

# Complex Valued Nonlinear Adaptive Filters

Noncircularity, Widely Linear and  
Neural Models

**Danilo P. Mandic**

*Imperial College London, UK*

**Vanessa Su Lee Goh**

*Shell EP, Europe*



A John Wiley and Sons, Ltd, Publication

# Contents

<b>Preface</b>	<b>xiii</b>
<b>Acknowledgements</b>	<b>xvii</b>
<b>1 The Magic of Complex Numbers</b>	<b>1</b>
1.1 History of Complex Numbers	2
1.1.1 Hypercomplex Numbers	7
1.2 History of Mathematical Notation	8
1.3 Development of Complex Valued Adaptive Signal Processing	9
<b>2 Why Signal Processing in the Complex Domain?</b>	<b>13</b>
2.1 Some Examples of Complex Valued Signal Processing	13
2.1.1 Duality Between Signal Representations in $\mathbb{R}$ and $\mathbb{C}$	18
2.2 Modelling in $\mathbb{C}$ is Not Only Convenient But Also Natural	19
2.3 Why Complex Modelling of Real Valued Processes?	20
2.3.1 Phase Information in Imaging	20
2.3.2 Modelling of Directional Processes	22
2.4 Exploiting the Phase Information	23
2.4.1 Synchronisation of Real Valued Processes	24
2.4.2 Adaptive Filtering by Incorporating Phase Information	25
2.5 Other Applications of Complex Domain Processing of Real Valued Signals	26
2.6 Additional Benefits of Complex Domain Processing	29
<b>3 Adaptive Filtering Architectures</b>	<b>33</b>
3.1 Linear and Nonlinear Stochastic Models	34
3.2 Linear and Nonlinear Adaptive Filtering Architectures	35
3.2.1 Feedforward Neural Networks	36
3.2.2 Recurrent Neural Networks	37
3.2.3 Neural Networks and Polynomial Filters	38
3.3 State Space Representation and Canonical Forms	39

<b>4</b>	<b>Complex Nonlinear Activation Functions</b>	<b>43</b>
4.1	Properties of Complex Functions	43
4.1.1	Singularities of Complex Functions	45
4.2	Universal Function Approximation	46
4.2.1	Universal Approximation in $\mathbb{R}$	47
4.3	Nonlinear Activation Functions for Complex Neural Networks	48
4.3.1	Split-complex Approach	49
4.3.2	Fully Complex Nonlinear Activation Functions	51
4.4	Generalised Splitting Activation Functions (GSAF)	53
4.4.1	The Clifford Neuron	53
4.5	Summary: Choice of the Complex Activation Function	54
<b>5</b>	<b>Elements of <math>\mathbb{C}\mathbb{R}</math> Calculus</b>	<b>55</b>
5.1	Continuous Complex Functions	56
5.2	The Cauchy–Riemann Equations	56
5.3	Generalised Derivatives of Functions of Complex Variable	57
5.3.1	$\mathbb{C}\mathbb{R}$ Calculus	59
5.3.2	Link between $\mathbb{R}$ - and $\mathbb{C}$ -derivatives	60
5.4	$\mathbb{C}\mathbb{R}$ -derivatives of Cost Functions	62
5.4.1	The Complex Gradient	62
5.4.2	The Complex Hessian	64
5.4.3	The Complex Jacobian and Complex Differential	64
5.4.4	Gradient of a Cost Function	65
<b>6</b>	<b>Complex Valued Adaptive Filters</b>	<b>69</b>
6.1	Adaptive Filtering Configurations	70
6.2	The Complex Least Mean Square Algorithm	73
6.2.1	Convergence of the CLMS Algorithm	75
6.3	Nonlinear Feedforward Complex Adaptive Filters	80
6.3.1	Fully Complex Nonlinear Adaptive Filters	80
6.3.2	Derivation of CNGD using $\mathbb{C}\mathbb{R}$ calculus	82
6.3.3	Split-complex Approach	83
6.3.4	Dual Univariate Adaptive Filtering Approach (DUAF)	84
6.4	Normalisation of Learning Algorithms	85
6.5	Performance of Feedforward Nonlinear Adaptive Filters	87
6.6	Summary: Choice of a Nonlinear Adaptive Filter	89
<b>7</b>	<b>Adaptive Filters with Feedback</b>	<b>91</b>
7.1	Training of IIR Adaptive Filters	92
7.1.1	Coefficient Update for Linear Adaptive IIR Filters	93
7.1.2	Training of IIR filters with Reduced Computational Complexity	96

7.2	Nonlinear Adaptive IIR Filters: Recurrent Perceptron	97
7.3	Training of Recurrent Neural Networks	99
7.3.1	Other Learning Algorithms and Computational Complexity	102
7.4	Simulation Examples	102
<b>8</b>	<b>Filters with an Adaptive Stepsize</b>	<b>107</b>
8.1	Benveniste Type Variable Stepsize Algorithms	108
8.2	Complex Valued GNGD Algorithms	110
8.2.1	Complex GNGD for Nonlinear Filters (CFANNGD)	112
8.3	Simulation Examples	113
<b>9</b>	<b>Filters with an Adaptive Amplitude of Nonlinearity</b>	<b>119</b>
9.1	Dynamical Range Reduction	119
9.2	FIR Adaptive Filters with an Adaptive Nonlinearity	121
9.3	Recurrent Neural Networks with Trainable Amplitude of Activation Functions	122
9.4	Simulation Results	124
<b>10</b>	<b>Data-reusing Algorithms for Complex Valued Adaptive Filters</b>	<b>129</b>
10.1	The Data-reusing Complex Valued Least Mean Square (DRCLMS) Algorithm	129
10.2	Data-reusing Complex Nonlinear Adaptive Filters	131
10.2.1	Convergence Analysis	132
10.3	Data-reusing Algorithms for Complex RNNs	134
<b>11</b>	<b>Complex Mappings and Möbius Transformations</b>	<b>137</b>
11.1	Matrix Representation of a Complex Number	137
11.2	The Möbius Transformation	140
11.3	Activation Functions and Möbius Transformations	142
11.4	All-pass Systems as Möbius Transformations	146
11.5	Fractional Delay Filters	147
<b>12</b>	<b>Augmented Complex Statistics</b>	<b>151</b>
12.1	Complex Random Variables (CRV)	152
12.1.1	Complex Circularity	153
12.1.2	The Multivariate Complex Normal Distribution	154
12.1.3	Moments of Complex Random Variables (CRV)	157
12.2	Complex Circular Random Variables	158
12.3	Complex Signals	159
12.3.1	Wide Sense Stationarity, Multicorrelations, and Multispectra	160
12.3.2	Strict Circularity and Higher-order Statistics	161
12.4	Second-order Characterisation of Complex Signals	161
12.4.1	Augmented Statistics of Complex Signals	161
12.4.2	Second-order Complex Circularity	164

<b>13 Widely Linear Estimation and Augmented CLMS (ACLMS)</b>	<b>169</b>
13.1 Minimum Mean Square Error (MMSE) Estimation in $\mathbb{C}$	169
13.1.1 Widely Linear Modelling in $\mathbb{C}$	171
13.2 Complex White Noise	172
13.3 Autoregressive Modelling in $\mathbb{C}$	173
13.3.1 Widely Linear Autoregressive Modelling in $\mathbb{C}$	174
13.3.2 Quantifying Benefits of Widely Linear Estimation	174
13.4 The Augmented Complex LMS (ACLMS) Algorithm	175
13.5 Adaptive Prediction Based on ACLMS	178
13.5.1 Wind Forecasting Using Augmented Statistics	180
<b>14 Duality Between Complex Valued and Real Valued Filters</b>	<b>183</b>
14.1 A Dual Channel Real Valued Adaptive Filter	184
14.2 Duality Between Real and Complex Valued Filters	186
14.2.1 Operation of Standard Complex Adaptive Filters	186
14.2.2 Operation of Widely Linear Complex Filters	187
14.3 Simulations	188
<b>15 Widely Linear Filters with Feedback</b>	<b>191</b>
15.1 The Widely Linear ARMA (WL-ARMA) Model	192
15.2 Widely Linear Adaptive Filters with Feedback	192
15.2.1 Widely Linear Adaptive IIR Filters	195
15.2.2 Augmented Recurrent Perceptron Learning Rule	196
15.3 The Augmented Complex Valued RTRL (ACRTRL) Algorithm	197
15.4 The Augmented Kalman Filter Algorithm for RNNs	198
15.4.1 EKF Based Training of Complex RNNs	200
15.5 Augmented Complex Unscented Kalman Filter (ACUKF)	200
15.5.1 State Space Equations for the Complex Unscented Kalman Filter	201
15.5.2 ACUKF Based Training of Complex RNNs	202
15.6 Simulation Examples	203
<b>16 Collaborative Adaptive Filtering</b>	<b>207</b>
16.1 Parametric Signal Modality Characterisation	207
16.2 Standard Hybrid Filtering in $\mathbb{R}$	209
16.3 Tracking the Linear/Nonlinear Nature of Complex Valued Signals	210
16.3.1 Signal Modality Characterisation in $\mathbb{C}$	211
16.4 Split vs Fully Complex Signal Natures	214
16.5 Online Assessment of the Nature of Wind Signal	216
16.5.1 Effects of Averaging on Signal Nonlinearity	216
16.6 Collaborative Filters for General Complex Signals	217
16.6.1 Hybrid Filters for Noncircular Signals	218
16.6.2 Online Test for Complex Circularity	220

<b>17 Adaptive Filtering Based on EMD</b>	<b>221</b>
17.1 The Empirical Mode Decomposition Algorithm	222
17.1.1 Empirical Mode Decomposition as a Fixed Point Iteration	223
17.1.2 Applications of Real Valued EMD	224
17.1.3 Uniqueness of the Decomposition	225
17.2 Complex Extensions of Empirical Mode Decomposition	226
17.2.1 Complex Empirical Mode Decomposition	227
17.2.2 Rotation Invariant Empirical Mode Decomposition (RIEMD)	228
17.2.3 Bivariate Empirical Mode Decomposition (BEMD)	228
17.3 Addressing the Problem of Uniqueness	230
17.4 Applications of Complex Extensions of EMD	230
<b>18 Validation of Complex Representations – Is This Worthwhile?</b>	<b>233</b>
18.1 Signal Modality Characterisation in $\mathbb{R}$	234
18.1.1 Surrogate Data Methods	235
18.1.2 Test Statistics: The DVV Method	237
18.2 Testing for the Validity of Complex Representation	239
18.2.1 Complex Delay Vector Variance Method (CDVV)	240
18.3 Quantifying Benefits of Complex Valued Representation	243
18.3.1 Pros and Cons of the Complex DVV Method	244
<b>Appendix A: Some Distinctive Properties of Calculus in <math>\mathbb{C}</math></b>	<b>245</b>
<b>Appendix B: Liouville's Theorem</b>	<b>251</b>
<b>Appendix C: Hypercomplex and Clifford Algebras</b>	<b>253</b>
C.1 Definitions of Algebraic Notions of Group, Ring and Field	253
C.2 Definition of a Vector Space	254
C.3 Higher Dimension Algebras	254
C.4 The Algebra of Quaternions	255
C.5 Clifford Algebras	256
<b>Appendix D: Real Valued Activation Functions</b>	<b>257</b>
D.1 Logistic Sigmoid Activation Function	257
D.2 Hyperbolic Tangent Activation Function	258
<b>Appendix E: Elementary Transcendental Functions (ETF)</b>	<b>259</b>
<b>Appendix F: The <math>\mathcal{O}</math> Notation and Standard Vector and Matrix Differentiation</b>	<b>263</b>
F.1 The $\mathcal{O}$ Notation	263
F.2 Standard Vector and Matrix Differentiation	263

<b>Appendix G: Notions From Learning Theory</b>	<b>265</b>
G.1 Types of Learning	266
G.2 The Bias–Variance Dilemma	266
G.3 Recursive and Iterative Gradient Estimation Techniques	267
G.4 Transformation of Input Data	267
<b>Appendix H: Notions from Approximation Theory</b>	<b>269</b>
<b>Appendix I: Terminology Used in the Field of Neural Networks</b>	<b>273</b>
<b>Appendix J: Complex Valued Pipelined Recurrent Neural Network (CPRNN)</b>	<b>275</b>
J.1 The Complex RTRL Algorithm (CRTRL) for CPRNN	275
J.1.1 Linear Subsection Within the PRNN	277
<b>Appendix K: Gradient Adaptive Step Size (GASS) Algorithms in <math>\mathbb{R}</math></b>	<b>279</b>
K.1 Gradient Adaptive Stepsize Algorithms Based on $\partial J/\partial \mu$	280
K.2 Variable Stepsize Algorithms Based on $\partial J/\partial \varepsilon$	281
<b>Appendix L: Derivation of Partial Derivatives from Chapter 8</b>	<b>283</b>
L.1 Derivation of $\partial e(k)/\partial w_n(k)$	283
L.2 Derivation of $\partial e^*(k)/\partial \varepsilon(k-1)$	284
L.3 Derivation of $\partial w(k)/\partial \varepsilon(k-1)$	286
<b>Appendix M: A <i>Posteriori</i> Learning</b>	<b>287</b>
M.1 <i>A Posteriori</i> Strategies in Adaptive Learning	288
<b>Appendix N: Notions from Stability Theory</b>	<b>291</b>
<b>Appendix O: Linear Relaxation</b>	<b>293</b>
O.1 Vector and Matrix Norms	293
O.2 Relaxation in Linear Systems	294
O.2.1 Convergence in the Norm or State Space?	297
<b>Appendix P: Contraction Mappings, Fixed Point Iteration and Fractals</b>	<b>299</b>
P.1 Historical Perspective	303
P.2 More on Convergence: Modified Contraction Mapping	305
P.3 Fractals and Mandelbrot Set	308
<b>References</b>	<b>309</b>
<b>Index</b>	<b>321</b>