb{E} \left\{ |r_{k_1,b_1}|^2 |r_{k_2,b_2}|^2 | \mathbf{h}_{b_1}, \mathbf{h}_{b_2} \right\} \right\} \quad (4.48)$$

where $N_1 \triangleq N_b(N_b - 1)N_c(N_c - 1)$. Since the signal, noise and channel are independent, and as $b_2 \neq b_1$ and $k_2 \neq k_1$, r_{k_1,b_1} and r_{k_2,b_2} are independent; hence, (4.48) can be written as

$$\mathbb{E}\{\hat{\zeta}\} = \frac{1}{N_1} \sum_{b_1:b_2}^{N_b} \sum_{k_1:k_2}^{N_c} \mathbb{E} \left\{ \mathbb{E} \left\{ |r_{k_1,b_1}|^2 | \mathbf{h}_{b_1} \right\} \right\} \mathbb{E} \left\{ \mathbb{E} \left\{ |r_{k_2,b_2}|^2 | \mathbf{h}_{b_2} \right\} \right\} \quad (4.49)$$

By substituting $\mathbb{E}\{\mathbb{E}\{|r_{k_1,b_1}|^2 | \mathbf{h}_{b_1}\}\}$ and $\mathbb{E}\{\mathbb{E}\{|r_{k_2,b_2}|^2 | \mathbf{h}_{b_2}\}\}$ with $\mu^{(1)}$, one easily obtains that $\mathbb{E}\{\hat{\zeta}\} = (\mu^{(1)})^2$.

Similarly, for $\hat{\mu}^{(2)}$ in (4.9), we can write

$$\mathbb{E}\{\hat{\mu}^{(2)}\} = \frac{1}{N_2} \sum_b^{N_b} \sum_{k_1:k_2}^{N_c} \mathbb{E} \left\{ \mathbb{E} \left\{ |r_{k_1,b}|^2 |r_{k_2,b}|^2 | \mathbf{h}_b \right\} \right\} \quad (4.50)$$

where $N_2 \triangleq N_b N_c (N_c - 1)$. Because the signal, noise and channel are independent, and as $k_2 \neq k_1$, $r_{k_1,b}$ and $r_{k_2,b}$ are independent in terms of the signal and noise distribution, and one can easily write

¹⁰For two random variables, w and z , $\mathbb{E}\{z\} = \mathbb{E}\{\mathbb{E}\{z|w\}\}$ [169].

$$\mathbb{E}\{\hat{\mu}^{(2)}\} = \frac{1}{N_2} \sum_b^{N_b} \sum_{k_1:k_2}^{N_c} \mathbb{E} \left\{ \mathbb{E} \left\{ |r_{k_1,b}|^2 \middle| \mathbf{h}_b \right\} \mathbb{E} \left\{ |r_{k_2,b}|^2 \middle| \mathbf{h}_b \right\} \right\} \quad (4.51)$$

By substituting $\mathbb{E}\{|r_{k_1,b}|^2|\mathbf{h}_b\}$ and $\mathbb{E}\{|r_{k_2,b}|^2|\mathbf{h}_b\}$ with μ_b , we can write

$$\mathbb{E}\{\hat{\mu}^{(2)}\} = \frac{1}{N_2} \sum_b^{N_b} \sum_{k_1:k_2}^{N_c} \mathbb{E}\{(\mu_b)^2\}. \quad (4.52)$$

Finally, by replacing $\mathbb{E}\{(\mu_b)^2\}$ with $\mu^{(2)}$ in (4.52), one easily obtains that $\mathbb{E}\{\hat{\mu}^{(2)}\} = \mu^{(2)}$.

4.C First-order Taylor Expansion of the Decision Statistics in (4.11)

Let us define

$$\mathbf{z} = \frac{\mathbf{z}_1}{\mathbf{z}_2}, \quad (4.53)$$

where \mathbf{z}_1 and \mathbf{z}_2 are two correlated random variables. One can write

$$\frac{\mathbf{z}_1}{\mathbf{z}_2} = \frac{\mu_{z_1} + u_1}{\mu_{z_2} + u_2}, \quad (4.54)$$

with $\mu_{z_1} \triangleq \mathbb{E}\{\mathbf{z}_1\}$ and $\mu_{z_2} \triangleq \mathbb{E}\{\mathbf{z}_2\}$, and u_1 and u_2 as two zero-mean correlated random variables. With the assumption that

$$\frac{\text{Var}\{u_2\}}{\mu_{z_2}^2} \ll 1 \quad (4.55)$$

with probability almost one, $|u_2/\mu_{z_2}| < 1$, and thus, by employing the first-order Taylor expansion, one can write

$$z = \left(\frac{\mu_{z_1}}{\mu_{z_2}} + \frac{u_1}{\mu_{z_2}} \right) \left(1 + \frac{u_2}{\mu_{z_2}} \right)^{-1} = \left(\frac{\mu_{z_1}}{\mu_{z_2}} + \frac{u_1}{\mu_{z_2}} \right) \left(1 - \frac{u_2}{\mu_{z_2}} \right) + \mathcal{O}(u_2^2). \quad (4.56)$$

By applying the statistical expectation, one can further write

$$\mathbb{E}\{z\} = \mathbb{E}\left\{ \frac{z_1}{z_2} \right\} \approx \frac{\mu_{z_1}}{\mu_{z_2}} - \frac{\mathbb{E}\{u_1 u_2\}}{\mu_{z_2}^2}. \quad (4.57)$$

With the assumption that

$$\frac{\mu_{z_1}}{\mu_{z_2}} \gg \frac{\mathbb{E}\{u_1 u_2\}}{\mu_{z_2}^2} \quad (4.58)$$

one obtains

$$\mathbb{E}\left\{ \frac{z_1}{z_2} \right\} \approx \frac{\mu_{z_1}}{\mu_{z_2}}. \quad (4.59)$$

Consequently, the expectation of the ratio of two correlated random variables can be approximated by the ratio of their expectations when the conditions in (4.55) and (4.58) are fulfilled. For the proposed algorithm, we have

$$\begin{aligned} z_1 &= \hat{\zeta} - 2\hat{\mu}^{(1)}\sigma_w^2 + \sigma_w^4 \\ z_2 &= \hat{\mu}^{(2)} - \hat{\zeta} \end{aligned} \quad (4.60)$$

with

$$\begin{aligned} \mu_{z_1} &= (\mu^{(1)} - \sigma_w^2)^2 = n_t^2 \sigma_h^4 \sigma_s^4 \\ \mu_{z_2} &= \mu^{(2)} - (\mu^{(1)})^2 = n_t \sigma_h^4 \sigma_s^4 \end{aligned} \quad (4.61)$$

and

$$\begin{aligned} u_1 &= \hat{\zeta} - (\mu^{(1)})^2 + 2(\mu^{(1)} - \hat{\mu}^{(1)})\sigma_w^2 \\ u_2 &= \hat{\mu}^{(2)} - \mu^{(2)} - \hat{\zeta} + (\mu^{(1)})^2. \end{aligned} \quad (4.62)$$

To check the conditions in (4.55) and (4.58), $\text{Var}\{u_2\}$ and $\mathbb{E}\{u_1 u_2\}$ are required to be computed. After several calculation, one obtains

$$\begin{aligned} \text{Var}\{u_2\} / \mu_{z_2}^2 & \quad (4.63) \\ &= n_t^{-2} \left[\beta_{20}^1 [2\alpha_{22}^2 \bar{\vartheta}^{(2)} + 4\alpha_{22}^3 \bar{\kappa}^{(3)} + \alpha_{22}^4 \bar{\mu}^{(4)}] + [\beta_{20}^2 (2\alpha_{22}^2 + 4\alpha_{22}^3 + \alpha_{22}^4) - 1] (\bar{\mu}^{(2)})^2 \right. \\ &+ \beta_{22}^2 [2\alpha_{22}^2 ((\bar{\vartheta}^{(1)})^2 + (\bar{\mu}^{(2)})^2) + 4\alpha_{22}^3 \bar{\vartheta}^{(1)} \bar{\mu}^{(2)} + (4\alpha_{22}^3 + 2\alpha_{22}^4) (\bar{\mu}^{(2)})^2] \\ &+ 4\beta_{22}^3 [\alpha_{22}^2 (\bar{\vartheta}^{(1)} + \bar{\mu}^{(2)}) + \alpha_{22}^3 \bar{\vartheta}^{(1)} + (3\alpha_{22}^3 + \alpha_{22}^4) \bar{\mu}^{(2)}] (\bar{\mu}^{(1)})^2 + [\beta_{22}^4 (2\alpha_{22}^2 + 4\alpha_{22}^3 + \alpha_{22}^4) - 1] (\bar{\mu}^{(1)})^4 \\ &\left. - 2[\beta_{21}^2 [4(\alpha_{22}^2 + \alpha_{22}^3) \bar{\kappa}^{(2)} + (4\alpha_{22}^3 + 2\alpha_{22}^4) \bar{\mu}^{(3)}] \bar{\mu}^{(1)} + [\beta_{21}^3 (2\alpha_{22}^2 + 4\alpha_{22}^3 + \alpha_{22}^4) - 1] \bar{\mu}^{(2)} (\bar{\mu}^{(1)})^2] \right] \end{aligned}$$

$$\begin{aligned} \mathbb{E}\{u_1 u_2\} / \mu_{z_2}^2 & \quad (4.64) \\ &= n_t^{-2} \left[-\beta_{22}^2 [2\alpha_{22}^2 ((\bar{\vartheta}^{(1)})^2 + (\bar{\mu}^{(2)})^2) + 4\alpha_{22}^3 \bar{\vartheta}^{(1)} \bar{\mu}^{(2)} + (4\alpha_{22}^3 + 2\alpha_{22}^4) (\bar{\mu}^{(2)})^2] \right. \\ &- 4\beta_{22}^3 [\alpha_{22}^2 (\bar{\vartheta}^{(1)} + \bar{\mu}^{(2)}) + \alpha_{22}^3 \bar{\vartheta}^{(1)} + (3\alpha_{22}^3 + \alpha_{22}^4) \bar{\mu}^{(2)}] (\bar{\mu}^{(1)})^2 - [\beta_{22}^4 (2\alpha_{22}^2 + 4\alpha_{22}^3 + \alpha_{22}^4) - 1] (\bar{\mu}^{(1)})^4 \\ &+ \beta_{21}^2 [4(\alpha_{22}^2 + \alpha_{22}^3) \bar{\kappa}^{(2)} + (4\alpha_{22}^3 + 2\alpha_{22}^4) \bar{\mu}^{(3)}] \bar{\mu}^{(1)} + [\beta_{21}^3 (2\alpha_{22}^2 + 4\alpha_{22}^3 + \alpha_{22}^4) - 1] \bar{\mu}^{(2)} (\bar{\mu}^{(1)})^2 \\ &- 2\gamma^{-1} [\beta_{20}^1 (2\alpha_{21}^2 \bar{\kappa}^{(2)} + \alpha_{21}^3 \bar{\mu}^{(2)}) + (\beta_{20}^2 (2\alpha_{21}^2 + \alpha_{21}^3) - 1) \bar{\mu}^{(1)} \bar{\mu}^{(2)}] \\ &\left. + 2\gamma^{-1} [2\beta_{21}^2 [\alpha_{21}^2 (\bar{\vartheta}^{(1)} + \bar{\mu}^{(2)}) + \alpha_{21}^3 \bar{\mu}^{(2)}] \bar{\mu}^{(1)} + [\beta_{21}^3 (2\alpha_{21}^2 + \alpha_{21}^3) - 1] (\bar{\mu}^{(1)})^3] \right]. \end{aligned}$$

The theoretical results for $\text{Var}\{u_2\} / \mu_{z_2}^2$ and $\mathbb{E}\{u_1 u_2\} / \mu_{z_2}^2$ for different values of N_b , N_c , and n_t at $\gamma = -10$ dB, as a worst SNR case, are shown in Fig. 4.C.1 and Fig. 4.C.2, respectively. As can be seen, for $N_c \geq 50$ and $N_b \geq 200$, $\text{Var}(u_2) / \mu_{z_2}^2 \ll 1$ and $n_t = \mu_{z_1} / \mu_{z_2} \gg \mathbb{E}\{u_1 u_2\} / \mu_{z_2}^2$; thus, the approximation (4.12) is accurate. This is also confirmed with the simulation results for $\mathbb{E}\{\Upsilon\}$ in Fig. 4.C.3. It is clear that

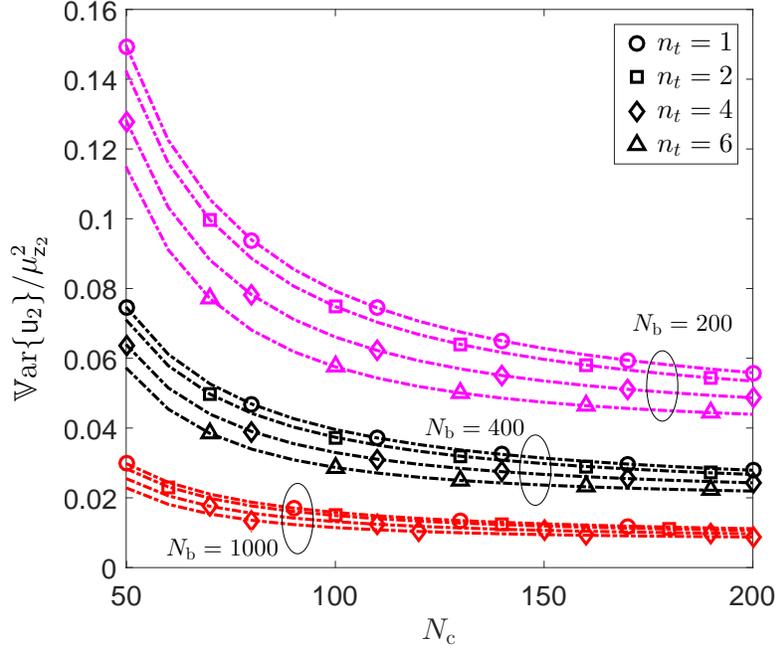


Fig. 4.C.1: $\text{Var}\{u_2\}/\mu_{z_2}^2$ for different values of N_c and N_b at $\text{SNR}=-10$ dB.

the approximation (4.11) is more accurate for high SNR values.

4.D Proof of the Mean Vector

By substituting (4.7), (4.8), and (4.9) in (4.25), one obtains

$$\begin{aligned}
 x_\ell &= \frac{\Gamma_\ell}{N_b N_c (N_c - 1)} \sum_b^{N_b} \sum_{k_1:k_2}^{N_c} |r_{k_1,b}|^2 |r_{k_2,b}|^2 - \frac{\Gamma_\ell + 1}{N_b (N_b - 1) N_c (N_c - 1)} \sum_{b_1:b_2}^{N_b} \sum_{k_1:k_2}^{N_c} |r_{k_1,b_1}|^2 |r_{k_2,b_2}|^2 \\
 &+ \frac{2\sigma_w^2}{N_b N_c} \sum_b^{N_b} \sum_k^{N_c} |r_{k,b}|^2 - \sigma_w^4, \quad \ell = n_t, n_t - 1.
 \end{aligned} \tag{4.65}$$

Let us first define

$$\mu_b^{k_d:k_{n+d-1}} \triangleq \mathbb{E}\{|r_{k_d,b}|^2 |r_{k_{d+1},b}|^2 \cdots |r_{k_{n+d-1},b}|^2 | \mathbf{h}_b\}, \tag{4.66}$$

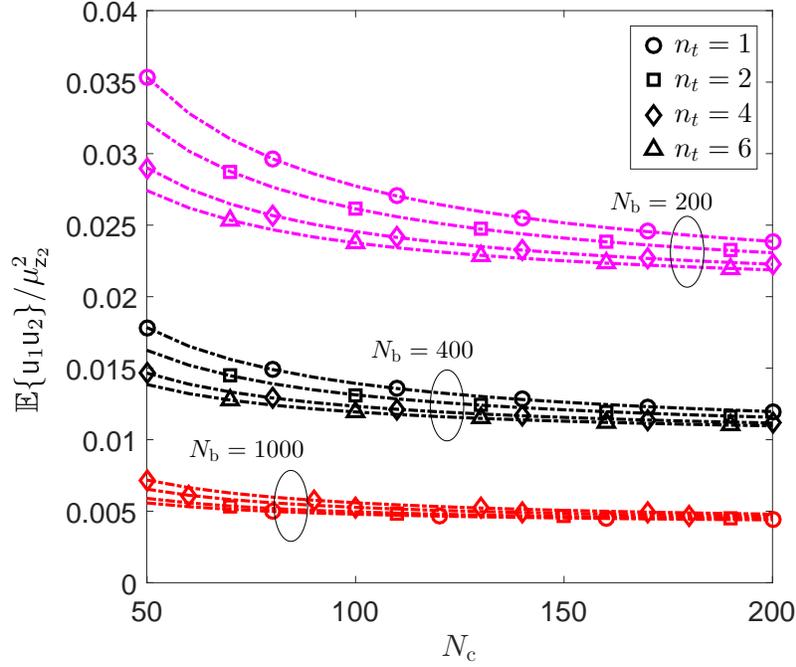


Fig. 4.C.2: $\mathbb{E}\{u_1 u_2\} / \mu_{z_2}^2$ for different values of N_c and N_b at SNR=-10 dB.

where d and n are positive integers. For simplicity of notation, let us consider $\mu_b^{k_d:k_d} = \mu_b^{k_d}$. By employing the linearity property of the statistical expectation and using (4.66), it can be easily shown that

$$\begin{aligned} \mathbb{E}\{x_\ell\} &= \frac{\Gamma_\ell}{N_b N_c (N_c - 1)} \sum_b^{N_b} \sum_{k_1:k_2}^{N_c} \mathbb{E}\{\mu_b^{k_1:k_2}\} - \frac{\Gamma_\ell + 1}{N_b (N_b - 1) N_c (N_c - 1)} \sum_{b_1:b_2}^{N_b} \sum_{k_1:k_2}^{N_c} \mathbb{E}\{\mu_{b_1}^{k_1}\} \mathbb{E}\{\mu_{b_2}^{k_2}\} \\ &\quad + \frac{2\sigma_w^2}{N_b N_c} \sum_b^{N_b} \sum_k^{N_c} \mathbb{E}\{\mu_b^k\} - \sigma_w^4. \end{aligned} \quad (4.67)$$

With the received symbols in each block being independent, it is obtained that $\mu_b^{k_d:k_{n+d-1}} = (\mu_b)^n$, and from (4.3) we have

$$\mathbb{E}\{\mu_b^{k_d:k_{n+d-1}}\} = \mathbb{E}\{(\mu_b)^n\} = \mu^{(n)}. \quad (4.68)$$

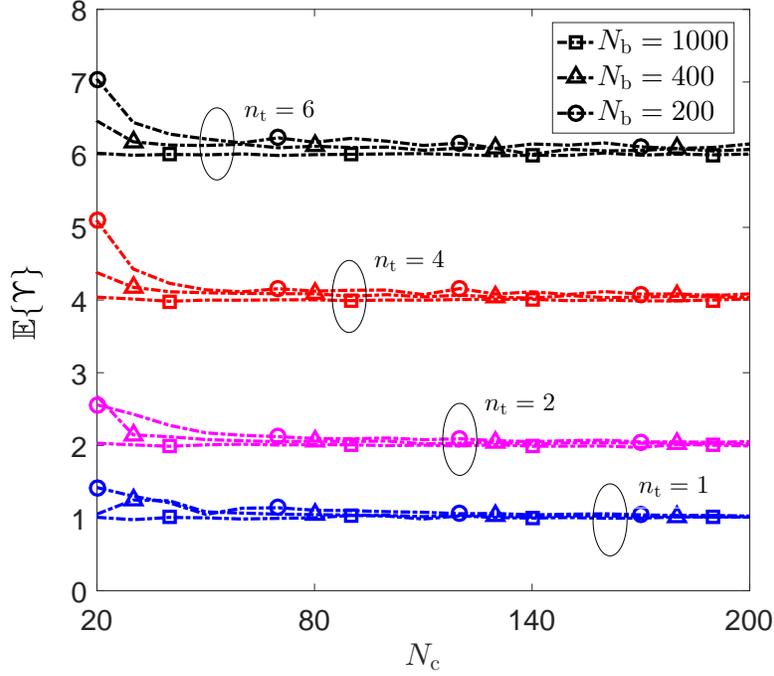


Fig. 4.C.3: $\mathbb{E}\{\Upsilon\}$ for different values of N_c and N_b at SNR=-10 dB.

By using (4.68) for $n = 2$, (4.67) can be re-written as

$$\mathbb{E}\{x_\ell\} = \Gamma_\ell \mu^{(2)} - (\Gamma_\ell + 1) (\mu^{(1)})^2 + 2\sigma_w^2 \mu^{(1)} - \sigma_w^4. \quad (4.69)$$

Finally, by substituting (4.4) and (4.5) in (4.69), and using that $\mathbb{E}\{y_\ell\} = \mathbb{E}\{x_\ell\} / \sigma_s^4 \sigma_h^4$, (4.30) is obtained.

4.E Proof of the Covariance Matrix

For the variance of y_ℓ , $\ell = n_t, n_t - 1$, we have

$$\text{Var}\{y_\ell\} = \frac{\text{Var}\{x_\ell\}}{\sigma_s^8 \sigma_h^8}, \quad (4.70)$$

where

$$\text{Var}\{x_\ell\} = \mathbb{E}\{x_\ell^2\} - \mathbb{E}\{x_\ell\}^2. \quad (4.71)$$

Let us define

$$\vartheta_b^{k_d} = \mathbb{E}\{|r_{k_d,b}|^4 | \mathbf{h}_b\} \quad (4.72)$$

and

$$\vartheta_b^{k_d:k_{d+1}} = \mathbb{E}\{|r_{k_d,b}|^4 |r_{k_{d+1},b}|^4 | \mathbf{h}_b\}. \quad (4.73)$$

By using (4.65), (4.66), (4.72), (4.73), and after several calculations, $\mathbb{E}\{x_\ell^2\}$ can be written as in (4.74).

$$\begin{aligned} \mathbb{E}\{x_\ell^2\} &= \frac{\Gamma_\ell^2}{N_b^2 N_c^2 (N_c - 1)^2} A_1 + \frac{(\Gamma_\ell + 1)^2}{N_b^2 (N_b - 1)^2 N_c^2 (N_c - 1)^2} A_2 + \left(\frac{2\sigma_w^2}{N_b N_c}\right)^2 A_3 + \sigma_w^8 \\ &\quad - \frac{2\Gamma_\ell (\Gamma_\ell + 1)}{N_b^2 N_c^2 (N_c - 1)^2 (N_b - 1)} A_4 + \frac{4\Gamma_\ell \sigma_w^2}{N_b^2 N_c^2 (N_c - 1)} A_5 - \frac{4(\Gamma_\ell + 1) \sigma_w^2}{N_b^2 (N_b - 1) N_c^2 (N_c - 1)} A_6 \\ &\quad - \frac{2\Gamma_\ell \sigma_w^4}{N_b N_c (N_c - 1)} A_7 - \frac{4\sigma_w^6}{N_b N_c} A_8 + \frac{2(\Gamma_\ell + 1) \sigma_w^4}{N_b (N_b - 1) N_c (N_c - 1)} A_9, \end{aligned} \quad (4.74)$$

where

$$\begin{aligned} A_1 &= 2 \sum_{b_1}^{N_b} \sum_{k_1:k_2}^{N_c} \mathbb{E}\{\vartheta_{b_1}^{k_1:k_2}\} + 4 \sum_{b_1}^{N_b} \sum_{k_1:k_3}^{N_c} \mathbb{E}\{\vartheta_{b_1}^{k_1} \mu_{b_1}^{k_2:k_3}\} + \sum_{b_1}^{N_b} \sum_{k_1:k_4}^{N_c} \mathbb{E}\{\mu_{b_1}^{k_1:k_4}\} + 2 \sum_{b_1:b_2}^{N_b} \sum_{k_1:k_2}^{N_c} \mathbb{E}\{\mu_{b_1}^{k_1:k_2} \mu_{b_2}^{k_1:k_2}\} \\ &\quad + 4 \sum_{b_1:b_2}^{N_b} \sum_{k_1:k_3}^{N_c} \mathbb{E}\{\mu_{b_1}^{k_1:k_2} \mu_{b_2}^{k_2:k_3}\} + \sum_{b_1:b_2}^{N_b} \sum_{k_1:k_4}^{N_c} \mathbb{E}\{\mu_{b_1}^{k_1:k_2} \mu_{b_2}^{k_3:k_4}\}, \end{aligned}$$

$$\begin{aligned}
A_2 = & 2 \sum_{b_1:b_2}^{N_b} \sum_{k_1:k_2}^{N_c} \mathbb{E}\{\vartheta_{b_1}^{k_1} \vartheta_{b_2}^{k_2}\} + 2 \sum_{b_1:b_2}^{N_b} \sum_{k_1:k_2}^{N_c} \mathbb{E}\{\mu_{b_1}^{k_1:k_2} \mu_{b_2}^{k_1:k_2}\} + 4 \sum_{b_1:b_2}^{N_b} \sum_{k_1:k_3}^{N_c} \mathbb{E}\{\vartheta_{b_1}^{k_1} \mu_{b_2}^{k_2:k_3}\} \\
& + 4 \sum_{b_1:b_2}^{N_b} \sum_{k_1:k_3}^{N_c} \mathbb{E}\{\mu_{b_1}^{k_1:k_2} \mu_{b_2}^{k_2:k_3}\} + 2 \sum_{b_1:b_2}^{N_b} \sum_{k_1:k_4}^{N_c} \mathbb{E}\{\mu_{b_1}^{k_1:k_2} \mu_{b_2}^{k_3:k_4}\} + 4 \sum_{b_1:b_3}^{N_b} \sum_{k_1:k_2}^{N_c} \mathbb{E}\{\vartheta_{b_1}^{k_1} \mu_{b_2}^{k_2} \mu_{b_3}^{k_2}\} \\
& + 4 \sum_{b_1:b_3}^{N_b} \sum_{k_1:k_2}^{N_c} \mathbb{E}\{\mu_{b_1}^{k_1:k_2} \mu_{b_2}^{k_1} \mu_{b_3}^{k_2}\} + 4 \sum_{b_1:b_3}^{N_b} \sum_{k_1:k_3}^{N_c} \mathbb{E}\{\vartheta_{b_1}^{k_1} \mu_{b_2}^{k_2} \mu_{b_3}^{k_3}\} + 8 \sum_{b_1:b_3}^{N_b} \sum_{k_1:k_3}^{N_c} \mathbb{E}\{\mu_{b_1}^{k_1:k_2} \mu_{b_2}^{k_2} \mu_{b_3}^{k_3}\} \\
& + 4 \sum_{b_1:b_3}^{N_b} \sum_{k_1:k_3}^{N_c} \mathbb{E}\{\mu_{b_1}^{k_1:k_2} \mu_{b_2}^{k_3} \mu_{b_3}^{k_3}\} + 4 \sum_{b_1:b_3}^{N_b} \sum_{k_1:k_4}^{N_c} \mathbb{E}\{\mu_{b_1}^{k_1} \mu_{b_2}^{k_2:k_3} \mu_{b_3}^{k_4}\} + 2 \sum_{b_1:b_4}^{N_b} \sum_{k_1:k_2}^{N_c} \mathbb{E}\{\mu_{b_1}^{k_1} \mu_{b_2}^{k_2} \mu_{b_3}^{k_1} \mu_{b_4}^{k_2}\} \\
& + 4 \sum_{b_1:b_4}^{N_b} \sum_{k_1:k_3}^{N_c} \mathbb{E}\{\mu_{b_1}^{k_1} \mu_{b_2}^{k_2} \mu_{b_3}^{k_2} \mu_{b_4}^{k_3}\} + \sum_{b_1:b_4}^{N_b} \sum_{k_1:k_4}^{N_c} \mathbb{E}\{\mu_{b_1}^{k_1} \mu_{b_2}^{k_2} \mu_{b_3}^{k_3} \mu_{b_4}^{k_4}\}
\end{aligned}$$

,

$$A_3 = \sum_{b_1}^{N_b} \sum_{k_1}^{N_c} \mathbb{E}\{\vartheta_{b_1}^{k_1}\} + \sum_{b_1}^{N_b} \sum_{k_1:k_2}^{N_c} \mathbb{E}\{\mu_{b_1}^{k_1:k_2}\} + \sum_{b_1:b_2}^{N_b} \sum_{k_1}^{N_c} \mathbb{E}\{\mu_{b_1}^{k_1} \mu_{b_2}^{k_1}\} + \sum_{b_1:b_2}^{N_b} \sum_{k_1:k_2}^{N_c} \mathbb{E}\{\mu_{b_1}^{k_1} \mu_{b_2}^{k_2}\},$$

$$\begin{aligned}
A_4 = & 4 \sum_{b_1:b_2}^{N_b} \sum_{k_1:k_2}^{N_c} \mathbb{E}\{\vartheta_{b_1}^{k_1} \mu_{b_1}^{k_2} \mu_{b_2}^{k_2}\} + 4 \sum_{b_1:b_2}^{N_b} \sum_{k_1:k_3}^{N_c} \mathbb{E}\{\vartheta_{b_1}^{k_1} \mu_{b_1}^{k_2} \mu_{b_2}^{k_3}\} + 4 \sum_{b_1:b_2}^{N_b} \sum_{k_1:k_3}^{N_c} \mathbb{E}\{\mu_{b_1}^{k_1:k_3} \mu_{b_2}^{k_1}\} \\
& + 2 \sum_{b_1:b_2}^{N_b} \sum_{k_1:k_4}^{N_c} \mathbb{E}\{\mu_{b_1}^{k_1:k_3} \mu_{b_2}^{k_4}\} + 2 \sum_{b_1:b_3}^{N_b} \sum_{k_1:k_2}^{N_c} \mathbb{E}\{\mu_{b_1}^{k_1:k_2} \mu_{b_2}^{k_1} \mu_{b_3}^{k_2}\} + 4 \sum_{b_1:b_3}^{N_b} \sum_{k_1:k_3}^{N_c} \mathbb{E}\{\mu_{b_1}^{k_1:k_2} \mu_{b_2}^{k_1} \mu_{b_3}^{k_3}\} \\
& + \sum_{b_1:b_3}^{N_b} \sum_{k_1:k_4}^{N_c} \mathbb{E}\{\mu_{b_1}^{k_1:k_2} \mu_{b_2}^{k_3} \mu_{b_3}^{k_4}\}
\end{aligned}$$

,

$$A_5 = 2 \sum_{b_1}^{N_b} \sum_{k_1:k_2}^{N_c} \mathbb{E}\{\vartheta_{b_1}^{k_1} \mu_{b_1}^{k_2}\} + \sum_{b_1}^{N_b} \sum_{k_1:k_3}^{N_c} \mathbb{E}\{\mu_{b_1}^{k_1:k_3}\} + 2 \sum_{b_1:b_2}^{N_b} \sum_{k_1:k_2}^{N_c} \mathbb{E}\{\mu_{b_1}^{k_1:k_2} \mu_{b_2}^{k_1}\} + \sum_{b_1:b_2}^{N_b} \sum_{k_1:k_3}^{N_c} \mathbb{E}\{\mu_{b_1}^{k_1:k_2} \mu_{b_2}^{k_3}\},$$

$$\begin{aligned}
A_6 &= 2 \sum_{b_1:b_2}^{N_b} \sum_{k_1:k_2}^{N_c} \mathbb{E}\{\vartheta_{b_1}^{k_1} \mu_{b_2}^{k_2}\} + 2 \sum_{b_1:b_2}^{N_b} \sum_{k_1:k_2}^{N_c} \mathbb{E}\{\mu_{b_1}^{k_1:k_2} \mu_{b_2}^{k_2}\} + 2 \sum_{b_1:b_2}^{N_b} \sum_{k_1:k_3}^{N_c} \mathbb{E}\{\mu_{b_1}^{k_1:k_2} \mu_{b_2}^{k_3}\} \\
&\quad + 2 \sum_{b_1:b_3}^{N_b} \sum_{k_1:k_2}^{N_c} \mathbb{E}\{\mu_{b_1}^{k_1} \mu_{b_2}^{k_2} \mu_{b_3}^{k_1}\} + \sum_{b_1:b_3}^{N_b} \sum_{k_1:k_3}^{N_c} \mathbb{E}\{\mu_{b_1}^{k_1} \mu_{b_2}^{k_2} \mu_{b_3}^{k_3}\}
\end{aligned}$$

$$, A_7 = \sum_{b_1}^{N_b} \sum_{k_1:k_2}^{N_c} \mathbb{E}\{\mu_{b_1}^{k_1:k_2}\} , A_8 = \sum_{b_1}^{N_b} \sum_{k_1}^{N_c} \mathbb{E}\{\mu_{b_1}^{k_1}\} , \text{ and } A_9 = \sum_{b_1:b_2}^{N_b} \sum_{k_1:k_2}^{N_c} \mathbb{E}\{\mu_{b_1}^{k_1} \mu_{b_2}^{k_2}\}.$$

Further, with the same justification as in Appendix 4.D with the received symbols in each block being independent, it can be seen that $\vartheta_b^{k_d:k_{d+1}} = (\vartheta_b^{k_d})^2$ and $\vartheta_b^{k_d} = \vartheta_b$, where ϑ_b is obtained by expanding $|r_{k,b}|^4$ and taking the statistical expectation over the signal and noise distributions, as

$$\vartheta_b = \mathbb{E}\left\{|r_{k_d,b}|^4 \mid \mathbf{h}_b\right\} = \Omega_s \sigma_s^4 \sum_{m_1}^{n_t} |h_b^{(m_1)}|^4 + 2\sigma_s^4 \sum_{m_1:m_2}^{n_t} |h_b^{(m_1)}|^2 |h_b^{(m_2)}|^2 + 4\sigma_s^2 \sigma_w^2 \sum_{m_1}^{n_t} |h_b^{(m_1)}|^2 + 2\sigma_w^4, \quad (4.75)$$

Moreover, by employing (4.2), (4.46), (4.66), (4.68), and (4.75), analogous to Appendix 4.A, it can easily be shown that

$$\begin{aligned}
\mu^{(3)} &= \mathbb{E}\{\mu_b^{k_d:k_{d+2}}\} = \mathbb{E}\{(\mu_b)^3\} = (n_t^3 + 3n_t^2 + 2n_t)\sigma_s^6 \sigma_h^6 + \sigma_w^6 \\
&\quad + 3(n_t^2 + n_t)\sigma_s^4 \sigma_h^4 \sigma_w^2 + 3n_t \sigma_s^2 \sigma_h^2 \sigma_w^4
\end{aligned} \quad (4.76)$$

$$\begin{aligned}
\mu^{(4)} &\triangleq \mathbb{E}\{\mu_b^{k_d:k_{d+3}}\} = \mathbb{E}\{(\mu_b)^4\} = (n_t^4 + 6n_t^3 + 11n_t^2 + 6n_t)\sigma_s^8 \sigma_h^8 + \sigma_w^8 \\
&\quad + 4(n_t^3 + 3n_t^2 + 2n_t)\sigma_s^6 \sigma_h^6 \sigma_w^2 + 6(n_t^2 + n_t)\sigma_s^4 \sigma_h^4 \sigma_w^4 + 4n_t \sigma_s^2 \sigma_h^2 \sigma_w^6
\end{aligned} \quad (4.77)$$

$$\vartheta^{(1)} \triangleq \mathbb{E}\{\vartheta_b^{k_d}\} = \mathbb{E}\{\vartheta_b\} = 2(n_t^2 + (\Omega_s - 1)n_t)\sigma_s^4\sigma_h^4 + 4n_t\sigma_s^2\sigma_h^2\sigma_w^2 + 2\sigma_w^4 \quad (4.78)$$

$$\begin{aligned} \vartheta^{(2)} \triangleq \mathbb{E}\{\vartheta_b^{k_d:k_{d+1}}\} &= \mathbb{E}\{(\vartheta_b)^2\} = (4n_t^4 + 8(\Omega_s + 1)n_t^3 + (4\Omega_s^2 + 24\Omega_s - 20)n_t^2 \\ &+ (20\Omega_s^2 - 32\Omega_s + 8)n_t)\sigma_s^8\sigma_h^8 + 4\sigma_w^8 + (16n_t^3 + 16(\Omega_s + 1)n_t^2 + 32(\Omega_s - 1)n_t)\sigma_s^6\sigma_h^6\sigma_w^2 \\ &+ (24n_t^2 + 8(\Omega_s + 1)n_t)\sigma_s^4\sigma_h^4\sigma_w^4 + 16n_t\sigma_s^2\sigma_h^2\sigma_w^6 \end{aligned} \quad (4.79)$$

$$\begin{aligned} \kappa^{(2)} \triangleq \mathbb{E}\{\vartheta_b^{k_d}\mu_b^{k_{d'}}\} &= \mathbb{E}\{\vartheta_b\mu_b\} = (2n_t^3 + 2(\Omega_s + 1)n_t^2 + (4\Omega_s - 4)n_t)\sigma_s^6\sigma_h^6 + 2\sigma_w^6 \\ &+ (6n_t^2 + 2(\Omega_s + 1)n_t)\sigma_s^4\sigma_h^4\sigma_w^2 + 6n_t\sigma_s^2\sigma_h^2\sigma_w^4 \end{aligned} \quad (4.80)$$

and

$$\begin{aligned} \kappa^{(3)} \triangleq \mathbb{E}\{\vartheta_b^{k_{d'}}\mu_b^{k_d:k_{d+1}}\} &= \mathbb{E}\{\vartheta_b(\mu_b)^2\} = (2n_t^4 + (8 + 2\Omega_s)n_t^3 + (2 + 10\Omega_s)n_t^2 \\ &+ 12(\Omega_s - 1)n_t)\sigma_s^8\sigma_h^8 + 2\sigma_w^8 + (8n_t^3 + 4(4 + \Omega_s)n_t^2 + 8\Omega_s n_t)\sigma_s^6\sigma_h^6\sigma_w^2 \\ &+ (12n_t^2 + (8 + 2\Omega_s)n_t)\sigma_s^4\sigma_h^4\sigma_w^4 + 8n_t\sigma_s^2\sigma_h^2\sigma_w^6. \end{aligned} \quad (4.81)$$

Finally, by using the assumption that the channel coefficients are independent for different blocks, substituting (4.4), (4.5), (4.76), (4.77), (4.78), (4.79), (4.80), and (4.81) into (4.74), and employing the corresponding result and (4.69) with (4.71) and then (4.70), (4.33) is easily obtained.

For the the covariance

$$\text{Cov}\{y_{n_t}, y_{n_t-1}\} = \frac{\text{Cov}\{x_{n_t}, x_{n_t-1}\}}{\sigma_s^8\sigma_h^8} = \frac{\mathbb{E}\{x_{n_t}x_{n_t-1}\} - \mathbb{E}\{x_{n_t}\}\mathbb{E}\{x_{n_t-1}\}}{\sigma_s^8\sigma_h^8}, \quad (4.82)$$

by following similar steps as for the variance, (4.34) is obtained.

Chapter 5

SNR and Noise Variance

Estimation in MIMO Time-varying Channels

5.1 Introduction

Wireless communication systems often require knowledge of the signal-to-noise ratio (SNR). The SNR is considered a key parameter whose *a priori* knowledge can be exploited at both receiver and transmitter (through feedback) in order to achieve the optimal performance in adaptive systems. Estimation of the SNR is typically employed in cognitive radio, power control, mobile assisted handoff, and adaptive modulation schemes, as well as soft decoding procedures [110, 170–174].

In general, SNR estimation can be divided into two major categories, depending on whether they base the estimation process on the knowledge of the transmitted symbols or not. Methods that base the estimation only on the received signal and do not need the *a priori* knowledge of the transmitted symbols are called non-data-aided

(NDA) methods, while data-aided (DA) methods assume perfect knowledge of some transmitted symbols (for example, training sequences provided for synchronization and equalization), to facilitate the estimation process. The DA methods have the drawback of limiting the system throughput due to the transmission of known data. Also, in many applications, such as cognitive radio systems, where the primary users' pilot symbol locations might not be known to the secondary users, DA methods cannot be employed [171]

Accurate SNR and noise variance estimation are extremely important tasks in the design of multiple-input multiple-output (MIMO) systems [125]. Typically, these parameters in MIMO systems are estimated through complex algorithms. Hence, the development of the MIMO systems has led to intense research on accurate and low-complexity SNR estimation [170].

5.1.1 Literature Review

Different types of DA and NDA SNR estimators exist in the literature. Those estimators are in general categorized as maximum likelihood (ML)-based, expectation maximization (EM)-based, Kolmogorov-Smirnov (K-S)-based, moments-based (MB), and decision-directed (DD)-based estimators [175–179].

The DA SNR estimators are mainly ML-based estimators and its variations. However, the NDA ML-based SNR estimators have also been developed. The NDA EM-based SNR estimators are employed to iteratively attain the ML estimate. This is achieved by averaging the likelihood function, at each iteration, over all the possible values of the unknown transmitted symbols [177].

The mainly employed NDA SNR estimators are the moment-based (MB) estimators. The second- and fourth-order moments of the received signal are mostly employed in the MB SNR estimators. However, for multilevel constellations, the es-

timation variance considerably increases as the SNR increases [179, 180]. Recently, NDA SNR estimators through higher-order moments have been proposed. They exhibit significant performance improvement for multilevel constellations at intermediate and high SNRs [178]. The NDA Kolmogorov-Smirnov (K-S)-based SNR estimators extract the empirical cumulative distribution function (ECDF) of a certain decision statistic from the received signal and compare it with pre-stored cumulative distribution function (CDF)s or ECDFs of the reference signals with known SNRs. Then, the specific SNR, with which the pre-stored CDFs or ECDFs is the most closest to the ECDF of the received signal, is selected as the estimate [175, 176]. It should be mentioned that the MB and K-S-based estimators do not attain the NDA-Cramer-Rao lower bound (CRLB) over a wide range of SNRs. The DD-based estimators base the estimation process on the detected transmitted symbols. Hence, these estimators require the transmission of fewer known data symbols, but they may suffer from erroneous detections [181, 182].

While SNR estimation for single-input single-output (SISO) systems has been extensively studied for both DA and NDA cases, a few number of estimators have been developed for MIMO systems. Moreover, to the best of our knowledge, most of the current SNR estimators for both SISO and MIMO systems, have been developed for time-invariant frequency-flat fading channels [183, 184], as well as for time-invariant, frequency-selective fading channels [170, 182]. However, in many applications requiring SNR estimation (e.g., mobile communications), this assumption is not valid.

In conjunction with single-input multiple-output (SIMO) and MIMO systems, the DA and NDA ML-based SNR estimators in time-invariant frequency-flat fading channel have been investigated in [170, 185, 186]. The ML-based SNR estimation in SISO time-varying channel was investigated in [171, 172, 187] for the DA and NDA scenarios. Moreover, in [110, 188], ML-based DA SNR estimation in SIMO and

MIMO time-varying channel was studied. Furthermore, the CRLBs of the NDA SNR estimation and efficient estimators in SISO time-invariant frequency-selective channel have been evaluated in [189]. These have been extended to SIMO systems in [173] for both DA as well as NDA models

The problem of the NDA SNR estimation through moments of the received signal in SISO time-invariant and frequency-flat fading channel has been investigated in [190–194], and M_1M_4 , M_2M_4 , and $M_1M_4M_6$ MB estimators have been developed. The classic M_1M_4 estimator performs well with respect to constant modulus schemes like M -PSK, whereas a severe degradation is observed for multilevel constellations in the medium-to-high SNR range. Recently, a new MB approach has been proposed, which exploits a linear combination of ratios for certain even-order moments (up to the eight-order moment) to improve on previous estimators of this class [178]. However, the weights of the linear combination should be tuned according to the constellation and the specific SNR operation range. The problem of MB SNR estimation in SIMO and MIMO time-invariant and frequency-flat fading channel has been explored in [179, 192]. To the best of our knowledge, the problem of MB SNR estimation in time-varying frequency-flat fading channel was investigated in [195].

5.1.2 Motivation

After reviewing the current SNR estimators in the literature, the following observations can be made:

- The problem of NDA SNR estimation in MIMO time-varying channel has not been extensively investigated;
- The current MB SNR estimators exhibit high estimation variance for multilevel constellations;

- The existing DA SNR estimators in MIMO time-varying fading channel suffer from huge computational complexity. Moreover, the periodic pilot transmission in DA estimators results in reduced system capacity;
- The performance of the DA SNR estimators in MIMO time-varying channels are severely affected by channel impairments, such as lack of time-frequency synchronization;
- The current DA and NDA SNR estimators for MIMO systems require *a priori* knowledge about the transmitted signals, such as the number of transmit antennas and modulation format.

5.1.3 Problem Statement

The specific research problems which are studied in Chapter 5 of this thesis are as follows:

- Two low-complexity MB-SNR estimators for MIMO time-varying frequency-flat fading channel are proposed;
- The proposed estimators are developed based on the coherence time of the fading channel which can be estimated through the proposed maximum Doppler spread (MDS) estimators in Chapter 3 of this thesis.

5.1.4 Methodology

The statistical moment-based approach is employed in Chapter 5 of this thesis. While the proposed NDA M_2M_4 SNR estimator relies on the second- and fourth-order statistics of the received signal, the proposed NDA $M_2M_4M_6$ SNR estimator additionally

employs the sixth-order statistics of the received signal, as well. By using the sixth-order statistics, the NDA $M_2M_4M_6$ SNR estimator removes the need for *a priori* knowledge about the number of transmit antennas.

5.2 Moment-based SNR Estimation

In this section, we derive the M_2M_4 and $M_2M_4M_6$ NDA SNR estimators for MISO systems by employing higher-order moments of the received signal. Moreover, an extension to the MIMO systems is provided.

5.2.1 System Model

We consider a multiple-input single-output (MISO) block fading channel [161, 162] with n_t transmit antennas. It is assumed that N_b observation blocks, $b = 1, 2, \dots, N_b$, each with length N_c symbols, are affected by independent and identically distributed (i.i.d.) fading characterized by the $(n_t \times 1)$ vector \mathbf{h}_b , and corrupted by additive white Gaussian noise. A typical value for the block length, N_c , in the case of the Clarke-Jakes Doppler spectrum is $N_c = \lceil 9/16\pi f_d T_s \rceil$ [163], where f_d and T_s are the maximum Doppler frequency and the symbol period, respectively. The received complex-valued signal, $r_{k,b}$, is expressed as

$$r_{k,b} = \mathbf{h}_b^\dagger \mathbf{s}_{k,b} + w_{k,b}, \quad k = 1, 2, \dots, N_c, \quad b = 1, 2, \dots, N_b, \quad (5.1)$$

where $r_{k,b}$ is the k th received symbol in the b th observation block, $\mathbf{s}_{k,b} = [s_{k,b}^{(1)}, s_{k,b}^{(2)}, \dots, s_{k,b}^{(n_t)}]^\dagger$ represents the zero-mean transmitted symbols from the n_t transmit antennas, whose variance $\mathbb{E}\{|s_{k,b}^{(m)}|^2\} = \sigma_s^2$, $m = 1, 2, \dots, n_t$ is unknown at the receive-side, $w_{k,b}$ is complex additive white Gaussian noise with variance σ_w^2 , which is unknown, and $\mathbf{h}_b =$

$[\mathbf{h}_b^{(1)}, \mathbf{h}_b^{(2)}, \dots, \mathbf{h}_b^{(n_t)}]^\dagger$ denotes the channel coefficients, with $h_b^{(m)}$, $m = 1, 2, \dots, n_t$ as the channel coefficient between the m th transmit antenna and the receive antenna for the b th observation block. It is assumed that the channel coefficients in each block are uncorrelated complex Gaussian random variables with $\mathbb{E}\{h_b^{(m_1)}(h_b^{(m_2)})^*\} = \sigma_h^2 \delta(m_1 - m_2)$, where σ_h^2 is unknown. The average SNR is defined as $\gamma \triangleq n_t \sigma_s^2 \sigma_h^2 / \sigma_w^2$ at the received-side.

5.2.2 M_2M_4 SNR Estimator

The proposed M_2M_4 SNR estimator exploits the second- and fourth-order moments of the received signal, as follows:

Let us consider the second-order moment of the received signal in the b th observation block, which can be expressed as

$$\mu_b \triangleq \mathbb{E} \left\{ |r_{k,b}|^2 | h_b \right\} = \sigma_s^2 \sum_{m=1}^{n_t} |h_b^{(m)}|^2 + \sigma_w^2. \quad (5.2)$$

Further, let us define

$$M_d \triangleq \mathbb{E} \left\{ (\mu_b)^{d/2} \right\} = \mathbb{E} \left\{ \left(\mathbb{E} \left\{ |r_{k,b}|^2 | h_b \right\} \right)^{d/2} \right\} \quad (5.3)$$

where d is a positive integer. With the channel coefficients corresponding to different transmit antennas being independent complex Gaussian random variables with variance σ_h^2 , i.e., $\mathbb{E} \left\{ \mathbf{h}_b \mathbf{h}_b^H \right\} = \sigma_h^2 \mathbf{I}$, and by using (5.2), when $d = 2$ (second-order statistic), one can easily show that (5.3) can be expressed as

$$M_2 \triangleq \mathbb{E} \left\{ \mu_b \right\} = \sigma_s^2 \sum_{m=1}^{n_t} \mathbb{E} \left\{ |h_b^{(m)}|^2 \right\} + \sigma_w^2 = n_t \sigma_s^2 \sigma_h^2 + \sigma_w^2. \quad (5.4)$$

Furthermore, when $d = 4$ (fourth-order statistic), one can show that (5.3) becomes

$$\begin{aligned}
M_4 &= \mathbb{E} \left\{ \left(\sigma_s^2 \sum_{m_1}^{n_t} |h_b^{(m_1)}|^2 + \sigma_w^2 \right)^2 \right\} = \sigma_s^4 \sum_{m_1}^{n_t} \mathbb{E} \left\{ |h_b^{(m_1)}|^4 \right\} + 2\sigma_s^2 \sigma_w^2 \sum_{m_1}^{n_t} \mathbb{E} \left\{ |h_b^{(m_1)}|^2 \right\} \\
&\quad + \sigma_s^4 \sum_{m_1:m_2}^{n_t} \mathbb{E} \left\{ |h_b^{(m_1)}|^2 \right\} \mathbb{E} \left\{ |h_b^{(m_2)}|^2 \right\} + \sigma_w^4 = (n_t^2 + n_t) \sigma_s^4 \sigma_h^4 + 2n_t \sigma_s^2 \sigma_h^2 \sigma_w^2 + \sigma_w^4.
\end{aligned} \tag{5.5}$$

Equation (5.5) along with (5.4) lead to the following systems of equations

$$M_2 = n_t \sigma_s^2 \sigma_h^2 + \sigma_w^2 \tag{5.6}$$

$$M_4 = (n_t^2 + n_t) \sigma_s^4 \sigma_h^4 + 2n_t \sigma_s^2 \sigma_h^2 \sigma_w^2 + \sigma_w^4,$$

with $\sigma_s^2 \sigma_h^2$ and σ_w^2 as unknowns. By solving these equations, one obtains

$$\sigma_h^2 \sigma_s^2 = \sqrt{\frac{M_4 - M_2^2}{n_t}}, \tag{5.7}$$

and

$$\sigma_w^2 = M_2 - n_t \sqrt{\frac{M_4 - M_2^2}{n_t}}. \tag{5.8}$$

Hence, one easily obtains the average SNR, γ , as

$$\gamma \triangleq \frac{n_t \sigma_h^2 \sigma_s^2}{\sigma_w^2} = \frac{n_t \sqrt{\frac{M_4 - M_2^2}{n_t}}}{M_2 - n_t \sqrt{\frac{M_4 - M_2^2}{n_t}}}. \tag{5.9}$$

It should be noted that the term inside the square root in the right hand-side of (5.9) is a positive value as $M_4 - M_2^2$ represents the variance of μ_b . In practice, the moments

in (5.9), i.e., M_2 and M_4 are replaced by their time-averages as,

$$\hat{m}_2 = \frac{1}{N_b N_c} \sum_{b_1}^{N_b} \sum_{k_1=1}^{N_c} |r_{k_1, b_1}|^2 \quad (5.10)$$

and

$$\hat{m}_4 = \frac{1}{N_b N_c (N_c - 1)} \sum_{b_1}^{N_b} \sum_{k_1:k_2}^{N_c} |r_{k_1, b_1}|^2 |r_{k_2, b_1}|^2. \quad (5.11)$$

By substituting (5.10) and (5.11) in (5.9), one easily obtains the estimated average SNR, $\hat{\gamma}$, as

$$\hat{\gamma} = \frac{n_t \sqrt{\frac{\hat{m}_4 - \hat{m}_2^2}{n_t}}}{\hat{m}_2 - n_t \sqrt{\frac{\hat{m}_4 - \hat{m}_2^2}{n_t}}}. \quad (5.12)$$

It should be mentioned that $\hat{m}_4 - \hat{m}_2^2$ represents a positive value as it is the sample variance of $\hat{\mu}_b$.

5.2.3 $M_2M_4M_6$ SNR Estimator

To remove the need for *a priori* knowledge about the number of transmit antennas in the proposed M_2M_4 estimator, the sixth-order moment of the received signal can be used along with the second- and fourth-order moments to derive the $M_2M_4M_6$ SNR estimator.

By replacing (5.2) in the definition of M_6 , it is straightforward to write that

$$M_6 \triangleq \mathbb{E} \left\{ \left(\sigma_s^2 \sum_{m=1}^{n_t} |h_b^{(m)}|^2 + \sigma_w^2 \right)^3 \right\}. \quad (5.13)$$

Furthermore, by employing $(a + b)^3 = a^3 + b^3 + 3a^2b + 3b^2a$ and applying the linearity

property of statistical expectation, one can easily express (5.13) as

$$\begin{aligned}
M_6 &= \sigma_s^6 \left[\sum_{m_1}^{n_t} \mathbb{E} \left\{ \left| h_b^{(m_1)} \right|^6 \right\} + 3 \sum_{m_1:m_2}^{n_t} \mathbb{E} \left\{ \left| h_b^{(m_1)} \right|^4 \right\} \mathbb{E} \left\{ \left| h_b^{(m_2)} \right|^2 \right\} \right. \\
&\quad + \sum_{m_1:m_3}^{n_t} \mathbb{E} \left\{ \left| h_b^{(m_1)} \right|^2 \right\} \mathbb{E} \left\{ \left| h_b^{(m_2)} \right|^2 \right\} \mathbb{E} \left\{ \left| h_b^{(m_3)} \right|^2 \right\} \left. \right] \\
&\quad + 3\sigma_s^2 \sigma_w^4 \sum_{m_1}^{n_t} \mathbb{E} \left\{ \left| h_b^{(m_1)} \right|^2 \right\} + 3\sigma_s^4 \sigma_w^2 \left[\sum_{m_1}^{n_t} \mathbb{E} \left\{ \left| h_b^{(m_1)} \right|^4 \right\} \right. \\
&\quad \left. + \sum_{m_1:m_2}^{n_t} \mathbb{E} \left\{ \left| h_b^{(m_1)} \right|^2 \right\} \mathbb{E} \left\{ \left| h_b^{(m_2)} \right|^2 \right\} \right] + \sigma_w^6. \tag{5.14}
\end{aligned}$$

With the channel coefficients corresponding to different antennas being independent circular complex Gaussian random variables with variance σ_h^2 , and employing the property of circular complex Gaussian random variable $x \sim \mathcal{N}_c(0, \sigma_x^2)$ that $\mathbb{E}\{|x|^{2n}\} = n! \sigma_x^{2n}$, one can obtain

$$\begin{aligned}
M_6 &= \sigma_s^6 \sigma_h^6 (6n_t + 6n_t(n_t - 1) + n_t(n_t - 1)(n_t - 2)) \\
&\quad + 3\sigma_s^4 \sigma_h^4 \sigma_w^2 (2n_t + n_t(n_t - 1)) + 3n_t \sigma_s^2 \sigma_h^2 \sigma_w^4 + \sigma_w^6 \\
&= (n_t^3 + 3n_t^2 + 2n_t) \sigma_s^6 \sigma_h^6 + 3(n_t^2 + n_t) \sigma_s^4 \sigma_h^4 \sigma_w^2 \\
&\quad + 3n_t \sigma_s^2 \sigma_h^2 \sigma_w^4 + \sigma_w^6. \tag{5.15}
\end{aligned}$$

Equation (5.15) along with (5.6) lead to the following system of equations

$$\begin{aligned}
M_6 &= (n_t^3 + 3n_t^2 + 2n_t) \sigma_s^6 \sigma_h^6 + 3(n_t^2 + n_t) \sigma_s^4 \sigma_h^4 \sigma_w^2 + 3n_t \sigma_s^2 \sigma_h^2 \sigma_w^4 + \sigma_w^6 \\
M_4 &= (n_t^2 + n_t) \sigma_s^4 \sigma_h^4 + 2n_t \sigma_s^2 \sigma_h^2 \sigma_w^2 + \sigma_w^4 \\
M_2 &= n_t \sigma_s^2 \sigma_h^2 + \sigma_w^2 \tag{5.16}
\end{aligned}$$

with n_t , σ_w^2 , and $\sigma_h^2 \sigma_s^2$ as unknowns. By solving these equations, it is straightforward

to obtain

$$n_t = \frac{4(M_4 - M_2^2)^3}{(M_6 + 2M_2^3 - 3M_2M_4)^2} \quad (5.17)$$

$$\sigma_h^2 \sigma_s^2 = \frac{2M_2^3 - 3M_4M_2 + M_6}{2(M_4 - M_2^2)}. \quad (5.18)$$

$$\sigma_w^2 = \frac{M_2^2M_4 + M_6M_2 - 2M_4^2}{2M_2^3 - 3M_4M_2 + M_6}. \quad (5.19)$$

Hence, by employing (5.17)-(5.19), one easily obtain the average SNR, γ , for the $M_2M_4M_6$ estimator as

$$\gamma \triangleq \frac{n_t \sigma_h^2 \sigma_s^2}{\sigma_w^2} = \frac{2(M_4 - M_2^2)^2}{M_2^2M_4 + M_6M_2 - 2M_4^2}. \quad (5.20)$$

In practice, analogous to the M_2M_4 estimator, M_2 and M_4 are replaced by their corresponding estimates given in (5.10) and (5.11), respectively. Also, for M_6 , the following estimator is employed

$$\hat{m}_6 = \frac{1}{N_{bc}} \sum_b^{N_b} \sum_{k_1:k_3}^{N_c} |r_{k_1,b}|^2 |r_{k_2,b}|^2 |r_{k_3,b}|^2 \quad (5.21)$$

where $N_{bc} \triangleq N_b N_c (N_c - 1)(N_c - 2)$. By substituting (5.10), (5.11), and (5.21) into (5.20), one easily obtains

$$\hat{\gamma} = \frac{2(\hat{m}_4 - \hat{m}_2^2)^2}{\hat{m}_2^2 \hat{m}_4 + \hat{m}_6 \hat{m}_2 - 2\hat{m}_4^2}. \quad (5.22)$$

5.2.4 Extension to MIMO Systems

Multiple receive antennas can improve the SNR estimation through space-diversity. When the noise variance is the same across the antenna elements, the performance of the M_2M_4 and $M_2M_4M_6$ estimators can be improved by estimating the SNR at each

receive antenna using (5.12) or (5.20), and then averaging over all antenna elements as

$$\hat{\gamma}_{\text{MIMO}} = \frac{1}{n_r} \sum_{i=1}^{n_r} \hat{\gamma}_i, \quad (5.23)$$

where n_r is the number of receive antennas, and $\hat{\gamma}_i$ is the estimated average SNR at the i th receive antenna.

5.3 Simulations

In this section, we examine the performance of the proposed NDA M_2M_4 and $M_2M_4M_6$ SNR estimators, through several simulation experiments.

5.3.1 Simulation Setup

Unless otherwise mentioned, the modulation was quadrature phase-shift-keying (QPSK) with spatial multiplexing transmission scheme. The channel coefficients were modeled as independent circular complex Gaussian random variables with variance σ_h^2 . The additive white noise was modeled as circular complex Gaussian random variables with variance σ_w^2 . Without loss of generality, it was assumed that $\sigma_w^2 = 1$, and the performance of the proposed estimators was evaluated in terms of the NRMSE, i.e., $\sqrt{\mathbb{E}\{(\hat{\gamma} - \gamma)^2\}}/\gamma$. Also, unless otherwise mentioned, $N_c = 100$, and $N_b = 2000$ and $N_b = 5000$ for the M_2M_4 and $M_2M_4M_6$ estimator, respectively.

5.3.2 Simulation Results

Fig. 5.1 shows the distributions of the estimated SNR, $\hat{\gamma}$, for the M_2M_4 and $M_2M_4M_6$ estimators for $n_t = 3$ at $\gamma = 0$ dB. As can be seen, the distributions can be approxi-

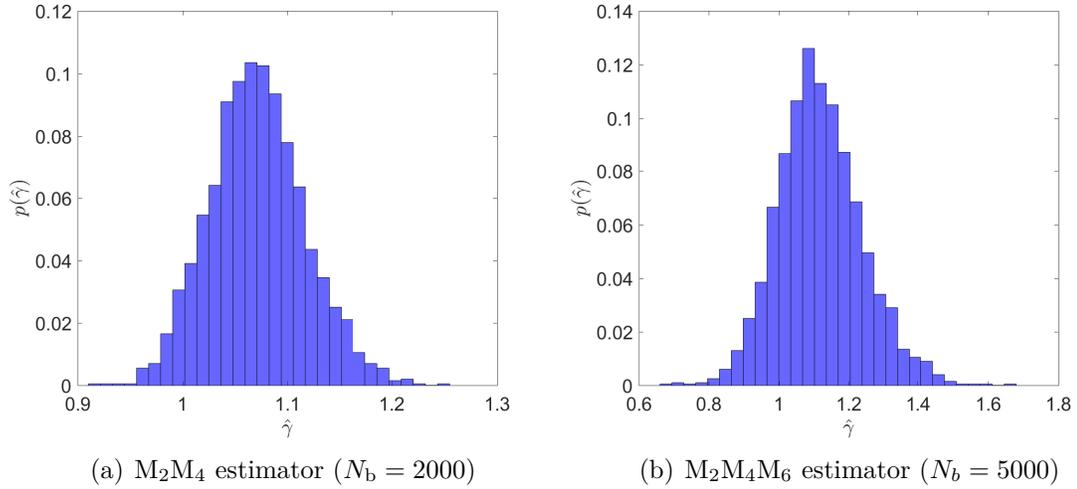


Fig. 5.1: Distribution of the estimated SNR, $\hat{\gamma}$, for the proposed estimators with $n_t = 2$ and at $\gamma = 0$ dB.

mated as Gaussian.

Fig. 5.2 illustrates the normalized bias, $\mathbb{E}\{\hat{\gamma} - \gamma\}/\gamma$, of the M_2M_4 estimator versus SNR, γ , for different numbers of transmit antennas, $n_t = 1, 2, 3$, and 4. As expected,

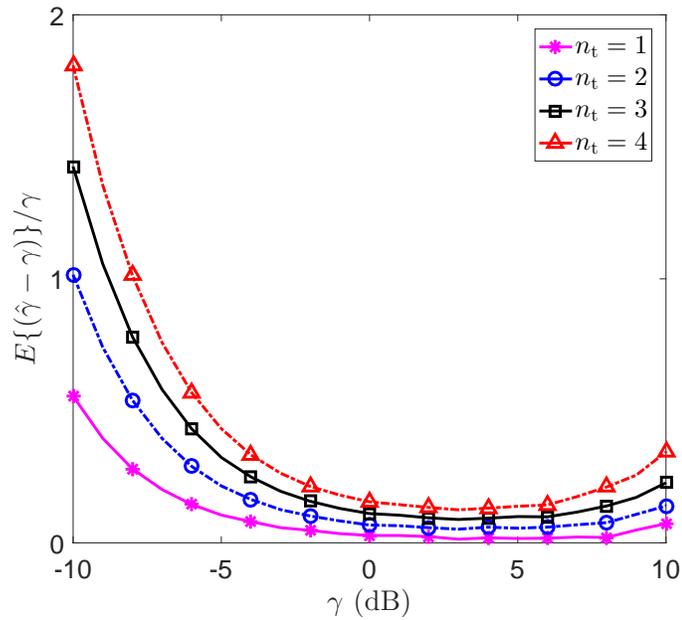


Fig. 5.2: The normalized bias of the M_2M_4 estimator versus SNR, γ , for $n_t = 1, 2, 3, 4$.

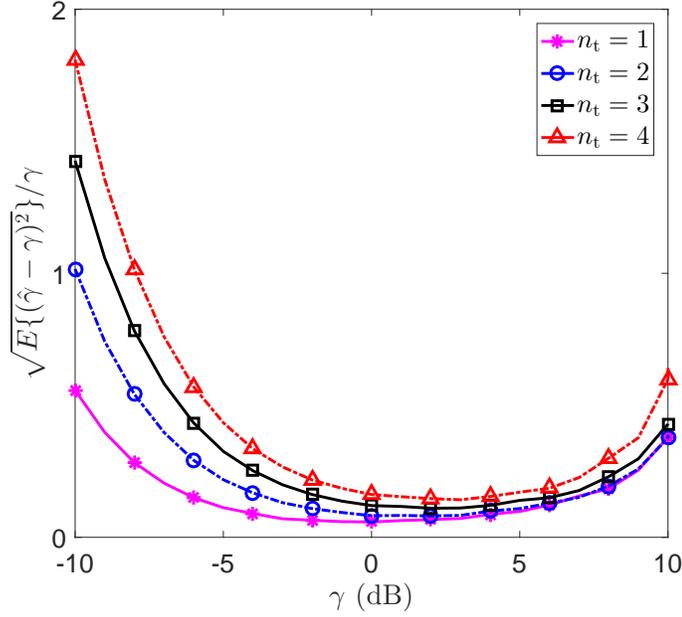


Fig. 5.3: NRMSE of the M_2M_4 estimator versus SNR, γ , for $n_t = 1, 2, 3, 4$.

the bias increases as the number of transmit antennas increases. As can also be seen, the bias is higher for $\gamma < -4$ dB and $\gamma > 8$ dB. This can be explained, as for $\gamma < -4$ dB and $\gamma > 8$ dB either σ_w^2 or $\sigma_h^2\sigma_s^2$ is not estimated accurately, which leads to higher estimation error in their ratio. However, for -4 dB $< \gamma < 8$ dB, both σ_w^2 and $\sigma_h^2\sigma_s^2$ are estimated reasonably accurate, which leads to trivial bias. It should be mentioned that a larger observation time results in a lower bias.

In Fig. 5.3, the NRMSE of the M_2M_4 estimator for different numbers of transmit antennas is illustrated. As can be seen, the proposed SNR estimator exhibits good performance over a wide-range of SNR. It should be mentioned that for $\gamma < -4$ dB and $\gamma > 8$ dB, the existence of the larger bias leads to higher NRMSE.

Fig. 5.4 presents the normalized bias, $\mathbb{E}\{\hat{\gamma} - \gamma\}/\gamma$, of the $M_2M_4M_6$ estimator for different numbers of transmit antennas, $n_t = 1, 2, 3$, and 4. As expected, in comparison to the performance of the M_2M_4 estimator in Fig. 5.2, the $M_2M_4M_6$ estimator

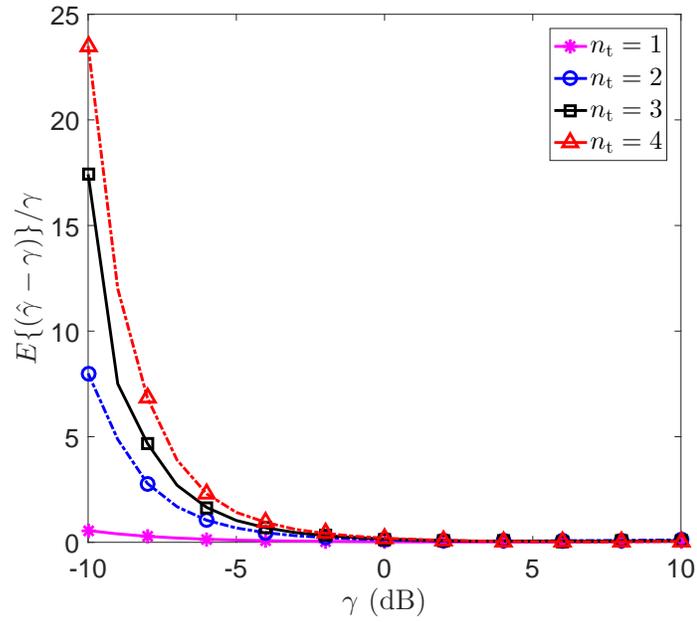


Fig. 5.4: The normalized bias of the $M_2M_4M_6$ estimator versus SNR, γ , for $n_t = 1, 2, 3, 4$.

exhibits a higher bias even for larger observation time at low SNR values. This can be explained, as the higher-order moment employed in the $M_2M_4M_6$ estimator results

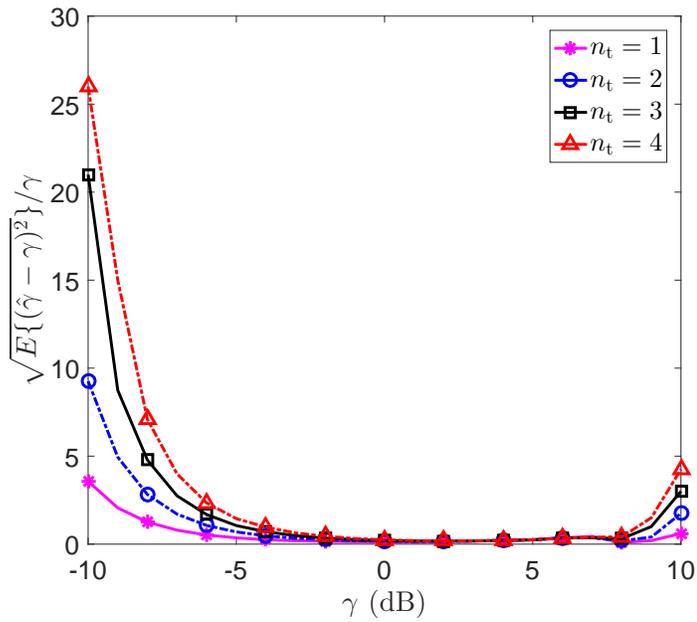


Fig. 5.5: NRMSE of the $M_2M_4M_6$ estimator versus SNR, γ , for $n_t = 1, 2, 3, 4$.

in larger estimation error.

In Fig. 5.5, the NRMSE of the $M_2M_4M_6$ estimator for different numbers of transmit antennas is illustrated. As can be seen, the proposed SNR estimator exhibits a good performance at $-4 \text{ dB} < \gamma < 8 \text{ dB}$ in terms of NRMSE.

5.4 Conclusions and Directions for Future Research

In Chapter 5 of this thesis, NDA SNR estimation was studied in MIMO time-varying frequency-flat fading channel. In this section, a summary of the main results in Chapter 5 is provided, and possible directions for future research are pointed out.

5.4.1 Summary

Motivated by the need for NDA SNR estimation in multiple antenna systems, two MB estimators, M_2M_4 and $M_2M_4M_6$, were developed for MISO systems in Chapter 5 of this thesis. Moreover, an extension of the proposed estimators to MIMO systems was investigated. The derived estimators rely on higher-order moments of the received signal and estimate SNR in time-varying frequency-flat fading channel. The effectiveness of the proposed SNR estimators was confirmed through simulations. It was shown that they exhibit small bias over a moderate range of SNRs. Furthermore, while the M_2M_4 estimator outperforms the $M_2M_4M_6$ estimator, they exhibit a good performance in terms of NRMSE. The proposed NDA SNR estimators exhibit the following advantages:

- They do not require perfect time synchronization;
- They are robust to the carrier frequency offset (CFO);
- They increase system capacity due to lack of pilots or preambles;

- They do not require *a priori* knowledge of other transmission parameters.
- The $M_2M_4M_6$ estimator does not require *a priori* knowledge of the number of transmit antennas.

5.4.2 Future research

The results in Chapter 5 of this thesis open interesting directions for a number of future research topics. Here, a few of them are outlined as follows:

- The optimal combining method for the proposed NDA MB SNR estimators in case of multiple receive antennas can be derived through the bootstrap technique;
- The proposed NDA MB SNR estimators can be extended for non-Gaussian noise;
- The NDA-CRLB for SNR estimation in MIMO time-varying frequency-selective fading channel can be investigated.

Chapter 6

Blind STBC Identification

6.1 Introduction

Signal identification is the process of identifying the type of the transmitted signal from a pool of candidates [11, 16]. It finds applications in radio surveillance, software defined radio, and spectrum awareness in cognitive radio [24, 25]. In the context of the adaptive transmission and software defined radio, due to the flexible architecture of the transmitters, the same hardware can be used for different transmission parameters, e.g., modulation format, coding rate, and antenna configuration. Accordingly, at the receive-side, robust algorithms are required for blind estimation of the signal parameters, such as modulation format, channel encoders, and transmit-side antenna configuration in a noisy environment [196]. In fact, signal identification is an intermediate step between signal detection and signal decoding/demodulation for adaptive transmission.

6.1.1 Literature Review

Signal identification has been extensively explored for single-input systems. These include identification of the modulation format [197–200], single versus multi-carrier transmissions [201], the type of multi-carrier technique [202, 203], and channel encoders [204], as well as blind parameter estimation [205].

With the advent of multiple-input systems, it has been adopted by different wireless standards, such as IEEE 802.11n, IEEE 802.16e, and 3GPP LTE [206]. Hence, the problem of signal identification for multiple-input systems, with new technical challenges, such as estimating the number of transmit antennas and type of transmit-side antenna configuration has emerged. Signal identification for multiple-input scenarios is at a very early stage. Estimation of the number of antennas [150–155], identification of the space-time block code (STBC) [7–9, 207–210], and modulation classification [211–213] have been considered in the literature.

Regarding the STBC identification, which is the focus of this chapter, the maximum likelihood (ML) [7] and feature based (FB)[8, 9, 208–210] approaches have been investigated. The former provides an optimal performance in the sense of maximizing the probability of correct identification. However, it requires knowledge of the channel and noise power, as well as timing and frequency synchronization, and is sensitive to model mismatches, such as carrier frequency offset (CFO) and timing offset (TO). With the FB approach, features are extracted from the received signal and a decision is made based on the observed feature values. Although the FB-based method is not optimal, it can provide a near optimal performance with reduced complexity.

Identification of STBCs has been studied for a pool of two candidate STBCs [8, 209, 210] and a larger pool of STBCs [7, 9, 208] in the literature. With the former, identification of spatial multiplexing (SM) and Alamouti space-time block code (AL-STBC) has been mainly considered, as these represent the most frequently

used STBCs in standards [206]. For the latter, the extension to a larger pool of STBCs is considered.

In [7], the likelihood-based algorithm for the identification of STBCs is investigated. While the proposed algorithm is optimal and capable of performing STBC identification with a single receive antenna, it requires perfect timing and frequency synchronization and its performance drops in the presence of model mismatches, such as TO, CFO and non-Gaussian noise. In [207], the space-time second-order correlation function is used as discriminating feature, and a subjective threshold setting is employed for decision making. In [214], the minimum distance between the theoretical and estimated space-time second-order correlation is employed for decision making. In [9, 215], the inherent cyclostationarity of the STBCs is employed as discriminating feature, and a cyclostationarity test is used for decision making. While these feature-based algorithms show lower sensitivity to the model mismatches when compared with the likelihood-based algorithm, they require multiple receive antennas and still do not exhibit a good performance under TO, CFO, and non-Gaussian noise. In [8], fourth-order moments are investigated as features for the identification of AL-STBC and SM. While the fourth-order moments algorithms can perform STBC identification with a single receive antenna, they are still sensitive to the CFO and TO. Furthermore, they do not show a good performance in the presence of non-Gaussian noise. In [208], the extension of the fourth-order moment algorithm for a larger STBC candidate pool by employing a decision binary tree is investigated.

6.1.2 Motivation

After reviewing the current algorithms for STBC identification, the following observations were made:

- New FB algorithm should be developed without requirement of a subjective

threshold setting for decision making;

- The existing algorithms are sensitive to the CFO and TO;
- STBC identification has not been explored for non-Gaussian noise;
- Due to the limitation in size and power of the intelligent/cognitive radios, signal identification with a single receive antenna is of practical interest.

6.1.3 Problem Statement

The specific research problems which are studied in Chapter 6 of this thesis are as follows:

- First, a non-parametric test is adopted to propose a new algorithm for identification of SM and AL-STBC, which provides significant advantages, such as:
 - It does not require a subjective threshold setting for decision making;
 - It does not require knowledge of the channel parameters, noise power, and modulation type. Also, it is robust to CFO and relatively robust to TO;
 - It is capable of performing the STBC identification with a single receive antenna, and it provides a good performance in the presence of non-Gaussian noise.
- Then, the proposed algorithm is extended for the identification of a larger STBC candidate pool.

6.1.4 Methodology

The proposed identification algorithm is developed based on the Kolmogrov-Smirnov (K-S) test, as a non-parametric goodness-of-fit test, along with decision binary tree.

6.2 Identification of SM and AL-STBC

6.2.1 Signal Model

Consider a wireless communication system using either SM or AL-STBC. Let n_t denote the number of transmit antennas and P be the number of time periods for the transmission of the b th block of symbols $\mathbf{s}_b = [s_{1,b}, \dots, s_{n_s,b}]^\dagger$. The encoder employs a transmission matrix $\mathbf{C}(\mathbf{s}_b)$ of size $n_t \times P$ to map the symbols to be transmitted. For SM, $n_t = n_s$ and $P = 1$, whereas $n_t = n_s = 2$ and $P = 2$ for AL-STBC. The coding matrix for SM and AL-STBC are given respectively as [8]

$$\mathbf{C}^{\text{SM}}(\mathbf{s}_b) = [s_{1,b}, \dots, s_{n_t,b}]^\dagger, \quad \mathbf{C}^{\text{AL}}(\mathbf{s}_b) = \begin{bmatrix} s_{1,b} & -s_{2,b}^* \\ s_{2,b} & s_{1,b}^* \end{bmatrix}. \quad (6.1)$$

Let us consider a receiver equipped with a single antenna. Without loss of generality, it is assumed that the first received symbol, denoted by $r(0)$, corresponds to the $(k_1 + 1)$ th column of the j th block, denoted by $\mathbf{C}_{k_1}(\mathbf{s}_j)$. Using this assumption, the k th received sample is given by [207]

$$r_k = \mathbf{h}^\dagger \mathbf{X}_k + w_k, \quad k = 0, 1, \dots, K - 1, \quad (6.2)$$

where $\mathbf{X}_k = \mathbf{C}_u(\mathbf{s}_v)$, with $u = (k + k_1) \bmod P$, $v = j + (k + k_1) \text{div } P$, and $z \bmod P$ and $z \text{div } P$ represent the remainder and the quotient of the division z/P , respectively, $w(k)$ is the additive white noise, and $\mathbf{h} = [h_1, \dots, h_{n_t}]^\dagger$ is the fading coefficients vector.

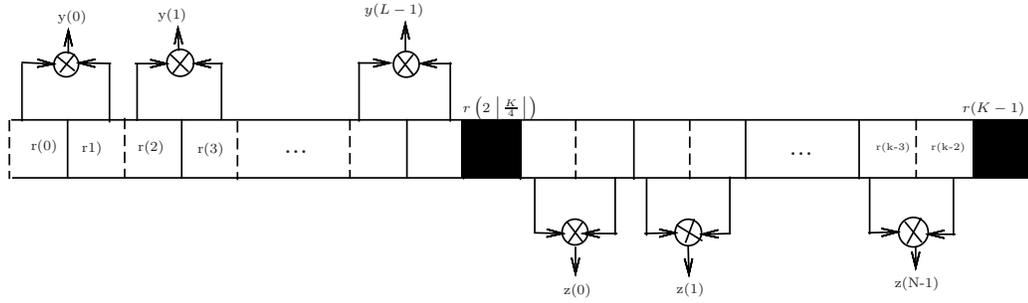


Fig. 6.1: Illustration of the event AL_1 . Solid lines are used to delimitate symbols which do not belong to the same block and dashed lines are employed to delimitate symbols which belong to the same block.

6.2.2 K-S Based Identification Algorithm

Let us define the random sequences \mathbf{y} and \mathbf{z} as

$$\mathbf{y} \triangleq \{y_0, y_1, \dots, y_{(L-1)}\} \quad (6.3)$$

and

$$\mathbf{z} \triangleq \{z_0, z_1, \dots, z_{(N-1)}\} \quad (6.4)$$

where

$$y_k \triangleq |r_{2k} r_{2k+1}| \quad (6.5)$$

and

$$z_k \triangleq |r_{(2k+2\lfloor K/4 \rfloor + 1)} r_{(2k+2\lfloor K/4 \rfloor + 2)}| \quad (6.6)$$

with $L = \lfloor K/4 \rfloor$, $N = K/2 - \lfloor K/4 \rfloor - 1$, and $|\cdot|$ and $\lfloor \cdot \rfloor$ are the absolute value operator and the floor function, respectively.

When SM is transmitted, the components of both \mathbf{y} and \mathbf{z} are independent and identically distributed (i.i.d.), being drawn from the same distribution, as \mathbf{r} is an i.i.d. sequence. On the other hand, for AL-STBC, the components of \mathbf{y} and \mathbf{z} are not

necessary i.i.d. since \mathbf{r} is not an i.i.d. sequence. In fact, with the assumption that the alignment of the STBC block is not known at the blind receiver, two different events can be considered for AL-STBC: (i) event AL_1 , for which the first received symbol does not correspond to the beginning of the AL-STBC block (see Fig. 6.1) and (ii) event AL_2 , for which it does. Given (6.3)-(6.6), for the event AL_1 , y_k , $k = 0, 1, \dots, L - 1$, represents the multiplication of consecutive received symbols from different blocks, whereas $z(k)$, $k = 0, 1, \dots, N - 1$, is the multiplication of successive received symbols from the same blocks. Accordingly, \mathbf{y} is not an i.i.d. sequence, whereas \mathbf{z} is an i.i.d. sequence for the event AL_1 . By following the same reasoning, \mathbf{y} is an i.i.d. sequence and \mathbf{z} is not for the event AL_2 . Thus, it is straightforward that detecting the presence of the SM transmission becomes equivalent to testing the null hypothesis,

$$H_0 : \begin{cases} \text{The components of both } \mathbf{y} \text{ and } \mathbf{z} \text{ are i.i.d.,} \\ \text{being drawn from the same distribution} \end{cases} \quad (6.7)$$

against the alternative hypothesis that the components of \mathbf{y} or \mathbf{z} are not i.i.d., and therefore, they are not drawn from the same distribution.

Let $\hat{F}_y(x)$ and $\hat{F}_z(x)$ denote the empirical cumulative distribution function (ECDF)s of the sequences \mathbf{y} and \mathbf{z} , which are defined as

$$\hat{F}_y(x) \triangleq \frac{1}{L} \sum_{n=0}^{L-1} \mathbb{I}\{y_n \leq x\} \quad (6.8)$$

and

$$\hat{F}_z(x) \triangleq \frac{1}{N} \sum_{n=0}^{N-1} \mathbb{I}\{z_n \leq x\}, \quad (6.9)$$

where $\mathbb{I}\{\cdot\}$ is the indicator function, which is equal to one if the input is true, and equal to zero otherwise. Under H_0 , it follows from the strong law of large numbers

that for each x , $\hat{F}_y(x)$ and $\hat{F}_z(x)$ converge to the same CDF with probability one as $K \rightarrow \infty$; in other words, if the null hypothesis is valid, $\hat{F}_y(x)$ will be fairly close to $\hat{F}_z(x)$ when K is large enough. Thus, it is reasonable to reject the null hypothesis H_0 in favor of its alternative hypothesis, i.e., the presence of AL-STBC, if a significant distance is observed between $\hat{F}_y(x)$ and $\hat{F}_z(x)$. Hence, the distance between $\hat{F}_y(x)$ and $\hat{F}_z(x)$ can be used as a feature for the identification of SM and AL-STBC. To measure the distance between two ECDFs, we employ the two-sample K-S test [216]. With this test, the distance is defined as

$$\xi \triangleq \max_v |\hat{F}_z(v) - \hat{F}_y(v)|, \quad (6.10)$$

where $v \in \{y(0), \dots, y(L-1), z(0), \dots, z(N-1)\}$ [216]. By comparing the goodness-of-fit statistic with a threshold, λ , the hypothesis H_0 is rejected if

$$\xi \geq \lambda. \quad (6.11)$$

Moreover, ξ is distribution-free under H_0 and the following result holds for all continuous distributions [216, 217]

$$\mathbb{P}\{\xi < \lambda | H_0\} = \beta, \quad (6.12)$$

where β is the confidence interval, represented as

$$\beta = 1 - \Phi \left(\lambda \left[\sqrt{\frac{NL}{N+L}} + 0.12 + \frac{0.11}{\sqrt{\frac{NL}{N+L}}} \right] \right), \quad (6.13)$$

where

$$\Phi(x) = 2 \sum_{m=1}^{\infty} (-1)^{m-1} e^{-2m^2 x^2}. \quad (6.14)$$

As such, for a certain β , the threshold, λ , is specified by (6.13). The proposed K-S-based identification algorithm is formally described as following.

Algorithm 12 : K-S Based STBC Identification Algorithm

- 1: **Input** the target confidence interval β
 - 2: **Output** Type of the transmitted STBC code
 - 3: Acquire the measurement $\mathbf{r} = [r(0), r(1), \dots, r(K-1)]$.
 - 4: Obtain the sequences \mathbf{y} and \mathbf{z} from \mathbf{r} using (6.5)-(6.6).
 - 5: Obtain the corresponding threshold λ using (6.13).
 - 6: Compute the ECDFs \hat{F}_y and \hat{F}_z using (6.8) and (6.9).
 - 7: Compute the maximum distance between the \hat{F}_y and \hat{F}_z ξ according to (6.10).
 - 8: **If** $\xi < \lambda$
 - 9: SM is present
 - 10: **else**
 - 11: AL-STBC is present
 - 12: **end**
-

6.3 Extension to a Larger Pool of STBCs

By employing a binary decision tree with $I - 1$ nodes, where I represents the cardinality of the STBC pool, the STBC codes can be distinguished via the two-sample K-S test at each node. At node i , two independent sequences $\mathbf{y}_i = \{y_0^i, y_1^i, \dots, y_{L_i-1}^i\}$ and $\mathbf{z}_i = \{z_0^i, z_1^i, \dots, z_{N_i-1}^i\}$ with lengths L_i and N_i are derived from the received signal with length K , such that the presence of the code C_i becomes equivalent to testing the null hypothesis,

$$H_0^i : \begin{cases} \text{The components of both } \mathbf{y}_i \text{ and } \mathbf{z}_i \text{ are i.i.d.,} \\ \text{being drawn from the same distribution} \end{cases} \quad (6.15)$$

against the alternative hypothesis, H_1^i , i.e., the presence of STBCs other than C_i , in which the components of \mathbf{y}_i or \mathbf{z}_i are not i.i.d., and thus, they are not drawn from the same distribution.

We denote $\hat{F}_{\mathbf{y}^i}(x)$ and $\hat{F}_{\mathbf{z}^i}(x)$ as the ECDFs of the sequences \mathbf{y}_i and \mathbf{z}_i at the i^{th} node, which are defined as

$$\hat{F}_{\mathbf{y}^i}(x) \triangleq \frac{1}{L_i} \sum_{n=0}^{L_i-1} \mathbb{I}\{y_n^i \leq x\}, \quad (6.16)$$

$$\hat{F}_{\mathbf{z}^i}(x) \triangleq \frac{1}{N_i} \sum_{n=0}^{N_i-1} \mathbb{I}\{z_n^i \leq x\}, \quad (6.17)$$

Under hypothesis H_0^i , $\hat{F}_{\mathbf{y}^i}(x)$ converges to $\hat{F}_{\mathbf{z}^i}(x)$ if L_i and N_i are large enough; accordingly, H_0^i can be rejected in favor of its alternative hypothesis if a large enough distance exists between $\hat{F}_{\mathbf{y}^i}(x)$ and $\hat{F}_{\mathbf{z}^i}(x)$. This property is used as a feature at the binary tree for the STBC identification at the i^{th} node. The two-sample K-S test is employed to measure the distance; the largest absolute difference between the two ECDFs is used, which is defined as

$$\xi^i \triangleq \max_{v_i} \left| \hat{F}_{\mathbf{z}^i}(v_i) - \hat{F}_{\mathbf{y}^i}(v_i) \right|, \quad (6.18)$$

where $v_i = \{y_0^i, y_2^i, \dots, y_{L_i-1}^i, z_0^i, z_2^i, \dots, z_{N_i-1}^i\}$. At the i^{th} node, the hypothesis H_0^i is decided on if

$$\xi^i < \lambda_i, \quad (6.19)$$

where λ_i is a threshold. It should be noted that ξ^i does not depend on the distribution of \mathbf{y}_i and \mathbf{z}_i under H_0^i and the following result is valid for all continuous distributions

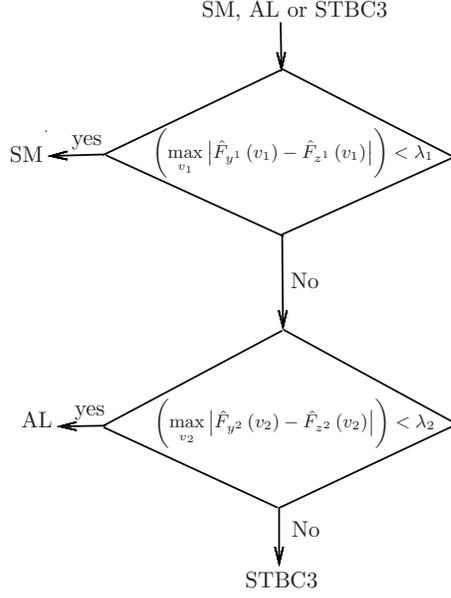


Fig. 6.2: Flowchart of the proposed binary decision tree for the identification of SM, AL, and STBC3.

regardless of their type [216]

$$\mathbb{P}\{\xi^i < \lambda_i | H_0^i\} = \beta_i, \quad (6.20)$$

where β_i is the confidence interval at the i^{th} node, given by

$$\beta_i = 1 - \Phi \left(\lambda_i \left[\sqrt{\frac{N_i L_i}{N_i + L_i}} + 0.12 + \frac{0.11}{\sqrt{\frac{N_i L_i}{N_i + L_i}}} \right] \right), \quad (6.21)$$

where $\Phi(x)$ is given in (6.14).

Fig. 6.2 shows the flowchart of the decision tree-based identification method for SM, AL-STBC, and STBC3. At the first node, SM is distinguished from AL and STBC3 by employing the sequences \mathbf{y}_1 and \mathbf{z}_1 given as

$$\mathbf{y}_1 = \{y_0^1, y_1^1, \dots, y_{L_1-1}^1\} \quad (6.22)$$

and

$$\mathbf{z}_1 = \{z_0^1, z_1^1, \dots, z_{N_1-1}^1\}, \quad (6.23)$$

where $y_k^1 \triangleq |r_{2k}r_{2k+1}|$ and $z_k^1 \triangleq |r'_{2k}r'_{2k+1}|$, with $L_1 = \lfloor K/4 \rfloor$, $N_1 = K/2 - \lfloor K/4 \rfloor - 1$, and $r'_k = r_{k+2\lfloor K/4 \rfloor + 1}$. At the first node, SM is declared present if the statistic defined in (6.18) for the \mathbf{y}_1 and \mathbf{z}_1 sequences does not exceed the threshold λ_1 corresponding to the preset confidence interval β_1 . Otherwise, either AL or STBC3 can be present. Note that in both cases, the components of \mathbf{y}_1 and \mathbf{z}_1 are not necessary i.i.d., and are not drawn from the same distribution.

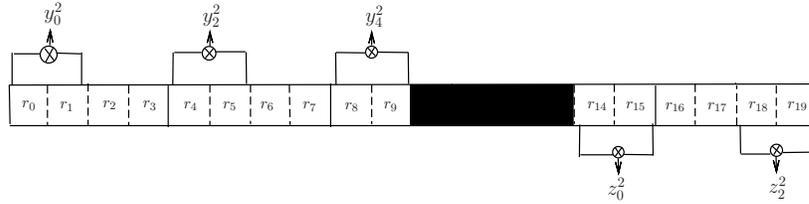
At the second node, AL is individualized from STBC3 by using two new sequences \mathbf{y}_2 and \mathbf{z}_2 , also obtained from the received sequence. The received sequence is split into two non-overlapping sequences $r'_1 = \{r_0, r_1, \dots, r_{M-1}\}$ and $r'_2 = \{r_{M+4}, r_{M+5}, \dots, r_{K-1}\}$, where $M = 4 \lfloor (K+4)/8 \rfloor - 2$, and the two new sequences \mathbf{y}_2 and \mathbf{z}_2 are defined as

$$\mathbf{y}_2 = \{y_0^2, y_1^2, \dots, y_{L_2-1}^2\}, \quad (6.24)$$

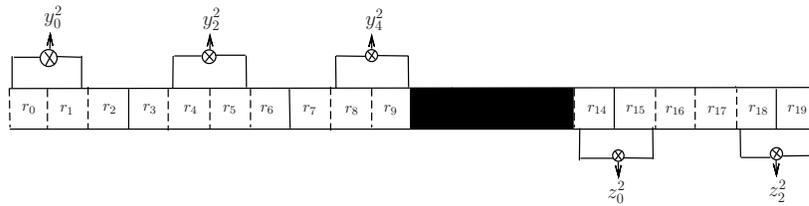
and

$$\mathbf{z}_2 = \{z_0^2, z_1^2, \dots, z_{N_2-1}^2\}, \quad (6.25)$$

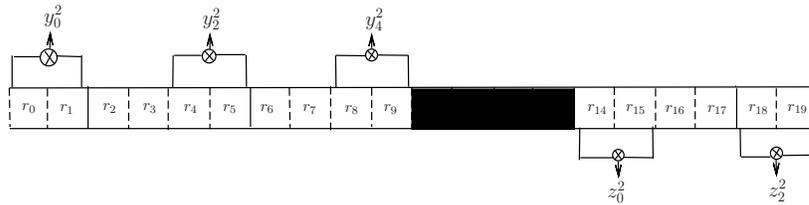
where $y_k^2 \triangleq |r_{4k}r_{4k+1}|$, $z_k^2 \triangleq |r_{4k+M+4}r_{4k+M+5}|$, $L_2 = \lfloor \frac{M+2}{4} \rfloor$, and $N_2 = \lfloor \frac{K-M-2}{4} \rfloor$. Under the assumption that the beginning of the STBC3 block is not known at the receiver, four different events can be considered for STBC3, corresponding to the four possible starting symbols. For each of these events, one can easily see that the components of \mathbf{y}_2 and \mathbf{z}_2 are drawn from different distributions (Fig. 6.3). On the other hand, for AL, the components of both \mathbf{y}_2 and \mathbf{z}_2 are i.i.d., being drawn from the same distribution (Fig. 6.4); this can be easily explained, as the components of



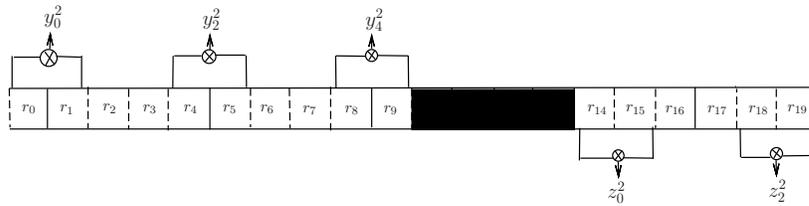
(a) Event 1



(b) Event 2

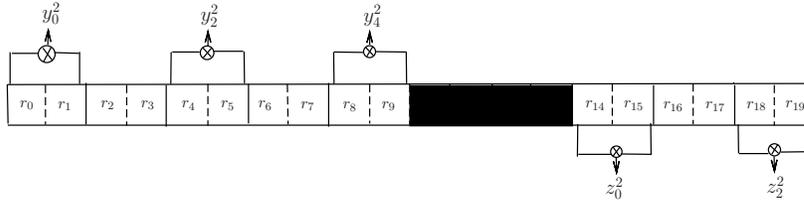


(c) Event 3

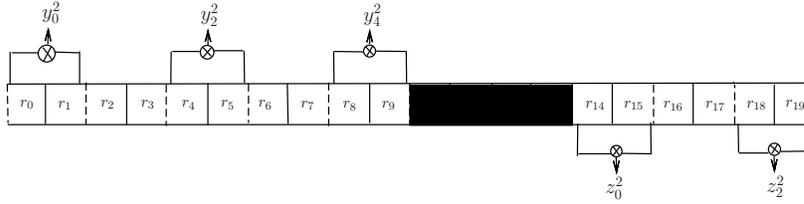


(d) Event 4

Fig. 6.3: Illustration of the sequences \mathbf{y}_2 and \mathbf{z}_2 for the four different STBC3 events, corresponding to the four possible starting symbols which belong to an STBC block, for $K = 20$. Solid lines are used to delimitate symbols which do not belong to the same block and dashed lines are employed to delimitate symbols which belong to the same block.



(a) Event 1



(b) Event 2

Fig. 6.4: Illustration of the sequences \mathbf{y}_2 and \mathbf{z}_2 for the two different AL events, corresponding to the two possible starting symbols which belong to an AL block, for $K = 20$. Solid lines are used to delimitate symbols which do not belong to the same block and dashed lines are employed to delimitate symbols which belong to the same block.

both \mathbf{y}_2 and \mathbf{z}_2 result either from the multiplication of received symbols belonging to the same AL block or to different AL blocks. Hence, detecting the presence of the AL transmission becomes equivalent to testing the null hypothesis in 6.15 for the \mathbf{y}_2 and \mathbf{z}_2 sequences. As such, AL is declared present if the statistic defined in 6.18 for the \mathbf{y}_2 and \mathbf{z}_2 sequences does not exceed the threshold λ_2 corresponding to β_2 ; otherwise, STBC3 is declared present.

6.4 Simulations

In this section, the performance of the proposed STBC identification algorithm is evaluated through several simulation experiments.

6.4.1 Simulation Setup

The performance of the proposed K-S-based identification algorithm was examined using 1000 Monte Carlo trials. Unless otherwise mentioned, for the identification of the SM and AL-STBC, the modulation was QPSK, $K = 2048$, $\beta = 0.99$, the CFO normalized to data rate was $\Delta f = 0.01$, Nakagami- m fading channel with $m = 2$ was considered¹, and $E\{|h_i^2|\} = 1$, $i = 0, 1$. In addition, for the identification of the SM, AL-STBC, and STBC3, we considered binary phase-shift keying (BPSK) modulation, $K = 8092$ observed symbols, and $\beta_1 = \beta_2 = 0.99$. The additive noise, w_k in 6.2, was modeled as impulsive noise due to a variety of natural and man-made sources in most radio channels, being characterized by the probability density function $f_w(x) = (1 - \varepsilon)f_N(x) + \varepsilon f_I(x)$, where $0 \leq \varepsilon < 1$ is the mixing parameter, and $f_N(x)$ and $f_I(x)$ are zero-mean complex Gaussian distributions with variance σ_N^2 and σ_I^2 , respectively [127]. We set $\varepsilon = 0.01$ and $\sigma_I^2/\sigma_N^2 = 100$. Under the assumption of unit variance constellations, regardless of the number of transmit antennas, the signal-to-noise ratio (SNR) is defined as $10 \log(P_s/\sigma_w^2)$, with $P_s = (1/p)\mathbb{E}\{\text{tr}[\mathbf{C}(\mathbf{s}_b)\mathbf{C}^H(\mathbf{s}_b)]\}$, where tr denotes the trace of a matrix [7], $n_t = 2$, and $\sigma_w^2 = (1 - \varepsilon)\sigma_N^2 + \varepsilon\sigma_I^2$. The average probability of correct identification $P_c = 2^{-1} \sum_{\theta \in \{\text{SM, AL}\}} \mathbb{P}\{\hat{\theta} = \theta | \theta\}$, $P_c = 3^{-1} \sum_{\theta \in \{\text{SM, AL, STBC3}\}} \mathbb{P}\{\hat{\theta} = \theta | \theta\}$ and the probability of correct identification $\mathbb{P}\{\hat{\theta} = \theta | \theta\}$, $\theta \in \{\text{SM, AL, STBC3}\}$ were used to evaluate the identification performance, with $\hat{\theta}$ as the identified signal type.

6.4.2 Simulation Results

Fig. 6.5 shows the probability of correct identification for SM and AL-STBC over Nakagami- m fading channel for different values of m . As expected, the probability

¹Note that results for the FOLP-C, versus which we compare our algorithm, are given for $m = 2$ in [8]; FOLP-C does not provide a good performance for $m = 1$.

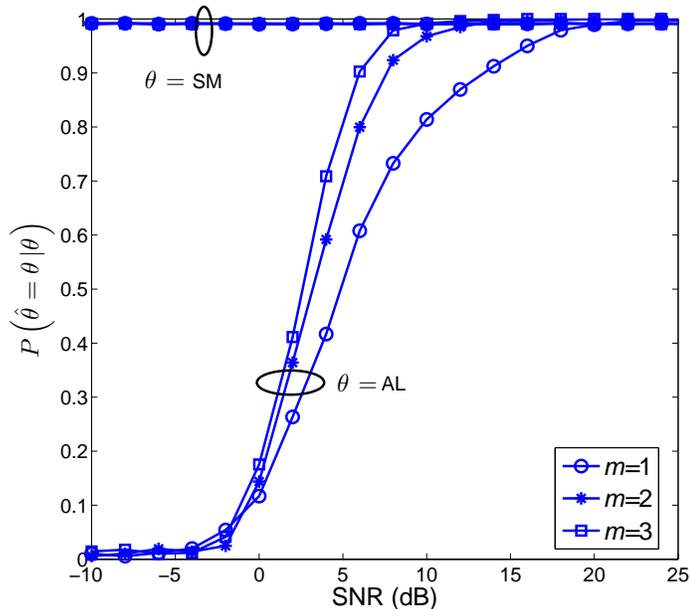


Fig. 6.5: The probability of correct identification, $\mathbb{P}\{\hat{\theta} = \theta | \theta\}$, $\theta \in \{\text{SM}, \text{AL}\}$, versus SNR for the K-S-based algorithm over different Nakagami- m fading channels in the presence of impulsive noise, with QPSK modulation, $K = 2048$, $\Delta f = 0.01$, and $\beta = 0.99$.

of correct identification for SM, $\mathbb{P}\{\hat{\theta} = \text{SM} | \text{SM}\}$, is equal to the target confidence interval $\beta = 0.99$ and is independent of the SNR and m . On the other hand, the probability of correct identification for AL-STBC, $\mathbb{P}\{\hat{\theta} = \text{AL} | \text{AL}\}$, improves as the SNR and m increase. This can be easily explained, as the noise and channel coefficients control the maximum distance between the ECDFs, $\hat{F}_y(x)$ and $\hat{F}_z(x)$; a reduced noise level and improved channel conditions lead to a larger distance, which in turn yields a better performance.

The probability of correct identification $\mathbb{P}\{\hat{\theta} = \theta | \theta\}$, $\theta \in \{\text{SM}, \text{AL}\}$, for different values of the confidence interval, β , is depicted in Fig. 6.6. As β increases, $\mathbb{P}\{\hat{\theta} = \text{SM} | \text{SM}\}$ also increases (as equals β), while $\mathbb{P}\{\hat{\theta} = \text{AL} | \text{AL}\}$ decreases. The latter can be easily explained, as according to (6.13), for fixed L and N values, a higher β

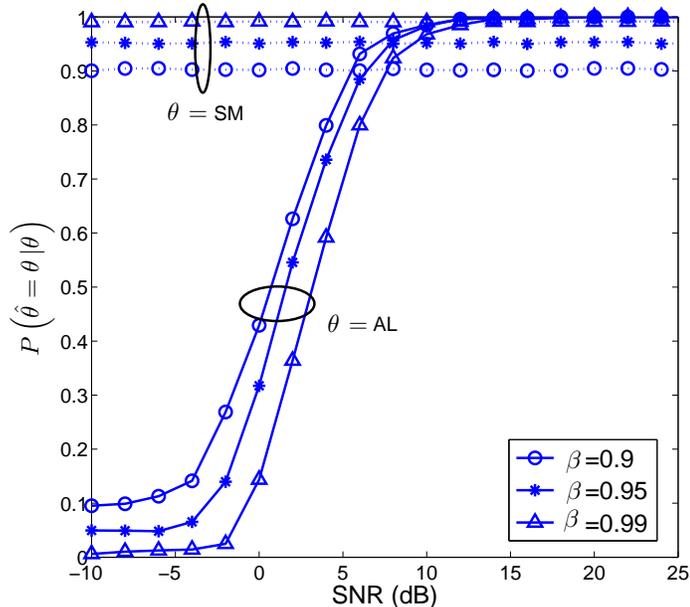


Fig. 6.6: The effect of the confidence interval, β , on the probability of correct identification, $\mathbb{P}\{\hat{\theta} = \theta | \theta\}$, $\theta \in \{\text{SM}, \text{AL}\}$, for the K-S-based algorithm over Nakagami- m fading channel, $m = 2$, in the presence of impulsive noise, with QPSK modulation, $K = 2048$, and $\Delta f = 0.01$.

Table 6.1: The effect of the SNR on the P_c .

SNR (dB)	5	10	15	20
K-S	0.864	0.979	0.994	0.999
FOLP-C	0.588	0.871	0.943	0.977

value yields an increase in the threshold λ .

In Fig. 6.7, the average probability of correct identification, P_c , is presented over Δf for the ML [7], FOLP-C² [8], and K-S-based algorithms. As can be seen, the proposed K-S-based algorithm is basically insensitive to the CFO; this is because of the absolute value operator in (6.5) and (6.6), which eliminates the CFO effect. Additionally, it outperforms both FOLP-C and ML, and the latter drastically fails as Δf increases.

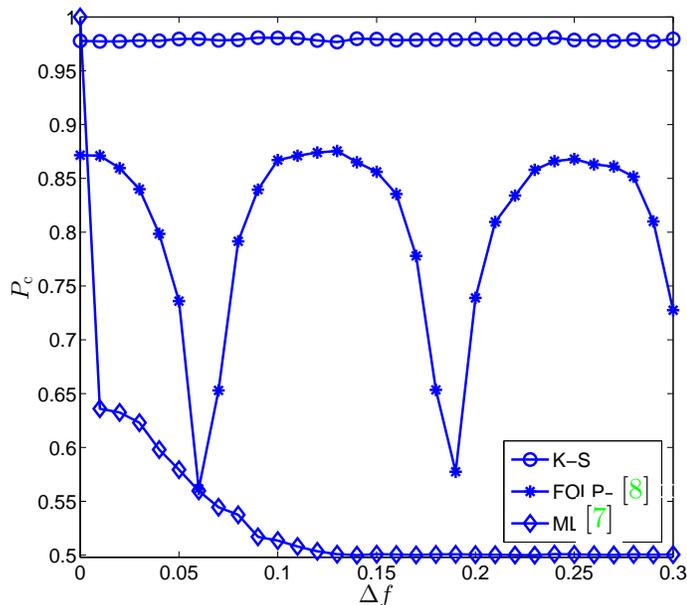


Fig. 6.7: The effect of the normalized CFO, Δf , on the average probability of correct identification, P_c , for the ML [7], FOLP-C [8], and K-S-based algorithms over Nakagami- m fading channel, $m = 2$, in the presence of impulsive noise, with QPSK modulation, $K = 2048$, $\beta = 0.99$, and at SNR=10 dB.

Table 6.2: The effect of the number symbols on the P_c .

K	512	1024	2048	4096
K-S	0.834	0.965	0.979	0.986
FOLP-C	0.516	0.790	0.871	0.911

Table 6.1 compares³ the average probability of correct identification, P_c , for the FOLP-C and K-S-based algorithms over Nakagami- m fading channel in the presence of impulsive noise and CFO for different values of SNR. As it is observed, the K-S-based algorithm exhibits a superior performance, especially at lower SNR, which is of interest.

In Table 6.2, the effect of the number of received symbols, K , on P_c is presented for the FOLP-C and K-S-based algorithms at SNR=10 dB. It can be observed that

²Note that FOLP-C was selected for comparison, as it outperforms all other FB algorithms proposed in the literature under transmission impairments [8].

³Henceforth, results for ML are omitted as its performance is not acceptable ($0.5 < P_c < 0.65$).

Table 6.3: The effect of the modulation type on the P_c .

Mod	BPSK	QPSK	8-PSK	16-PSK	16-QAM	32-QAM
K-S	0.883	0.979	0.978	0.9764	0.736	0.742
FOLP-C	0.502	0.871	0.855	0.8401	0.721	0.711

Table 6.4: The Effect of the TO on the P_c .

$\Delta T/T$	0	0.125	0.25	0.375
K-S	0.974	0.964	0.824	0.552
FOLP-C	0.871	0.802	0.691	0.505

increasing the number of symbols results in a performance improvement for the proposed algorithm. The reason is that as L and N increase, the corresponding threshold, λ , decreases for a given β in (6.13), yielding an increase in $\mathbb{P}\{\hat{\theta} = \text{AL} | \text{AL}\}$, while $\mathbb{P}\{\hat{\theta} = \text{SM} | \text{SM}\}$ is equal to β . It is worth noting that the K-S-based algorithm outperforms FOLP-C.

Table 6.3 illustrates the effect of the modulation type on P_c for the FOLP-C and K-S-based algorithms at SNR=10 dB. As it is observed, the proposed algorithm is independent of M for the M -PSK modulations ($M \geq 4$), and it degrades as M increases for the M -QAM modulations. It is worth noting that while $\mathbb{P}\{\hat{\theta} = \text{SM} | \text{SM}\}$ is independent of the modulation type (as equals β), $\mathbb{P}\{\hat{\theta} = \text{AL} | \text{AL}\}$ has the same behavior as P_c . Additionally, it should be mentioned that larger K is needed to attain an identification performance for M -QAM similar to the one for M -PSK modulations ($M \geq 4$). For instance, with $K=8192$, $P_c = 0.972$ and 0.963 for the proposed algorithm, while $P_c = 0.861$ and 0.852 for FOLP-C when using 16-QAM and 32-QAM, respectively. The proposed algorithm outperforms FOLP-C; furthermore, it can identify SM and AL-STBC with the binary phase-shift-keying (BPSK)⁴ modulation, opposite to FOLP-C.

⁴Note that there is a small amount of performance degradation for BPSK when compared with M -PSK ($M \geq 4$), as the difference between the two CDFs is reduced for AL-STBC due to the lack of the imaginary part in the BPSK constellation.

Table 6.4 illustrates the performance of the FOLP-C and K-S-based algorithms in the presence of TO for the case of root-raised-cosine pulse shaping with a roll-off factor of 0.2 at the transmit- and receive-sides. The receive filter output is sampled at times $T + k\Delta T$, where T is the symbol period and ΔT is the deviation from the ideal sampling time. As can be seen, the proposed algorithm exhibits a lower sensitivity to TO when compared with FOLP-C, being relatively robust for $\Delta T/T \leq 0.25$.

The good identification performance of the proposed K-S-based algorithm does not come at an additional computational cost; both the K-S-based and FOLP-C algorithms have the complexity order of $\mathcal{O}(K \log K)$.

Fig. 6.8 shows the average probability of correct identification for $\theta \in \{\text{SM}, \text{AL}, \text{STBC3}\}$ versus SNR, for different Nakagami- m fading channels, $m = 1, 2$ and ∞ . As can be seen, P_c improves as the SNR and m increase. This is because the maximum distance between the ECDFs at each node increases as the noise level decreases or as the channel conditions improve.

In Fig. 6.9, the effect of the CFO normalized to the data rate, Δf , on P_c is illustrated. As can be seen, the performance of the proposed method does not depend on Δf . This can be easily explained, as the CFO effect is eliminated through the absolute value operator in the definition of the derived sequences at each node. Thus, the proposed method is completely robust to Δf .

Fig. 6.10 compares the average probability of correct identification, P_c , achieved by the proposed method, fourth-order lag product (FOLP) method [8], and the second-order cyclostationarity-based (CS) method [9]. As can be seen, the proposed method outperforms both the FOLP and the second-order cyclostationarity-based methods.

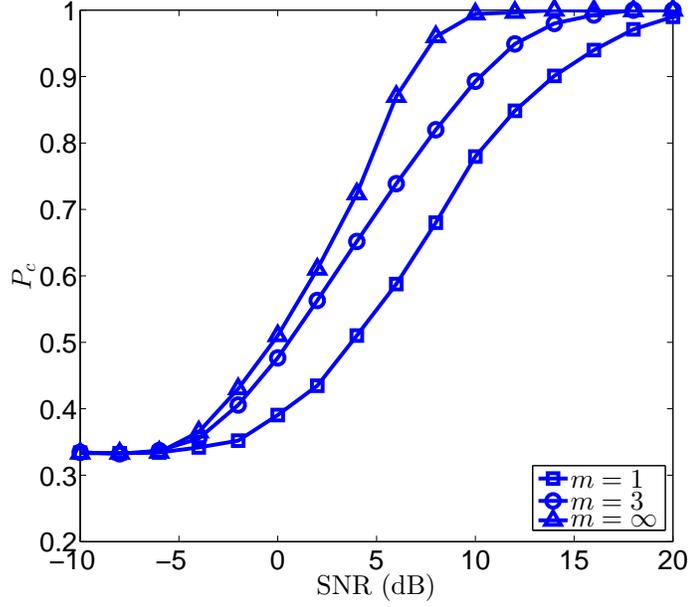


Fig. 6.8: Average probability of correct identification, P_c , $\theta \in \{\text{SM}, \text{AL}, \text{STBC3}\}$, versus SNR for diverse Nakagami- m fading channels.

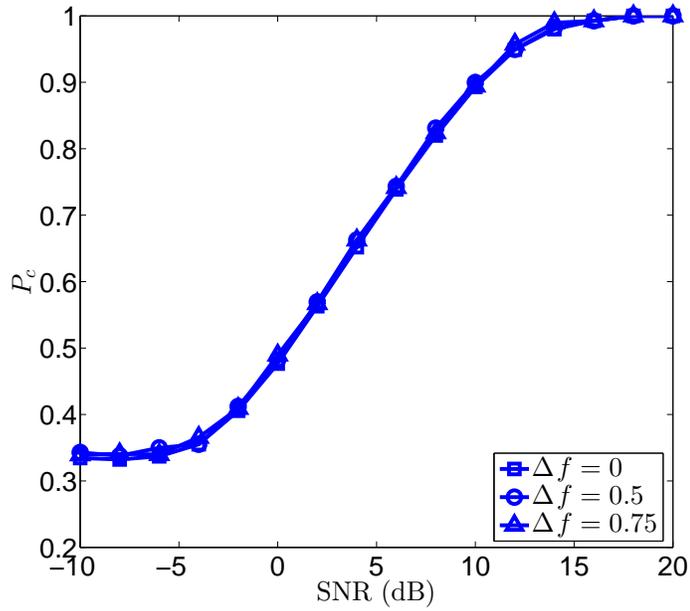


Fig. 6.9: Effect of the CFO normalized to the data rate, Δf , on P_c for $\theta \in \{\text{SM}, \text{AL}, \text{STBC3}\}$ in Nakagami- m fading channel, $m = 3$.

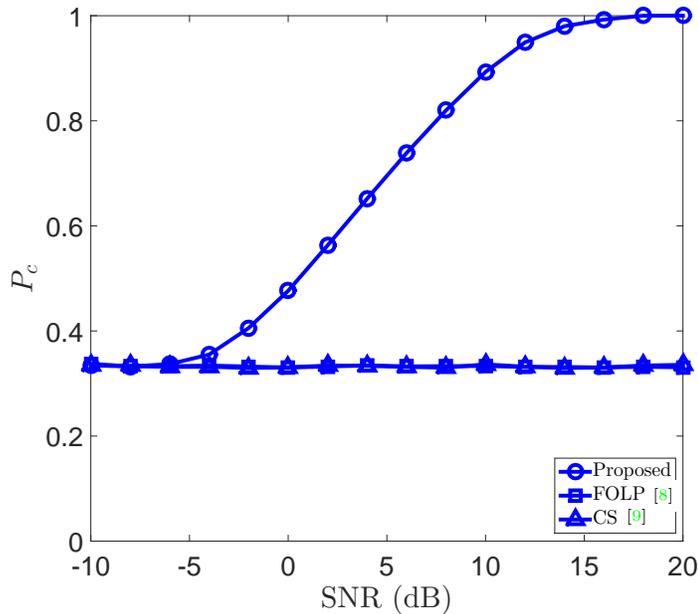


Fig. 6.10: Performance comparison of the proposed method, FOLP [8], and second-order cyclostationarity-based [9] for $\theta \in \{\text{SM}, \text{AL}, \text{STBC3}\}$ in Nakagami- m fading channel, $m = 3$ and $\Delta f = 0$.

6.5 Conclusions and Directions for Future Research

In Chapter 6 of this thesis, STBC identification was studied. In this section, a summary of the main results in Chapter 6 is provided, and possible directions for future research are pointed out .

6.5.1 Summary

In Chapter 6, a new algorithm was proposed for blind identification of STBCs. A single receive antenna configuration was considered, with unknown channel coefficients, noise parameters, and modulation type. The identification problem was formulated as a goodness-of-fit test and the K-S test was employed for decision making. The obtained results showed that the proposed K-S-based identification algorithm offers a superior performance when compared to related works in the literature, particularly in the

presence of CFO, TO, and non-Gaussian noise. By employing a decision binary tree, the proposed algorithm was extended for the identification of a larger STBC candidate pool.

6.5.2 Future research

The results in Chapter 6 of this thesis open interesting directions for a number of future research topics; two of them are as follows:

- The first one is an extension of the proposed blind STBC identification algorithm for single receive antenna to multiple receive antennas. In this case, the spatial diversity due to the multiple receive antennas would result in performance improvement;
- The second one pertains to developing a joint antenna enumeration and STBC identification algorithm.

Chapter 7

Overall Conclusions

In this thesis, five signal processing problems in conjunction with parameter estimation and signal identification in modern wireless communication systems were studied.

In Chapter 2, a new uplink MA scheme for IoT was proposed. The proposed MA scheme develops a new alternative mechanism to the MAC address through parametric parameter estimation and signal identification in order to identify the number and identity of the active IoT devices in the gateway. Furthermore, a new non-linear MUD algorithm, which does not require FCCs and CP estimation was designed.

In Chapter 3, MDS estimation in MIMO frequency-selective fading channel was studied, and different DA and NDA estimators were proposed. Both DA- and NDA-MLEs for MDS in MIMO frequency-selective channel were derived. Furthermore, a new low-complexity NDA-MBE for MDS was proposed. The proposed NDA-MBE does not require joint parameter estimation, such as CFO, signal power, noise power, and channel delay profile estimation.

In Chapter 4, the problem of antenna enumeration in time-varying fading channel was investigated and the semi-blind and blind antenna enumeration algorithms were proposed. Moreover, a closed-form expression for the probability of correct antenna

detection of the blind algorithm was derived. On the contrary to the existing antenna enumeration algorithms, the proposed algorithms can detect a large number of transmit antennas with a single receive antenna. This capability is obtained due to the variation of the fading channel. Simulation results show that the proposed antenna enumeration algorithms outperform the AIC and MDL algorithms in time-varying fading channel and at low SNR values.

In Chapter 5, the problem of SNR and noise variance estimation in MIMO time-varying fading channel was studied, and M_2M_4 and $M_2M_4M_6$ SNR estimators were proposed through statistical MB approach. The proposed MBEs increase the system capacity due to lack of pilots and preambles. Moreover, they are categorized as low-complexity SNR estimators.

In Chapter 6, the problem of STBC identification in Rayleigh fading and non-Gaussian noise was investigated. First, a new algorithm based on the K-S test was proposed for the identification of SM versus AL-STBC in the presence of channel impairments, such as CFO and TO. Then, the proposed algorithm was extended for the identification of a larger STBC candidate pool through a decision binary tree.

Bibliography

- [1] F. Bellili and S. Affes, “A low-cost and robust maximum likelihood Doppler spread estimator,” in *Proc. GLOBECOM*, Atlanta, USA, Dec. 2013, pp. 4325–4330.
- [2] F. Bellili, Y. Selmi, S. Affes, and A. Ghayeb, “A low-cost and robust maximum likelihood joint estimator for the doppler spread and CFO parameters over flat-fading rayleigh channels,” *IEEE Trans. Commun.*, pp. 1–1, 2017.
- [3] M. Souden, S. Affes, J. Benesty, and R. Bahroun, “Robust Doppler spread estimation in the presence of a residual carrier frequency offset,” *IEEE Trans. Signal Process.*, vol. 57, no. 10, pp. 4148–4153, 2009.
- [4] H. Zhang and A. Abdi, “Cyclostationarity-based Doppler spread estimation in mobile fading channels,” *IEEE Trans. Commun.*, vol. 57, no. 4, pp. 1061–1067, Apr. 2009.
- [5] A. Mariani, A. Giorgetti, and M. Chiani, “Effects of noise power estimation on energy detection for cognitive radio applications,” *IEEE Trans. Wireless Commun.*, vol. 59, no. 12, pp. 3410–3420, Dec. 2011.
- [6] O. Somekh, O. Simeone, Y. Bar-Ness, and W. Su, “Detecting the number of transmit antennas with unauthorized or cognitive receivers in MIMO systems,” in *Proc. IEEE MILCOM*, 2007, pp. 478–482.

- [7] V. Choqueuse, M. Marazin, L. Collin, K. C. Yao, and G. Burel, “Blind recognition of linear space–time block codes: a likelihood-based approach,” *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1290–1299, 2010.
- [8] Y. Eldemerdash, M. Marey, O. Dobre, G. K. Karagiannidis, R. Inkol *et al.*, “Fourth-order statistics for blind classification of spatial multiplexing and alam-outi space-time block code signals,” *IEEE Trans. Commun.*, vol. 61, no. 6, pp. 2420–2431, 2013.
- [9] M. Marey, O. Dobre, R. Inkol *et al.*, “Classification of space-time block codes based on second-order cyclostationarity with transmission impairments,” *IEEE Trans. Commun.*, vol. 11, no. 7, pp. 2574–2584, 2012.
- [10] H. W. Sorenson, *Parameter estimation: principles and problems*. M. Dekker, 1980, vol. 9.
- [11] O. Dobre, “Signal identification for emerging intelligent radios: classical problems and new challenges,” *IEEE Trans. Instrum. Meas.*, vol. 18, no. 2, pp. 11–18, 2015.
- [12] J. V. Beck and K. J. Arnold, *Parameter estimation in engineering and science*. James Beck, 1977.
- [13] Z. Lu and A. M. Zoubir, “Generalized bayesian information criterion for source enumeration in array processing,” *IEEE Trans. Signal Process.*, vol. 61, no. 6, pp. 1470–1480, 2013.
- [14] H. Saarnisaari, “Estimation of numbers of multipaths in DS-CDMA by the MDL principle,” in *Spread Spectrum Techniques and Applications, 1998. Proceedings., 1998 IEEE 5th International Symposium on*, vol. 1. IEEE, 1998, pp. 258–261.

- [15] Z. Lu and A. M. Zoubir, “Flexible detection criterion for source enumeration in array processing,” *IEEE Trans. Signal Process.*, vol. 61, no. 6, pp. 1303–1314, 2013.
- [16] O. A. Dobre, A. Abdi, Y. Bar-Ness, and W. Su, “Survey of automatic modulation classification techniques: classical approaches and new trends,” *Communications, IET*, vol. 1, no. 2, pp. 137–156, 2007.
- [17] M. R. Palattella, M. Dohler, A. Grieco, G. Rizzo, J. Torsner, T. Engel, and L. Ladid, “Internet of things in the 5G era: Enablers, architecture, and business models,” *IEEE J. Sel. Areas Commun.*, vol. 34, no. 3, pp. 510–527, Mar. 2016.
- [18] H. Arslan, *Cognitive radio, software defined radio, and adaptive wireless systems*. Springer, 2007.
- [19] J. D. Gibson, *Mobile communications handbook*. CRC press, 2012.
- [20] A. F. Molisch, *Wireless Communications*. John Wiley & Sons, 2007.
- [21] G. L. Stüber, *Principles of Mobile Communication*. Springer Science & Business Media, 2011.
- [22] M. D. Austin and G. L. Stüber, “Velocity adaptive handoff algorithms for microcellular systems,” *IEEE Trans. Veh. Technol.*, vol. 43, no. 3, pp. 549–561, Aug. 1994.
- [23] B. L. Mark and A. E. Leu, “Local averaging for fast handoffs in cellular networks,” *IEEE Trans. Wireless Commun.*, vol. 6, no. 3, pp. 866–874, Mar. 2007.
- [24] A. B. MacKenzie, J. H. Reed, P. Athanas, C. W. Bostian, R. M. Buehrer, L. DaSilva, S. W. Ellingson, Y. T. Hou, M. Hsiao, J.-M. Park *et al.*, “Cognitive

- radio and networking research at virginia tech,” *Proc. IEEE*, vol. 97, no. 4, pp. 660–688, 2009.
- [25] B. Wang and K. Liu, “Advances in cognitive radio networks: A survey,” *IEEE Trans. Emerg. Sel. Topics Power Electron.*, vol. 5, no. 1, pp. 5–23, 2011.
- [26] A. Gupta and R. K. Jha, “A survey of 5G network: Architecture and emerging technologies,” *IEEE Access*, vol. 3, pp. 1206–1232, Aug. 2015.
- [27] L. Atzori, A. Iera, and G. Morabito, “The internet of things: A survey,” *Computer Networks*, vol. 54, no. 15, pp. 2787–2805, 2010.
- [28] S. R. Islam, D. Kwak, M. H. Kabir, M. Hossain, and K.-S. Kwak, “The internet of things for health care: a comprehensive survey,” *IEEE Access*, vol. 3, pp. 678–708, June 2015.
- [29] A. Zanella, N. Bui, A. Castellani, L. Vangelista, and M. Zorzi, “Internet of things for smart cities,” *IEEE Internet Things J.*, vol. 1, no. 1, pp. 22–32, Feb. 2014.
- [30] L. Da Xu, W. He, and S. Li, “Internet of things in industries: A survey,” *IEEE Trans. Ind. Informat.*, vol. 10, no. 4, pp. 2233–2243, Nov. 2014.
- [31] M. R. Palattella, N. Accettura, X. Vilajosana, T. Watteyne, L. A. Grieco, G. Boggia, and M. Dohler, “Standardized protocol stack for the internet of (important) things,” *IEEE Commun. Surveys Tuts.*, vol. 15, no. 3, pp. 1389–1406, Third Quarter 2013.
- [32] Z. Sheng, S. Yang, Y. Yu, A. Vasilakos, J. Mccann, and K. Leung, “A survey on the ietf protocol suite for the internet of things: Standards, challenges, and opportunities,” *IEEE Wireless Commun.*, vol. 20, no. 6, pp. 91–98, Dec. 2013.

- [33] C. Goursaud and J.-M. Gorce, “Dedicated networks for IoT: PHY/MAC state of the art and challenges,” *EAI endorsed Transactions on Internet of Things*, Oct. 2015.
- [34] M. Centenaro, L. Vangelista, A. Zanella, and M. Zorzi, “Long-range communications in unlicensed bands: The rising stars in the iot and smart city scenarios,” *IEEE Wireless Commun.*, vol. 23, no. 5, pp. 60–67, Oct. 2016.
- [35] E. Soltanmohammadi, K. Ghavami, and M. Naraghi-Pour, “A survey of traffic issues in machine-to-machine communications over LTE,” *IEEE Internet Things J.*, vol. 3, no. 6, pp. 865–884, Dec. 2016.
- [36] H. S. Dhillon, H. Huang, and H. Viswanathan, “Wide-area wireless communication challenges for the Internet of Things,” *IEEE Commun. Mag.*, vol. 55, no. 2, pp. 168–174, Feb. 2017.
- [37] G. Durisi, T. Koch, and P. Popovski, “Toward massive, ultrareliable, and low-latency wireless communication with short packets,” *Proc. IEEE*, vol. 104, no. 9, pp. 1711–1726, Sep. 2016.
- [38] A. Biral, M. Centenaro, A. Zanella, L. Vangelista, and M. Zorzi, “The challenges of M2M massive access in wireless cellular networks,” *DCN*, vol. 1, no. 1, pp. 1–19, Feb. 2015.
- [39] A. Zanella, M. Zorzi, A. F. dos Santos, P. Popovski, N. Pratas, C. Stefanovic, A. Dekorsy, C. Bockelmann, B. Busropan, and T. A. Norp, “M2M massive wireless access: challenges, research issues, and ways forward,” in *Globecom. IEEE*, 2013, pp. 151–156.
- [40] A. Laya, C. Kalalas, F. Vazquez-Gallego, L. Alonso, and J. Alonso-Zarate, “Goodbye, ALOHA!” *IEEE Access*, vol. 4, pp. 2029–2044, May 2016.

- [41] C. Bockelmann, N. Pratas, H. Nikopour, K. Au, T. Svensson, C. Stefanovic, P. Popovski, and A. Dekorsy, “Massive machine-type communications in 5g: Physical and mac-layer solutions,” *IEEE Commun. Mag.*, vol. 54, no. 9, pp. 59–65, 2016.
- [42] E. MORIN, M. Maman, R. GUIZZETTI, and D. Andrzej, “Comparison of the device lifetime in wireless networks for the Internet of Things,” *IEEE Access*, vol. 5, pp. 7097–7114, June 2017.
- [43] T. Gomes, S. Pinto, F. Salgado, A. Tavares, and J. Cabral, “Building IEEE 802.15. 4 accelerators for heterogeneous wireless sensor nodes,” *IEEE Sensors Lett.*, vol. 1, no. 1, pp. 1–4, Feb. 2017.
- [44] D. De Guglielmo, F. Restuccia, G. Anastasi, M. Conti, and S. K. Das, “Accurate and efficient modeling of 802.15. 4 unslotted CSMA/CA through event chains computation,” *IEEE Trans. Mobile Comput.*, vol. 15, no. 12, pp. 2954–2968, Dec. 2016.
- [45] R. V. P. Yerra, M. P. R. S. Kiran, and R. Pachamuthu, “Reliability and delay analysis of slotted anycast multi-hop wireless networks targeting dense traffic iot applications,” *IEEE Commun. Lett.*, vol. 19, no. 5, pp. 727–730, May 2015.
- [46] J. DeCuir, “Introducing bluetooth smart: Part 1: A look at both classic and new technologies.” *IEEE Consumer Electronics Magazine*, vol. 3, no. 1, pp. 12–18, 2014.
- [47] M. S. Khan, M. S. Islam, and H. Deng, “Design of a reconfigurable rfid sensing tag as a generic sensing platform toward the future internet of things,” *IEEE Internet Things J.*, vol. 1, no. 4, pp. 300–310, Aug. 2014.

- [48] K.-H. Chang, “Bluetooth: a viable solution for IoT?[industry perspectives],” *IEEE Wireless Commun.*, vol. 21, no. 6, pp. 6–7, Dec. 2014.
- [49] T. Adame, A. Bel, B. Bellalta, J. Barcelo, and M. Oliver, “IEEE 802.11 AH: the WiFi approach for M2M communications,” *IEEE Wireless Commun.*, vol. 21, no. 6, pp. 144–152, Dec. 2014.
- [50] C. W. Park, D. Hwang, and T.-J. Lee, “Enhancement of IEEE 802.11 ah MAC for M2M communications,” *IEEE Commun. Lett.*, vol. 18, no. 7, pp. 1151–1154, July 2014.
- [51] L. Zheng, M. Ni, L. Cai, J. Pan, C. Ghosh, and K. Doppler, “Performance analysis of group-synchronized DCF for dense IEEE 802.11 networks,” *IEEE Trans. Wireless Commun.*, vol. 13, no. 11, pp. 6180–6192, Nov. 2014.
- [52] A. Ali, W. Hamouda, and M. Uysal, “Next generation M2M cellular networks: challenges and practical considerations,” *IEEE Communications Magazine*, vol. 53, no. 9, pp. 18–24, Sep. 2015.
- [53] M. Agiwal, A. Roy, and N. Saxena, “Next generation 5G wireless networks: A comprehensive survey,” *IEEE Commun. Surveys Tuts.*, vol. 18, no. 3, pp. 1617–1655, Third Quarte 2016.
- [54] M. S. Ali, E. Hossain, and D. I. Kim, “LTE/LTE-A random access for massive machine-type communications in smart cities,” *IEEE Commun. Mag.*, vol. 55, no. 1, pp. 76–83, 2017.
- [55] T. P. de Andrade, C. A. Astudillo, L. R. Sekijima, and N. L. da Fonseca, “The random access procedure in Long Term Evolution networks for the Internet of Things,” *IEEE Commun. Mag.*, vol. 55, no. 3, pp. 124–131, Mar. 2017.

- [56] Z. Wu, V. D. Park, and J. Li, “Enabling device to device broadcast for LTE cellular networks,” *IEEE J. Sel. Areas Commun.*, vol. 34, no. 1, pp. 58–70, 2016.
- [57] N. Prasad, H. Zhang, H. Zhu, and S. Rangarajan, “Multiuser scheduling in the 3GPP LTE cellular uplink,” *IEEE Trans. Mobile Comput.*, vol. 13, no. 1, pp. 130–145, Jan. 2014.
- [58] Z. Ding, Y. Liu, J. Choi, Q. Sun, M. ElKashlan, I. Chih-Lin, and H. V. Poor, “Application of non-orthogonal multiple access in LTE and 5G networks,” *IEEE Commun. Mag.*, vol. 55, no. 2, pp. 185–191, Feb. 2017.
- [59] D. Astély, E. Dahlman, A. Furuskär, Y. Jading, M. Lindström, and S. Parkvall, “LTE: the evolution of mobile broadband,” *IEEE Commun. Mag.*, vol. 47, no. 4, 2009.
- [60] R. Ratasuk, N. Mangalvedhe, A. Ghosh, and B. Vejlgaard, “Narrowband LTE-M system for M2M communication,” in *VTC. IEEE*, 2014, pp. 1–5.
- [61] A. Rico-Alvarino, M. Vajapeyam, H. Xu, X. Wang, Y. Blankenship, J. Bergman, T. Tirronen, and E. Yavuz, “An overview of 3gpp enhancements on machine to machine communications,” *IEEE Commun. Mag.*, vol. 54, no. 6, pp. 14–21, 2016.
- [62] W. Yang, M. Wang, J. Zhang, J. Zou, M. Hua, T. Xia, and X. You, “Narrow-band wireless access for low-power massive Internet of Things: A bandwidth perspective,” *IEEE Trans. Wireless Commun.*, vol. 24, no. 3, pp. 124–131, May 2017.
- [63] U. Raza, P. Kulkarni, and M. Sooriyabandara, “Low power wide area networks: An overview,” *IEEE Commun. Surveys Tuts.*, vol. 19, no. 2, pp. 855–873, Secondquarter 2017.

- [64] Y. Roth, J.-B. Doré, L. Ros, and V. Berg, “A comparison of physical layers for low power wide area networks,” in *Crowncom*. Springer, 2016, pp. 261–272.
- [65] K. E. Nolan, W. Guibene, and M. Y. Kelly, “An evaluation of low power wide area network technologies for the Internet of Things,” in *IWCMC*. IEEE, 2016, pp. 439–444.
- [66] S. Chen, R. Ma, H.-H. Chen, H. Zhang, W. Meng, and J. Liu, “Machine-to-machine communications in ultra-dense networks—a survey,” *IEEE Commun. Surveys Tuts.*, 2017.
- [67] M.-T. Do, C. Goursaud, and J.-M. Gorce, “Interference modelling and analysis of random FDMA scheme in ultra narrowband networks,” in *AICT 2014*, 2014.
- [68] T. J. Myers, “Random phase multiple access system with meshing,” Aug. 2010, uS Patent 7,773,664.
- [69] L. Vangelista, A. Zanella, and M. Zorzi, “Long-range IoT technologies: The dawn of LoRa™,” in *FABULOUS*. Springer, Fourthquarter 2015, pp. 51–58.
- [70] O. Georgiou and U. Raza, “Low power wide area network analysis: Can LoRa scale?” *IEEE Wireless Commun. Lett.*, vol. 6, no. 2, pp. 162–165, Apr. 2017.
- [71] X. Wang and H. V. Poor, “Robust multiuser detection in non-Gaussian channels,” *IEEE Trans. Signal Process.*, vol. 47, no. 2, pp. 289–305, Feb. 1999.
- [72] D. S. Pham, A. M. Zoubir, R. F. Brcic, and Y. H. Leung, “A nonlinear m -estimation approach to robust asynchronous multiuser detection in Non-Gaussian noise,” *IEEE Trans. Signal Process.*, vol. 55, no. 5, pp. 1624–1633, May. 2007.

- [73] A. Mantravadi and V. V. Veeravalli, “On chip-matched filtering and discrete sufficient statistics for asynchronous band-limited CDMA systems,” *IEEE Trans. Commun.*, vol. 49, no. 8, pp. 1457–1467, Aug. 2001.
- [74] B. Babadi, “Fundamental limits and constructive methods for estimation and sensing of sparse signals,” Ph.D. dissertation, Harvard University Cambridge, Massachusetts, May 2011.
- [75] R. Rubinstein, M. Zibulevsky, and M. Elad, “Double sparsity: Learning sparse dictionaries for sparse signal approximation,” *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1553–1564, Mar. 2010.
- [76] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.
- [77] A. P. Ruszczyński, *Nonlinear optimization*. Princeton university press, 2006, vol. 13.
- [78] M. Luessi, S. D. Babacan, R. Molina, and A. K. Katsaggelos, “Bayesian simultaneous sparse approximation with smooth signals,” *IEEE Trans. Signal Process.*, vol. 61, no. 22, pp. 5716–5729, Nov. 2013.
- [79] A. E. Hoerl and R. W. Kennard, “Ridge regression: Biased estimation for nonorthogonal problems,” *Technometrics*, vol. 12, no. 1, pp. 55–67, 1970.
- [80] P. A. Devijver and J. Kittler, *Pattern recognition: A statistical approach*. Prentice-Hall London, 1982, vol. 761.
- [81] R. Kohavi *et al.*, “A study of cross-validation and bootstrap for accuracy estimation and model selection,” in *IJCAI*, vol. 14, no. 2. Stanford, CA, 1995, pp. 1137–1145.

- [82] G. H. Golub, M. Heath, and G. Wahba, “Generalized cross-validation as a method for choosing a good ridge parameter,” *Technometrics*, vol. 21, no. 2, pp. 215–223, May 1979.
- [83] S. Geisser, *Predictive inference*. CRC press, 1993, vol. 55.
- [84] R. Ward, “Compressed sensing with cross validation,” *IEEE Trans. Inf. Theory*, vol. 55, no. 12, pp. 5773–5782, Dec. 2009.
- [85] S. M. Kay, “Fundamentals of Statistical Signal Processing, Vol. I: Estimation Theory,” *Prentice Hall*, 1993.
- [86] H. L. Van Trees, *Detection, estimation, and modulation theory*. John Wiley & Sons, 2004.
- [87] H. V. Poor, *An introduction to signal detection and estimation*. Springer Science & Business Media, 2013.
- [88] I. Daubechies, M. Defrise, and C. De Mol, “An iterative thresholding algorithm for linear inverse problems with a sparsity constraint,” *Communications on pure and applied mathematics*, vol. 57, no. 11, pp. 1413–1457, 2004.
- [89] S. Arlot, A. Celisse *et al.*, “A survey of cross-validation procedures for model selection,” *Statistics surveys*, vol. 4, pp. 40–79, July 2009.
- [90] R. Zdunek and A. Cichocki, “Improved M-FOCUSS algorithm with overlapping blocks for locally smooth sparse signals,” *IEEE Trans. Signal Process.*, vol. 56, no. 10, pp. 4752–4761, Oct. 2008.
- [91] S. Verdu, *Multiuser detection*. Cambridge university press, 1998.
- [92] ———, “Minimum probability of error for asynchronous Gaussian multiple-access channels,” *IEEE Trans. Inf. Theory*, vol. 32, no. 1, pp. 85–96, Jan. 1986.

- [93] J. Lehnert and M. Pursley, “Error probabilities for binary direct-sequence spread-spectrum communications with random signature sequences,” *IEEE Trans. Wireless Commun.*, vol. 35, no. 1, pp. 87–98, Jan. 1987.
- [94] T. S. Rappaport *et al.*, *Wireless communications: principles and practice*. Prentice Hall, 1996, vol. 2.
- [95] P. S. Boonstra, B. Mukherjee, and J. M. Taylor, “A small-sample choice of the tuning parameter in ridge regression,” *Statistica Sinica*, vol. 25, no. 3, p. 1185, July 2015.
- [96] P. Billingsley *et al.*, “Limit theorems for randomly selected partial sums,” *Ann. Math. Stat.*, vol. 33, no. 1, pp. 85–92, Nov. 1962.
- [97] N. A. Weiss, *A course in probability*. Addison-Wesley, 2006.
- [98] J. A. Tropp, “Algorithms for simultaneous sparse approximation. Part II: Convex relaxation,” *Signal Processing*, vol. 86, no. 3, pp. 589–602, Mar. 2006.
- [99] R. T. Rockafellar, *Convex analysis*. Princeton university press, 2015.
- [100] J. Wu and P. Fan, “A survey on high mobility wireless communications: Challenges, opportunities and solutions,” *IEEE Access*, vol. 4, pp. 450–476, Jan. 2016.
- [101] M. J. Chu and W. E. Stark, “Effect of mobile velocity on communications in fading channels,” *IEEE Trans. Veh. Technol.*, vol. 49, no. 1, pp. 202–210, Jan. 2000.
- [102] A. J. Goldsmith and S.-G. Chua, “Variable-rate variable-power M-QAM for fading channels,” *IEEE Trans. Commun.*, vol. 45, no. 10, pp. 1218–1230, Oct. 1997.

- [103] K. Balachandran, S. R. Kadaba, and S. Nanda, "Channel quality estimation and rate adaptation for cellular mobile radio," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 7, pp. 1244–1256, July 1999.
- [104] C. Tepedelenlioglu, A. Abdi, G. B. Giannakis, and M. Kaveh, "Estimation of Doppler spread and signal strength in mobile communications with applications to handoff and adaptive transmission," *Wirel. Commun. Mob. Comput.*, vol. 1, no. 2, pp. 221–242, Mar. 2001.
- [105] A. Duel-Hallen, S. Hu, and H. Hallen, "Long-range prediction of fading signals," *IEEE Signal Process. Mag.*, vol. 17, no. 3, pp. 62–75, May 2000.
- [106] D.-S. Shiu, G. J. Foschini, M. J. Gans, and J. M. Kahn, "Fading correlation and its effect on the capacity of multielement antenna systems," *IEEE Trans. Commun.*, vol. 48, no. 3, pp. 502–513, Mar. 2000.
- [107] C.-N. Chuah *et al.*, "Capacity scaling in MIMO wireless systems under correlated fading," *IEEE Trans. Inf. Theory*, vol. 48, no. 3, pp. 637–650, Mar. 2002.
- [108] T. Yucek, R. Tannious, and H. Arslan, "Doppler spread estimation for wireless OFDM systems," in *Proc. IEEE Advances in Wired and Wireless Communication*, 2005, pp. 233–236.
- [109] S. Coleri, M. Ergen, A. Puri, and A. Bahai, "Channel estimation techniques based on pilot arrangement in OFDM systems," *IEEE Trans. Broadcast.*, vol. 48, no. 3, pp. 223–229, Sept. 2002.
- [110] F. Bellili, R. Meftehi, S. Affes, and A. Stéphenne, "Maximum likelihood SNR estimation of linearly-modulated signals over time-varying flat-fading SIMO channels," *IEEE Trans. Signal Process.*, vol. 63, no. 2, pp. 441–456, Jan. 2015.

- [111] M. Marey, M. Samir, and O. A. Dobre, “EM-based joint channel estimation and IQ imbalances for OFDM systems,” *IEEE Trans. Broadcast.*, vol. 58, no. 1, pp. 106–113, Mar. 2012.
- [112] A. Khansefid and H. Minn, “On channel estimation for massive MIMO with pilot contamination,” *IEEE Commun. Lett.*, vol. 19, no. 9, pp. 1660–1663, Sept. 2015.
- [113] Y. Li, H. Minn, and M. Z. Win, “Frequency offset estimation for MB-OFDM-based UWB systems,” *IEEE Trans. Commun.*, vol. 56, no. 6, pp. 968–979, Jun. 2008.
- [114] M. Mohammadkarimi, O. A. Dobre, and M. Z. Win, “Non-Data-Aided SNR Estimation for Multiple Antenna Systems,” in *Proc. GLOBECOM*, Washington, USA, Dec. 2016, pp. 1–5.
- [115] H. Wang, O. A. Dobre, C. Li, and D. C. Popescu, “Blind cyclostationarity-based symbol period estimation for FSK signals,” *IEEE Commun. Lett.*, vol. 19, no. 7, pp. 1149–1152, July 2015.
- [116] A. Masmoudi, F. Bellili, S. Affes, and A. Stephenne, “A non-data-aided maximum likelihood time delay estimator using importance sampling,” *IEEE Trans. Signal Process.*, vol. 59, no. 10, pp. 4505–4515, Oct. 2011.
- [117] M. Mohammadkarimi, E. Karami, O. A. Dobre, and M. Z. Win, “Number of transmit antennas detection using time-diversity of the fading channel,” *IEEE Trans. Signal Process.*, vol. 65, no. 15, pp. 4031–4046, Aug. 2017.
- [118] Y.-R. Tsai and K.-J. Yang, “Approximate ML Doppler spread estimation over flat rayleigh fading channels,” *IEEE Trans. Signal Process.*, vol. 16, no. 11, pp. 1007–1010, Nov. 2009.

- [119] A. Dogandžić and B. Zhang, “Estimating Jakes’ Doppler power spectrum parameters using the whittle approximation,” *IEEE Trans. Signal Process.*, vol. 53, no. 3, pp. 987–1005, Mar. 2005.
- [120] K. E. Baddour and N. C. Beaulieu, “Nonparametric Doppler spread estimation for flat fading channels,” in *Proc. WCNC*, New Orleans, USA, 2003, pp. 953–958.
- [121] K. D. Anim-Appiah, “On generalized covariance-based velocity estimation,” *IEEE Trans. Veh. Technol.*, vol. 48, no. 5, pp. 1546–1557, Sep. 1999.
- [122] A. Abdi, H. Zhang, and C. Tepedelenlioglu, “A unified approach to the performance analysis of speed estimation techniques in mobile communication,” *IEEE Trans. Wireless Commun.*, vol. 56, no. 1, pp. 126–135, Jan. 2008.
- [123] G. Park, D. Hong, and C. Kang, “Level crossing rate estimation with doppler adaptive noise suppression technique in frequency domain,” in *Proc. IEEE VTC*, Orlando, USA, Oct. 2003, pp. 1192–1195.
- [124] L. Krasny, H. Arslan, D. Koilpillai, and S. Chennakeshu, “Doppler spread estimation in mobile radio systems,” *IEEE Commun. Lett.*, vol. 5, no. 5, pp. 197–199, May. 2001.
- [125] E. Biglieri, R. Calderbank, A. Constantinides, A. Goldsmith, A. Paulraj, and H. V. Poor, *MIMO Wireless Communications*. Cambridge University press, 2007.
- [126] N. T. Longford, “A fast scoring algorithm for maximum likelihood estimation in unbalanced mixed models with nested random effects,” *Biometrika*, vol. 74, no. 4, pp. 817–827, 1987.

- [127] A. Swami and B. M. Sadler, “Hierarchical digital modulation classification using cumulants,” *IEEE Trans. Commun.*, vol. 48, no. 3, pp. 416–429, Mar. 2000.
- [128] K. E. Baddour and N. C. Beaulieu, “Autoregressive modeling for fading channel simulation,” *IEEE Trans. Wireless Commun.*, vol. 4, no. 4, pp. 1650–1662, July 2005.
- [129] P. Sadeghi, R. A. Kennedy, P. B. Rapajic, and R. Shams, “Finite-state markov modeling of fading channels-a survey of principles and applications,” *IEEE Signal Process. Mag.*, vol. 25, no. 5, pp. 57–80, Sept. 2008.
- [130] I. Reed, “On a moment theorem for complex Gaussian processes,” *IEEE Trans. Inf. Theory*, vol. 3, no. 8, pp. 194–195, Apr. 1962.
- [131] A. R. Gallant, *Nonlinear Statistical Models*. John Wiley & Sons, 2009, vol. 310.
- [132] N. R. Draper, H. Smith, and E. Pownell, *Applied Regression Analysis*. Wiley New York, 1966, vol. 3.
- [133] S. S. Rao and S. Rao, *Engineering Optimization: Theory and Practice*. John Wiley & Sons, 2009.
- [134] S. M. Kay, “Fundamentals of Statistical Signal Processing, Vol. I: Estimation Theory,” *Prentice Hall*, 1993.
- [135] A. M. Zoubir and B. Boashash, “The bootstrap and its application in signal processing,” *IEEE Signal Process. Mag.*, vol. 15, no. 1, pp. 56–76, Jan. 1998.
- [136] A. M. Zoubir and D. R. Iskandler, “Bootstrap methods and applications,” *IEEE Signal Process. Mag.*, vol. 24, no. 4, pp. 10–19, July 2007.

- [137] A. M. Zoubir and D. R. Iskander, "Bootstrap modeling of a class of nonstationary signals," *IEEE Trans. Signal Process.*, vol. 48, no. 2, pp. 399–408, Feb. 2000.
- [138] M. L. Honig, I. of Electrical, and E. Engineers, *Advances in multiuser detection*. Wiley Online Library, 2009.
- [139] S. Valaee and S. Shahbazpanahi, "Detecting the number of signals in wireless DS-CDMA networks," *IEEE Trans. Commun.*, vol. 56, no. 7, pp. 1189–1197, 2008.
- [140] P. Comon and C. Jutten, *Handbook of Blind Source Separation: Independent component analysis and applications*. Academic press, 2010.
- [141] A. Alexiou and M. Haardt, "Smart antenna technologies for future wireless systems: trends and challenges," *IEEE Commun. Mag.*, vol. 42, no. 9, pp. 90–97, 2004.
- [142] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: benefits and challenges," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 742–758, 2014.
- [143] E. Larsson, O. Edfors, F. Tufvesson, and T. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, 2014.
- [144] M. Tubaishat and S. Madria, "Sensor networks: an overview," *IEEE Potentials*, vol. 22, no. 2, pp. 20–23, 2003.
- [145] L. Huang and H. C. So, "Source enumeration via MDI criterion based on linear

- shrinkage estimation of noise subspace covariance matrix,” *IEEE Trans. Signal Process.*, vol. 61, no. 19, pp. 4806–4821, 2013.
- [146] L. Huang, C. Qian, H. C. So, and J. Fang, “Source enumeration for large array using shrinkage-based detectors with small samples,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 51, no. 1, pp. 344–357, 2015.
- [147] F. Gao, A. B. Gershman *et al.*, “A generalized ESPRIT approach to direction-of-arrival estimation,” *IEEE Signal Process. Lett.*, vol. 12, no. 3, pp. 254–257, 2005.
- [148] R. T. Suryaprakash and R. Nadakuditi, “Consistency and MSE performance of MUSIC-based DOA of a single source in white noise with randomly missing data,” *IEEE Trans. Signal Process.*, 2015.
- [149] M. Wax and T. Kailath, “Detection of signals by information theoretic criteria,” *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 33, no. 2, pp. 387–392, 1985.
- [150] M. Shi, Y. Bar-Ness, and W. Su, “Adaptive estimation of the number of transmit antennas,” in *Proc. IEEE MILCOM*, 2007, pp. 1–5.
- [151] K. Hassan, C. N. Nzeza, R. Gautier, E. Radoi, M. Berbineau, and I. Dayoub, “Blind detection of the number of transmitting antennas for spatially-correlated MIMO systems,” in *ITS Telecommunications (ITST), 2011 11th International Conference on*. IEEE, 2011, pp. 458–462.
- [152] M.-R. Oularbi, S. Gazor, A. Aissa-El-Bey, and S. Houcke, “Enumeration of base station antennas in a cognitive receiver by exploiting pilot patterns,” *IEEE Commun. Lett.*, vol. 17, no. 1, pp. 8–11, Jan. 2013.

- [153] —, “Exploiting the pilot pattern orthogonality of OFDMA signals for the estimation of base stations number of antennas,” in *Proc. IEEE WOSSPA*, 2013, pp. 465–470.
- [154] E. Ohlmer, T. Liang, and G. Fettweis, “Algorithm for detecting the number of transmit antennas in MIMO-OFDM systems,” in *Proc. IEEE VTC*, 2008, pp. 478–482.
- [155] —, “Algorithm for detecting the number of transmit antennas in MIMO-OFDM systems: Receiver integration,” in *Proc. IEEE VTC*, 2008, pp. 1–5.
- [156] M. Mohammadkarimi and O. A. Dobre, “A novel non-parametric method for blind identification of stbc codes,” in *Information Theory (CWIT), 2015 IEEE 14th Canadian Workshop on*. IEEE, 2015, pp. 97–100.
- [157] S. Sanayei and A. Nosratinia, “Antenna selection in MIMO systems,” *IEEE Commun. Mag.*, vol. 42, no. 10, pp. 68–73, Oct. 2004.
- [158] A. Afana, T. Ngatched, and O. A. Dobre, “Spatial modulation in MIMO limited-feedback spectrum-sharing systems with mutual interference and channel estimation errors,” *IEEE Commun. Lett.*, vol. 19, no. 10, pp. 1754–1757, Oct. 2015.
- [159] I. Berenguer, X. Wang, and V. Krishnamurthy, “Adaptive MIMO antenna selection via discrete stochastic optimization,” *IEEE Trans. Signal Process.*, vol. 53, no. 11, pp. 4315–4329, Nov. 2005.
- [160] J. Yuan, “Adaptive transmit antenna selection with pragmatic space-time trellis codes,” *IEEE Trans. Wireless Commun.*, vol. 5, no. 7, pp. 1706–1715, July. 2006.
- [161] F. Rusek, A. Lozano, and N. Jindal, “Mutual information of IID complex Gaus-

- sian signals on block Rayleigh-faded channels,” *IEEE Trans. Inf. Theory*, vol. 58, no. 1, pp. 331–340, Jan. 2012.
- [162] F. Rusek, “Achievable rates of IID Gaussian symbols on the non-coherent block-fading channel without channel distribution knowledge at the receiver,” *IEEE Trans. Wireless Commun.*, vol. 11, no. 4, pp. 1277–1282, Apr. 2012.
- [163] T. S. Rappaport, *Wireless Communications: Principles and Practice*. 2nd ed. prentice hall, 1996.
- [164] A. W. Van der Vaart, *Asymptotic Statistics*. Cambridge University Press, 2000.
- [165] M. K. Simon, *Probability Distributions Involving Gaussian Random Variables: A Handbook for Engineers and Scientists*. Springer Science & Business Media, 2007.
- [166] J. P. Kermoal, L. Schumacher, K. I. Pedersen, P. E. Mogensen, and F. Frederiksen, “A stochastic MIMO radio channel model with experimental validation,” *IEEE J. Sel. Areas Commun.*, vol. 20, no. 6, pp. 1211–1226, Aug. 2002.
- [167] I. Reed, “On a moment theorem for complex gaussian processes,” *IEEE Trans. Inform. Theory*, vol. 8, no. 3, pp. 194–195, Apr. 1962.
- [168] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory, vol. 1*. Prentice Hall, 1993.
- [169] A. Papoulis and S. U. Pillai, *Probability, Random Variables, and Stochastic Processes*. Tata McGraw-Hill Education, 2002.
- [170] A. Das and B. D. Rao, “SNR and noise variance estimation for MIMO systems,” *IEEE Trans. Signal Process.*, vol. 60, no. 8, pp. 3929–3941, Aug. 2012.

- [171] A. Wiesel, J. Goldberg, and H. Messer-Yaron, “SNR estimation in time-varying fading channels,” *IEEE Trans. Commun.*, vol. 54, no. 5, pp. 841–848, May 2006.
- [172] H. Abeida, “Data-aided SNR estimation in time-variant Rayleigh fading channels,” *IEEE Trans. Signal Process.*, vol. 58, no. 11, pp. 5496–5507, Nov. 2010.
- [173] R. López-Valcarce, J. Villares, J. Riba, W. Gappmair, and C. Mosquera, “Cramer-rao bounds for SNR estimation of oversampled linearly modulated signals,” *IEEE Trans. Signal Process.*, vol. 63, no. 7, pp. 1675–1683, Apr. 2015.
- [174] J. Wang, B. Li, M. Liu, and J. Li, “SNR estimation of time-frequency overlapped signals for underlay cognitive radio,” *IEEE Commun. Lett.*, vol. 19, no. 11, pp. 1925–1928, Sep. 2015.
- [175] Y. Fu, J. Zhu, S. Wang, and Z. Xi, “Reduced complexity SNR estimation via Kolmogorov-Smirnov test,” *IEEE Commun. Lett.*, vol. 19, no. 9, pp. 1568–1571, Apr. 2015.
- [176] Y. Fu, J. Zhu, S. Wang, and H. Zhai, “Robust non-data-aided SNR estimation for multilevel constellations via Kolmogorov–Smirnov test,” *IEEE Commun. Lett.*, vol. 18, no. 10, pp. 1707–1710, Oct. 2014.
- [177] W. Gappmair, R. López-Valcarce, and C. Mosquera, “Cramer-Rao lower bound and EM algorithm for envelope-based SNR estimation of nonconstant modulus constellations,” *IEEE Trans. Commun.*, vol. 57, no. 6, pp. 1622–1627, June 2009.
- [178] M. Alvarez-Diaz, R. Lopez-Valcarce, and C. Mosquera, “SNR estimation for multilevel constellations using higher-order moments,” *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1515–1526, Mar. 2010.

- [179] A. Stéphenne, F. Bellili, and S. Affes, “Moment-based SNR estimation over linearly-modulated wireless SIMO channels,” *IEEE Trans. Wireless Commun.*, vol. 9, no. 2, pp. 714—722, Feb. 2010.
- [180] P. Gao and C. Tepedelenlioglu, “SNR estimation for non-constant modulus constellations,” in *IEEE Trans. Signal Process.*, vol. 53, no. 3. IEEE, Mar. 2005, pp. 865—870.
- [181] J. Sun and M. C. Valenti, “Joint synchronization and SNR estimation for turbo codes in AWGN channels,” *IEEE Trans. Commun.*, vol. 53, no. 7, pp. 1136–1144, June 2005.
- [182] G. Ren, H. Zhang, and Y. Chang, “SNR estimation algorithm based on the preamble for OFDM systems in frequency selective channels,” *IEEE Trans. Commun.*, vol. 57, no. 8, pp. 2230–2234, Aug. 2009.
- [183] D. R. Pauluzzi and N. C. Beaulieu, “A comparison of SNR estimation techniques for the AWGN channel,” *IEEE Trans. Commun.*, vol. 48, no. 10, pp. 1681–1691, Oct. 2000.
- [184] N. C. Beaulieu, A. S. Toms, and D. R. Pauluzzi, “Comparison of four SNR estimators for QPSK modulations,” *IEEE Commun. Lett.*, vol. 4, no. 2, pp. 43–45, Feb. 2000.
- [185] M. A. Boujelben, F. Bellili, S. Affes, and A. Stéphenne, “SNR estimation over SIMO channels from linearly modulated signals,” *IEEE Trans. Signal Process.*, vol. 58, no. 12, pp. 6017–6028, Dec. 2010.
- [186] A. Wiesel, *Signal-to-noise-ratio estimation in time-selective fading channels*. Tel Aviv University, 2002.

- [187] M. Morelli, M. Moretti, G. Imbarlina, and N. Dimitriou, “Low complexity SNR estimation for transmissions over time-varying flat-fading channels,” in *in Proc. WCNC*. IEEE, Apr. 2009, pp. 1–4.
- [188] F. Bellili, A. Stéphenne, and S. Affes, “SNR estimation of QAM-modulated transmissions over time-varying SIMO channels,” in *in Proc. ISWCS*. IEEE, 2008, pp. 199–203.
- [189] W. Gappmair, “Cramer–Rao lower bound for non-data-aided SNR estimation of linear modulation schemes,” *IEEE Trans. Commun.*, vol. 56, no. 5, pp. 689–693, 2008.
- [190] R. Lopez-Valcarce and C. Mosquera, “Sixth-order statistics-based non-data-aided SNR estimation,” *IEEE Commun. Lett.*, vol. 11, no. 4, 2007.
- [191] I. Bousnina, M. B. B. Salah, A. Samet, and I. Dayoub, “Ricean K-factor and SNR estimation for m -psk modulated signals using the fourth-order cross-moments matrix,” *IEEE Commun. Lett.*, vol. 16, no. 8, pp. 1236–1239, June 2012.
- [192] M. Hafez, T. Khattab, and H. M. Shalaby, “Blind SNR estimation of gaussian-distributed signals in Nakagami fading channels,” *IEEE Trans. Wireless Commun.*, vol. 14, no. 7, pp. 3509–3518, July 2015.
- [193] H. Zhao, P. Fan, P. T. Mathiopoulos, and S. Papaharalabos, “On SNR estimation techniques for turbo decoding over uncorrelated Rayleigh fading channels with unknown fading parameters,” *IEEE Trans. Veh. Commun.*, vol. 58, no. 9, pp. 4955–4961, Nov. 2009.
- [194] J. F. Sevillano, I. Velez, M. Leyh, S. Lipp, A. Irizar, and L. Fontan, “In-service SNR estimation without symbol timing recovery for QPSK data transmission

- systems,” *IEEE Trans. Wireless Commun.*, vol. 6, no. 9, pp. 3202–3207, Sep. 2007.
- [195] M. B. B. Salah and A. Samet, “NDA SNR estimation using fourth-order cross-moments in time-varying single-input multiple-output channels,” *IET Commun.*, vol. 10, no. 11, pp. 1348–1354, July 2016.
- [196] J. L. Xu, W. Su, and M. Zhou, “Software-defined radio equipped with rapid modulation recognition,” *IEEE Trans. Veh. Technol.*, vol. 59, no. 4, pp. 1659–1667, 2010.
- [197] H.-C. Wu, M. Saquib, and Z. Yun, “Novel automatic modulation classification using cumulant features for communications via multipath channels,” *IEEE Trans. Wireless Commun.*, vol. 7, no. 8, pp. 3098–3105, 2008.
- [198] O. A. Dobre, M. Oner, S. Rajan, and R. Inkol, “Cyclostationarity-based robust algorithms for QAM signal identification,” *IEEE Commun. Lett.*, vol. 16, no. 1, pp. 12–15, 2012.
- [199] W. Su, “Feature space analysis of modulation classification using very high-order statistics,” *IEEE Commun. Lett.*, vol. 17, no. 9, pp. 1688–1691, 2013.
- [200] W. Su, J. L. Xu, and M. Zhou, “Real-time modulation classification based on maximum likelihood,” *IEEE Commun. Lett.*, vol. 12, no. 11, pp. 801–803, 2008.
- [201] Q. Zhang, O. Dobre, Y. Eldemerdash, S. Rajan, R. Inkol *et al.*, “Second-order cyclostationarity of BT-SCLD signals: Theoretical developments and applications to signal classification and blind parameter estimation,” *IEEE Trans. Wireless Commun.*, vol. 12, no. 4, pp. 1501–1511, 2013.

- [202] A. Al-Habashna, O. Dobre, R. Venkatesan, D. C. Popescu *et al.*, “Second-order cyclostationarity of mobile WiMAX and LTE OFDM signals and application to spectrum awareness in cognitive radio systems,” *IEEE J. Sel. Topics Signal Process.*, vol. 6, no. 1, pp. 26–42, 2012.
- [203] A. Bouzegzi, P. Ciblat, and P. Jallon, “New algorithms for blind recognition of OFDM based systems,” *Signal Process*, vol. 90, no. 3, pp. 900–913, 2010.
- [204] T. Xia and H.-C. Wu, “Blind identification of nonbinary LDPC codes using average LLR of syndrome a posteriori probability,” *IEEE Commun. Lett.*, vol. 17, no. 7, pp. 1301–1304, 2013.
- [205] H.-C. Wu, X. Huang, and D. Xu, “Pilot-free dynamic phase and amplitude estimations for wireless ICI self-cancellation coded OFDM systems,” *IEEE Trans. Broadcast.*, vol. 51, no. 1, pp. 94–105, 2005.
- [206] L. Korowajczuk, *LTE, WiMAX and WLAN network design, optimization and performance analysis*. John Wiley & Sons, 2011.
- [207] V. Choqueuse, K. Yao, L. Collin, and G. Burel, “Hierarchical space-time block code recognition using correlation matrices,” *IEEE Commun. Lett.*, vol. 7, no. 9, pp. 3526–3534, 2008.
- [208] Y. Eldemerdash, O. Dobre, M. Marey, G. K. Karagiannidis, B. Liao *et al.*, “An efficient algorithm for space-time block code classification,” in *Global Communications Conference (GLOBECOM), 2013 IEEE*. IEEE, 2013, pp. 3329–3334.
- [209] E. Karami, O. Dobre *et al.*, “Identification of SM-OFDM and AL-OFDM signals based on their second-order cyclostationarity,” *IEEE Trans. Veh. Technol.*, vol. 64, no. 3, pp. 942–953, 2015.

- [210] Y. Eldemerdash, O. Dobre, B. J. Liao *et al.*, “Blind identification of SM and alamouti STBC-OFDM signals,” *IEEE Trans. Wireless Commun.*, vol. 14, no. 2, pp. 972–982, 2015.
- [211] M. Marey, O. Dobre, R. Inkol *et al.*, “Blind STBC identification for multiple-antenna OFDM systems,” *IEEE Trans. Commun.*, vol. 62, no. 5, pp. 1554–1567, 2014.
- [212] K. Hassan, I. Dayoub, W. Hamouda, C. N. Nzéza, and M. Berbineau, “Blind digital modulation identification for spatially-correlated MIMO systems,” *IEEE Trans. Wireless Commun.*, vol. 11, no. 2, pp. 683–693, 2012.
- [213] M. S. Muhlhaus, M. Oner, O. Dobre, F. K. Jondral *et al.*, “A low complexity modulation classification algorithm for MIMO systems,” *IEEE Commun. Lett.*, vol. 17, no. 10, pp. 1881–1884, 2013.
- [214] V. Choqueuse, K. Yao, L. Collin, and G. Burel, “Blind recognition of linear space time block codes,” in *in Proc. ICASSP*. IEEE, 2008, pp. 2833–2836.
- [215] M. R. DeYoung, R. W. Heath, and B. L. Evans, “Using higher order cyclostationarity to identify space-time block codes,” in *in Proc. GLOBECOM*. IEEE, 2008, pp. 1–5.
- [216] P. H. Kvam and B. Vidakovic, *Nonparametric statistics with applications to science and engineering*. John Wiley & Sons, 2007, vol. 653.
- [217] G. Zhang, X. Wang, Y.-C. Liang, and J. Liu, “Fast and robust spectrum sensing via kolmogorov-smirnov test,” *IEEE Trans. Commun.*, vol. 58, no. 12, pp. 3410–3416, 2010.