

Serdar Yüksel
Tamer Başar

Stochastic Teams, Games, and Control under Information Constraints

Systems & Control: Foundations & Applications

Series Editor

Tamer Başar, University of Illinois at Urbana-Champaign, Urbana, IL, USA

Editorial Board

Karl Johan Åström, Lund University, Lund, Sweden

Han-Fu Chen, Academia Sinica, Beijing, China

Bill Helton, University of California, San Diego, CA, USA

Alberto Isidori, Sapienza University of Rome, Rome, Italy

Miroslav Krstic, University of California, San Diego, CA, USA

H. Vincent Poor, Princeton University, Princeton, NJ, USA

H. Mete Soner, Princeton University, Princeton, NJ, USA

Former Editorial Board Member

Roberto Tempo, (1956-2017), CNR-IEIIT, Politecnico di Torino, Italy

Serdar Yüksel • Tamer Başar

Stochastic Teams, Games, and Control under Information Constraints

Serdar Yüksel
Mathematics & Statistics
Queen's University
Kingston, ON, Canada

Tamer Başar 
Coordinated Science Laboratory
University of Illinois Urbana-Champaign
Urbana, IL, USA

ISSN 2324-9749 ISSN 2324-9757 (electronic)
Systems & Control: Foundations & Applications
ISBN 978-3-031-54070-7 ISBN 978-3-031-54071-4 (eBook)
<https://doi.org/10.1007/978-3-031-54071-4>

Mathematics Subject Classification: 93E20, 93E03, 91A10, 94A15, 93E15

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Switzerland AG 2024

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This book is published under the imprint Birkhäuser, www.birkhauser-science.com by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

If disposing of this product, please recycle the paper.

*To
Adem, Helin, Kate, Sait, and Saibe Yüksel
(S.Y.)
and
Altan, Koray, Eren, Zara, Alya, and Reya
(T.B.)*

Preface

The progress in the field of stochastic decision making since the publication of our 2013 book *Stochastic Networked Control Systems: Stabilization and Optimization Under Information Constraints* in the same Series, and expansion and evolution of the topical areas the field encompasses, has given us the impetus to expand on our earlier coverage, to go also beyond stochastic teams to include stochastic games as well.

Given our initial motivation, this book may first be viewed as an addendum to our 2013 book. However, the progress in the field and the encouraging response we have received from the community have led us to make such substantial changes in the coverage (as noted below) that the current book can no longer be viewed as a *2nd edition*. Accordingly, some of the material in our 2013 book have not been included in this book and most of the new material here cannot be found anywhere else in such a complete and unified manner.

Our goal here has been to provide a comprehensive, mathematically rigorous, and accessible treatment of the interaction between information, decision, and control in multi-agent systems. We provide a comprehensive theory and coverage on information structures for stochastic teams as well as stochastic games.

The fields of decentralized stochastic control and networked control have reached a high level of maturity. On decentralization, there has been significant progress on existence, approximations, and optimality analyses, and many outstanding issues have been resolved, even though some still remain open. On networked control, there have been major findings on solution methods for both stabilization and optimization under various information constraints for both linear and non-linear models, and it is perhaps fair to say that the field is slowly converging to a phase where the rate of progress does not have the bursts of initial stages of discovery. Notably, non-linear systems have also found their place toward a broad treatment of information design in control theory and the theory of multi-agent decision making. Importantly, the fields noted above are now important, closely related sub-disciplines of their own, and will likely keep generations of students immersed in the problems involving the interaction of information and control, and decision making in general.

Additionally, we have included in the book a comprehensive treatment of information structures as they arise in stochastic game theory. Game theoretic framework brings in a multitude of additional intricacies with regard to the value and shaping of information, and even the most elementary arguments on the role of information in stochastic teams do not find universal applicability in game theory due to the competitive nature of equilibrium concepts, and fragility of information structure dependent properties, which we cover in detail in the book. We present foundational material regarding the interaction of information and decisions, toward a comprehensive theory of information structures in stochastic teams, games, and (networked) control.

As in the 2013 book, we take information structures both as passively provided (as, for example, in the theory of partially observed Markov decision processes, decentralized stochastic control as well as dynamic game theory) and as actively designed (via encoding and decoding policies when a physical information channel is present, as in communication theoretic problems or signaling game problems). For the latter case, in the context of networked control and signaling games, we may view encoders and decoders also as agents.

Finally, we provide in the book a comprehensive discussion on reinforcement learning theory in the context of the topics presented in the book. Reinforcement learning algorithms use feedback information to improve performance in a control task. The field of reinforcement learning has advanced significantly both in terms of fundamental theoretical contributions and applications. However, it is important to also understand the limitations of learning theory in the presence of model uncertainty and limited information. To this end, we found it appropriate to present several fundamental results toward generating a more complete book. Accordingly, a systematic discussion on reinforcement learning theory under a variety of information structures is presented.

To keep the length of the book within a manageable size, proofs of certain results with particularly tedious derivations have not been included, with relevant references provided as appropriate. Also, we have kept the analysis throughout the book primarily in the discrete-time domain, while we recognize that much research has been done in the continuous-time setup as well. Bibliographic remarks provide some references in this domain.

The book draws and utilizes a diverse set of tools (of both conceptual and analytical nature) from various disciplines, including stochastic control, team theory, game theory, information theory, probability theory, and dynamical systems, and amalgamates them into a unified, coherent, and applicable theory. It could be used as a textbook or as an accompanying text in a graduate-level course on the topic.

Kingston, ON, Canada
Urbana, IL, USA

Serdar Yüksel
Tamer Başar

Acknowledgements

As in the 2013 book, the work here would not have been possible had we not have the opportunity of interacting with several outstanding scholars and extremely resourceful students.

We would like to thank many individuals whose works, either independently or through collaboration, have shaped the contents of the book. Of particular mention are Fady Alajaji, Gürdal Arslan, Graeme Baker, Diego Cardenas, Giacomo Como, Alexander Condello, Liam Cregg, Philippe Furrer, Nicolas Garcia, Sinan Gezici, Meysam Ghomi, Adina Goldberg, Abhishek Gupta, Ian Hogeboom-Burr, Nuh Aygün Dalkıran Alireza Jalaeian, Andrew Johnston, Ali Devran Kara, Christoph Kawan, Ertan Kazıklı, Jonathan Keeler, Tamás Linder, Aditya Mahajan, Nuno C. Martins, Curtis McDonald, Sean Meyn, Tobias Oechtering, Ayça Özçelikkale, Daniel Quevedo, Naci Saldi, Sina Sanjari, Serkan Sarıtaş, Ryan Simpson, Mikael Skoglund, Nevroz Şen, Sekhar Tatikonda, Richard Wood, Emily Wright, Bora Yongaçoğlu, and Ali Zaidi.

We also had the honor of working with three outstanding scholars who are unfortunately not with us today. Prof. Ari Arapostathis, Prof. Milan Derpich, and Ramiro Zurkowski contributed to several results in the book, both directly and indirectly. It has been a privilege to be able to work with them.

Awni AlTabaa, Liam Cregg, Nicolas Garcia, Meysam Ghomi, Ian Hogeboom-Burr, Ali D. Kara, Ertan Kazıklı, Jonathan Keeler, Curtis McDonald, Lorenzo Miretti, Somnath Pradhan, Naci Saldi, Nevroz Şen Sina Sanjari, Ryan Simpson, Serkan Sarıtaş, Bora Yongaçoğlu, and Omar Mrani-Zentar have read different various sections or versions of the book, and provided substantial technical comments.

The book has also benefitted from technical discussions with many of our colleagues and students, including Tansu Alpcan, Yücel Altuğ, Orhan Arıkan, Charalambos D. Charalambous, Debasish Chatterjee, Alexander Condello, Tolga Duman, Atilla Eryılmaz, Adina Goldberg, Orhan Çağrı Imer, Akshay Kashyap, Ankur Kulkarni, Marcos M. Vasconcelos, Ather Gattami, Rayadurgam Srikant, Sekhar Tatikonda, Anant Sahai, Demos Teneketzis, and Venu Veeravalli, to whom we also convey our thanks.

Finally, we acknowledge the stimulating, conducive environment of the Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, where the initial seeds of this book project were planted when the first author was a graduate student under the supervision of the second. The bulk of the work was carried out after the first author joined Queen's University, which provided a very stimulating environment in the Department of Mathematics and Statistics for the project to be completed in the company of outstanding scholars and students, notably through the Mathematics and Engineering Program (also known as Apple-Math at Queen's). Bilkent University, where the first author holds an adjunct appointment, also provided a productive venue for collaborations over many years. We thank all these institutions and many of our colleagues there.

Contents

1	Introduction	1
1.1	Information in Decision and Control Theory, Networked Control, and Stochastic Games	1
1.2	Contents of the Book	7
1.2.1	Parts I and II: Mathematical Theory of Information Structures for Stochastic Teams and Games	8
1.2.2	Parts III and IV: Information Structure Design Under Data Constraints via Signaling/Coding in Stochastic Teams and Games	8
1.2.3	Part V: Reinforcement Learning Under Full and Partial, Decentralized, and Networked Information	9
 Part I Stochastic Teams and a Mathematical Theory of Information Structures		
2	An Introduction to Stochastic Team Theory and Solutions to Static Teams	13
2.1	Introduction	13
2.2	Static Decision Problems	14
2.2.1	A General Introduction to Static Decision Problems...	14
2.2.2	An Example of a Finite Stochastic Team	19
2.3	Dynamic Decision Problems	27
2.3.1	A General Introduction to Dynamic Decision Problems	27
2.3.2	An Example of a Finite Dynamic Team	34
2.4	Team-Optimal Solutions for Static Teams	40
2.4.1	Teams with Finite Measurement Spaces	40
2.4.2	Teams on Finite-Dimensional Spaces	50

- 2.4.3 Static Teams with Quadratic Loss and a Hilbert Space Argument to Optimality and Stationarity 54
- 2.4.4 Positively Exponentiated Quadratic Loss 71
- 2.5 Concluding Remarks 77
- 2.6 Bibliographic Notes 77
- 3 Information Structures in Decentralized Stochastic Control 81**
 - 3.1 Introduction 81
 - 3.2 Witsenhausen’s Intrinsic Model and Classification of Information Structures 84
 - 3.3 Nonclassical Information Structures and Witsenhausen’s Counterexample 88
 - 3.3.1 Witsenhausen’s Counterexample with Discrete Distributions 90
 - 3.3.2 Witsenhausen’s Counterexample (in Original Form) .. 91
 - 3.3.3 Numerical Procedures and a Convergence Result for Asymptotic Optimality of Quantized Policies 94
 - 3.3.4 Generalized Gaussian Test Channel 95
 - 3.4 Static Reduction of Sequential Dynamic Teams: Policy-Dependence Under Partially Nested and Policy-Independence Under Non-classical Information 97
 - 3.4.1 Static Reduction I: Classical and Quasi-Classical Information Structures and Their Policy-Dependent Reduction 98
 - 3.4.2 Static Reduction II: Non-classical Information Structures and Their Policy-Independent Reduction ... 100
 - 3.4.3 Further Differences Between Policy-Dependent and Policy-Independent Static Reductions 104
 - 3.4.4 Two Refinements of Policy-Dependent Static Reduction 105
 - 3.4.5 All Stochastic Dynamic Teams Are Nearly Static (with Independent Measurements) Reducible ... 106
 - 3.5 Beyond the One-Shot DM Model: An Agent-Wise Intrinsic Model and Agent-Wise Reductions 106
 - 3.5.1 An Agent-Wise Intrinsic Model with Agents (or Sub-teams of DMs) Acting Multiple Times 107
 - 3.5.2 Agent-Wise (or Sub-team-wise) Independent-Data, Nested Independent and Partially Nested Independent Reductions 107
 - 3.6 Equivalent Static Reductions Preserve Optimality but May Not Person-by-Person-Optimality or Stationarity 108

- 3.7 Convexity of Stochastic Team Problems: Information Structure and Cost Dependent Convexity Properties 113
 - 3.7.1 Convexity of Static Team Problems and an Equivalent Representation of Cost Functions 114
 - 3.7.2 Convexity of Dynamic Teams 117
 - 3.7.3 Revisiting Non-classical Information, Witsenhausen’s Counterexample and Lack of Convexity Due to Signaling for LQG Teams 118
- 3.8 Expansion of Information Structures, and Signaling via Control and Stochastic Relaxations of Nestedness 120
 - 3.8.1 Expansion of Information Structures: A Recipe for Identifying Sufficient Information 120
 - 3.8.2 Signaling Through Control Actions and Bismut’s Method 125
 - 3.8.3 No-Incentive for Signaling Under Partial Nestedness and Stochastic Partial Nestedness 127
- 3.9 Non-sequential Decentralized Stochastic Control: Causality, Solvability and Static Reduction 129
 - 3.9.1 Intrinsic Model for Non-sequential Stochastic Control 130
 - 3.9.2 Causality, Causal Implementability and Solvability ... 131
 - 3.9.3 Static Reduction of Non-sequential Stochastic Dynamic Teams Under Property C 137
- 3.10 Concluding Remarks 139
- 3.11 Bibliographic Notes 139
- 4 Existence, Structure, and Approximations to Optimality: Strategic Measures in Decentralized Stochastic Control 141**
 - 4.1 Decentralized Strategic Measures 142
 - 4.1.1 Measurable Policies as a Subset of Randomized Policies and Strategic Measures 143
 - 4.1.2 Sets of Strategic Measures for Static Teams 144
 - 4.1.3 Sets of Strategic Measures for Dynamic Teams in the Absence of Static Reduction 148
 - 4.1.4 Measurability Properties of Sets of Strategic Measures 152
 - 4.2 Relative Compactness and Closedness of Strategic Measures, and Existence of Optimal Policies 153
 - 4.3 Convex Relaxations 162
 - 4.3.1 Non-signaling Relaxation 163
 - 4.3.2 Quantum Correlations and Quantum Team Decision Policies 166
 - 4.3.3 Relations Between Convex Relaxations 170

- 4.4 Near Optimality of Finite Approximations in Decentralized Stochastic Control via Quantization of Measurements and Actions 174
 - 4.4.1 Finite Measurement Approximate Models 178
 - 4.4.2 Specific Results for Witsenhausen’s Problem 181
- 4.5 Concluding Remarks..... 182
- 4.6 Bibliographic Notes..... 183
- 5 Control Policy Topologies and Dynamic Programming Methods in Decentralized Stochastic Control 185**
 - 5.1 Introduction 185
 - 5.2 Dynamic Programming I: Common Information as Information State..... 186
 - 5.3 Dynamic Programming II: Witsenhausen’s Standard Form for Sequential Finite Teams..... 190
 - 5.4 Dynamic Programming III: Universal Dynamic Program for all Sequential Teams in General Spaces 192
 - 5.4.1 Controlled Markov Model Formulation for the Dynamic Program 193
 - 5.4.2 Well-Posedness of the Dynamic Program (5.8) 196
 - 5.4.3 Existence of Optimal Policies Through the Dynamic Programming Formulation 199
 - 5.5 Control Topology via a Cartesian Product of Individual Agent Policies 203
 - 5.5.1 On Classical Deterministic and Stochastic Control and Utility of Young Measures..... 203
 - 5.5.2 Trading-Off Continuity with Compactness: A Product Space Approach on Individual Control Policies 206
 - 5.5.3 A Weak*-Topology on Randomized Team Decision/Control Policies 207
 - 5.5.4 Hilbert Space Construction for Static LQG Teams..... 211
 - 5.6 Some Applications and Revisiting Existence Results for Classical (Single-DM) Stochastic Control 212
 - 5.6.1 Witsenhausen’s Counterexample with Gaussian Variables 213
 - 5.6.2 Existence for Partially Observable Markov Decision Processes (POMDPs) 213
 - 5.6.3 Revisiting Fully Observable Markov Decision Processes with the Construction Presented in the Chapter 214
 - 5.7 Concluding Remarks..... 215
 - 5.8 Bibliographic Notes..... 215

6 Many-Agent Convex and Non-convex Exchangeable (Mean-Field) Teams and Optimality of Symmetric Policies 217

6.1 Introduction 218

6.2 Convex Static Teams with Large Numbers of DMs 219

6.2.1 Asymptotically Optimal Policies as a Limit of Finite Team Optimal Policies 221

6.2.2 Symmetric Teams 222

6.2.3 Optimal Solutions for Mean-Field Teams as Limits of Optimal Policies for Finite Symmetric Teams 224

6.2.4 An Existence Theorem on Globally Optimal Policies for Mean-Field Team Problems 225

6.2.5 Examples 226

6.3 Convex Dynamic Teams with Large Numbers of DMs 228

6.3.1 Finite Horizon Convex Dynamic Team Problems with a Symmetric Information Structure 230

6.3.2 Optimality of Symmetric Policies for Convex Dynamic Teams with a Symmetric Information Structure 232

6.3.3 Convex Mean-Field Teams with a Symmetric Information Structure 234

6.4 Non-convex Exchangeable Teams and Mean-Field Optimality of Randomized Symmetric Policies with No Common Randomness 238

6.4.1 Exchangeable Static Teams with Large Numbers of DMs 238

6.4.2 Dynamic Exchangeable Teams 245

6.4.3 Approximate Optimality of Symmetrically Independent Policies for Exchangeable N -DM Stochastic Teams 251

6.5 Appendix: Proofs 254

6.5.1 Proof of Theorem 6.2.1 254

6.5.2 Proof of Theorem 6.2.5 255

6.5.3 Proof of Proposition 6.2.1 257

6.6 Concluding Remarks 258

6.7 Bibliographic Notes 258

7 Comparison of Information Structures and Their Blackwell Ordering 261

7.1 Introduction 261

7.2 Value of Information and Preliminary Results on Comparison of Information Structures 264

7.3 Blackwell Ordering and Comparison of Information Structures . 270

7.3.1 Blackwell Ordering: Forward 271

7.3.2 Blackwell Ordering: Converse 273

7.4	Multi-Agent Setting and Correlated Garblings	274
7.5	Appendix: Proofs	275
	7.5.1 Proof of Theorem 7.3.4	275
	7.5.2 Proof of Theorem 7.4.1	278
7.6	Concluding Remarks	280
7.7	Bibliographic Notes	281
8	Continuity, Existence and Topological Properties of Optimization on the Space of Information Structures	285
8.1	Introduction	285
8.2	Topological Characterization of Measurement Channels	286
8.3	Single Stage: Continuity of the Optimal Cost in Channels	288
	8.3.1 Continuity Properties Under Weak Convergence	288
	8.3.2 Continuity Properties Under Setwise Convergence	291
	8.3.3 Continuity Under Total Variation	293
	8.3.4 Further Sufficient Conditions for Continuity Under Weak Convergence	294
8.4	Single Stage: Compactness Properties of Spaces of Channels	296
8.5	Quantizers as a Class of Channels	298
8.6	The Multi-Stage Case	304
8.7	Continuity and Robustness to the Prior Probability Model	307
	8.7.1 Single-Stage Setup	307
	8.7.2 Robustness to Incorrect Priors and Mismatch Bounds for Single Stage Problems	309
	8.7.3 Continuity and Robustness to Incorrect Priors and Mismatch Bounds for Multi-Stage Problems	309
8.8	Multi-Agent Setting	311
8.9	All Stochastic Dynamic Teams Are Nearly Static Reducible via Additive Perturbation of Information Structures ..	315
8.10	Appendix: Proofs	317
	8.10.1 Proof of Lemma 8.2.1	317
	8.10.2 Proof of Theorem 8.3.1	318
	8.10.3 Proof of Lemma 8.4.3	319
	8.10.4 Proof of Theorem 8.5.4	320
	8.10.5 Proof of Theorem 8.7.2	321
	8.10.6 Proof of Theorem 8.9.1	322
	8.10.7 Proof of Theorem 8.9.2	327
8.11	Concluding Remarks	330
8.12	Bibliographic Notes	331
 Part II Stochastic Game Theory and Information Structures		
9	An Introduction to Stochastic Game Theory	335
9.1	Introduction	335
9.2	A General Framework for Static Stochastic Games	335

- 9.3 An Illustrative Example of a Finite Stochastic Game: Variations on the Team Theoretic Examples 342
 - 9.3.1 Multiple Subjective Probabilities 343
 - 9.3.2 Directly Conflicting Goals 347
- 9.4 Concluding Remarks..... 350
- 9.5 Bibliographic Notes..... 350
- 10 Information Structures in Stochastic Dynamic Games and Informational Properties of Equilibria..... 351**
 - 10.1 Information Structures in Stochastic Game Theory 352
 - 10.1.1 A Mathematical Characterization of Information Structures for Stochastic Dynamic Games 352
 - 10.1.2 An Agent-Wise Intrinsic Model for Games with Players Acting Multiple Times 353
 - 10.2 Solution Concepts for Dynamic Games 354
 - 10.2.1 Nash Equilibrium 355
 - 10.2.2 Stackelberg Equilibrium 355
 - 10.2.3 Perfect Bayesian Equilibrium and Backward Induction 356
 - 10.2.4 Markov Perfect Equilibria 358
 - 10.2.5 Time-Consistency Revisited..... 358
 - 10.2.6 Global vs. Stage-Wise Stackelberg Equilibria for Multi-Stage Games 360
 - 10.2.7 Saddle-Point (Nash) Equilibria via Markov Policies Under State Information for Zero-Sum Games 363
 - 10.2.8 Some Additional Equilibrium Refinements 364
 - 10.3 Information Structure Dependent Properties of Equilibria in Game Theory 365
 - 10.3.1 Value of Additional Information Can Be Positive or Negative in Non-Zero Sum Games 365
 - 10.3.2 Salient Information-Structure Dependent Properties... 367
 - 10.3.3 Zero-Sum Games Admit Several Regularity Properties 368
 - 10.3.4 Lack of Equilibrium Equivalence Under Policy-Dependent vs. Policy-Independent Static Reductions 369
 - 10.3.5 Deterministic Games and Information Structures 371
 - 10.4 Linear Quadratic Games and Dependence of Equilibrium Solutions on Information Structures 373
 - 10.4.1 Static LQG Games: Nash Equilibria 373
 - 10.4.2 LQG Multi-Stage Non-zero-Sum Games: Information Expansion May Lead to Non-unique Nash Equilibrium Costs 376

- 10.4.3 Linear Quadratic Games with Subjective Priors 381
- 10.4.4 Stackelberg Equilibria for Static and Dynamic LQG Games and Optimality of Affine Policies 383
- 10.5 Existence of Equilibria: Implications of the Geometric Properties of Information Structures and Strategic Measures Approach 384
- 10.6 Subjective Priors in Dynamic Games Revisited: Common Knowledge, Agreement, Robustness and Learning 387
 - 10.6.1 Common Knowledge 388
 - 10.6.2 Asymptotic Agreement and Common Knowledge with Common Priors but Different Posteriors 389
 - 10.6.3 Inconsistent Priors, Merging with Increasing Information, and Bayesian and Non-Bayesian Subjective Equilibria 390
- 10.7 Concluding Remarks 392
- 10.8 Bibliographic Notes 392
- 11 Information Dependent Properties of Equilibria: Existence, Comparison, Continuity and Team-against-Team Games 395**
 - 11.1 Introduction 395
 - 11.2 Stochastic Zero-Sum Games and Existence of Saddle-Point Equilibria 397
 - 11.2.1 Existence of Saddle-Point Equilibria (SPE) 399
 - 11.3 Comparison of Information Structures in Zero-Sum Games 400
 - 11.4 Continuity Properties of Equilibrium Values in the Space of Information Structures 402
 - 11.5 General Games: Lack of Continuity and Lack of Positive Value of Information 404
 - 11.6 Team Against Team Games: Zero-Sum, General and Exchangeable Games, and Role of Common Randomness 405
 - 11.6.1 Zero-Sum Games Between Teams of Finitely Many Agents: Existence of Equilibria and Comparison and Regularity in Information 406
 - 11.6.2 Non-Zero-Sum Games Between Teams of Finitely Many Agents: Existence of Equilibrium 409
 - 11.6.3 Nash Equilibria for Exchangeable Team Against Team Games, the Mean-Field Limit and Absence of Common Randomness 410
 - 11.7 Appendix: Proofs 418
 - 11.7.1 Proof of Theorem 11.2.1 418
 - 11.7.2 Proof of Theorem 11.2.2 420
 - 11.7.3 Proof of Lemma 11.3.1 421
 - 11.7.4 Proof of Lemma 11.3.2 422
 - 11.7.5 Proof of Theorem 11.4.1 424

- 11.7.6 Proof of Theorem 11.4.2 425
- 11.7.7 Proof of Theorem 11.4.3 428
- 11.7.8 Proof of Theorem 11.4.4 428
- 11.7.9 Proof of Theorem 11.6.1 429
- 11.7.10 Proof of Theorem 11.6.3 430
- 11.7.11 Proof of Theorem 11.6.6 431
- 11.7.12 Proof of Theorem 11.6.7 433
- 11.8 Concluding Remarks..... 434
- 11.9 Bibliographic Notes..... 434

Part III Information Design via Information Theory in Networked Systems and Control

- 12 Coding for Control and Connections with Information Theory..... 439**
 - 12.1 Introduction 439
 - 12.2 Quantization and Real-Time Coding..... 440
 - 12.2.1 Real-Time Coding 440
 - 12.2.2 Information Structures for Real-Time Encoders and Controllers: Policies, Actions and Measurability 441
 - 12.3 Information Theoretic Preliminaries and Performance of Quantizers 446
 - 12.3.1 Information Theoretic Preliminaries 446
 - 12.3.2 Fixed or Variable Rates of a Quantizer/Encoder 448
 - 12.3.3 Rate-Distortion Theory 449
 - 12.3.4 Channel Coding and Shannon Capacity 450
 - 12.4 Two Salient Aspects where *Information Theory in Control* Differs from *Classical Shannon Theory* 452
 - 12.4.1 Infinite-Dimensional Coding vs. Finite-Dimensional Coding 452
 - 12.4.2 Coding for Stable vs. Unstable Sources 454
 - 12.4.3 Adaptive vs. Non-adaptive Coding, Uniform Distribution Approximation and Design Considerations 456
 - 12.5 A Fundamental Bound on Information Rates for Real-Time Stabilization over Noiseless Channels..... 457
 - 12.6 Appendix: Proofs 458
 - 12.6.1 Proof of Theorem 12.5.1 458
 - 12.7 Concluding Remarks..... 460
 - 12.8 Bibliographic Notes..... 461
- 13 Stochastic Stabilization of Linear Systems Under Information Constraints 463**
 - 13.1 Introduction 463
 - 13.2 The Case of Discrete Noiseless Channels 466
 - 13.2.1 Stochastic Stability Analysis for a Scalar System..... 467

13.2.2	The Multi-Dimensional Case.....	472
13.2.3	The Partially Observed Case	475
13.2.4	Multi-Sensor and Multi-Controller Case: Decentralized Stabilization	476
13.3	Stabilization of Linear Systems Over Noisy Channels: Fundamental Bounds	484
13.3.1	Control and Communication Model	484
13.3.2	Converse Theorem on Stochastic Stability Over a Discrete Memoryless Channel (DMC)	485
13.4	Stochastic Stabilization Over Erasure Channels with Feedback..	487
13.4.1	Connections with Random-Time Drift Criteria	491
13.4.2	Simulation.....	492
13.5	Stochastic Stabilization Over General DMCs with Feedback	493
13.6	Channels with Memory and Multi-Dimensional Sources	497
13.7	Stabilization with Noisy Forward and Feedback/Reverse Channels: When Adaptation Is Infeasible	499
13.8	Appendix: Proofs	500
13.8.1	Proof of Theorem 13.2.1	500
13.8.2	Proof of Theorem 13.2.3	501
13.8.3	Proof of Theorem 13.2.5	503
13.8.4	Proof of Theorem 13.2.8	505
13.8.5	Proof of Theorem 13.6.2	506
13.8.6	Proof of Lemma 13.6.1	508
13.8.7	Proof of Theorem 13.6.3	511
13.8.8	Proof of Proposition 13.4.1	512
13.9	Concluding Remarks.....	514
13.10	Bibliographic Notes.....	515
14	Stabilization of Non-Linear Systems: Information Theoretic vs. Geometric Analysis	517
14.1	Introduction	518
14.2	Information Theoretic Approach to Arriving at Lower Bounds and Fundamental Limits	519
14.2.1	Sublinear Entropy Growth and a Generalization of Bode’s Integral Formula for Non-linear Systems	519
14.2.2	A Generalization of Bode’s Integral Formula for Non-linear Systems	521
14.2.3	Asymptotic Mean Stationarity and Ergodicity	522
14.2.4	Stationarity and Positive Harris Recurrence Under Structured (Stationary) Policies	523
14.3	Stochastic and Geometric Volume Growth Approach.....	525
14.3.1	Some Intuition via a Visit to the Deterministic Setup ..	525
14.3.2	A New Concept: Stabilization Entropy	528
14.3.3	Volume-Expanding Systems	530

- 14.3.4 The Noisy Channel Case 532
- 14.4 Case with Ergodicity via the Stochastic Geometric Method 534
 - 14.4.1 Proof Ideas Under the Ergodicity Criterion 535
 - 14.4.2 A Utility of Geometric Analysis: Refined
Bounds via Coordinate Splitting 538
- 14.5 Achievability: Discrete Noiseless Channels and a
Stationary and Ergodic Construction 542
- 14.6 Discussion: Comparison of the Information Theoretic
and Stochastic Growth (Dynamical Systems) Approaches 544
- 14.7 Appendix: Proofs 546
 - 14.7.1 Proof of Theorem 14.2.1 546
 - 14.7.2 Proof of Lemma 14.2.1 549
 - 14.7.3 Proof of Theorem 14.2.5 554
 - 14.7.4 Proof of Lemma 14.3.1 557
- 14.8 Concluding Remarks 558
 - 14.8.1 Comparison with Invariance Entropy and
Deterministic Non-linear Systems Controlled
Over Noiseless Channels 559
 - 14.8.2 Some Open Directions on Stationary Coding
and Control Policies and Information Theory 560
- 14.9 Bibliographic Notes 560
- 15 Optimal Real-Time Coding and Control Under Information
Constraints: Structural and Existence Results 563**
 - 15.1 Introduction 563
 - 15.2 Policies and Action Spaces for Encoding 564
 - 15.3 Optimal Causal Coding of a Markov Source 566
 - 15.3.1 Single-Terminal, Fully Observed Case 566
 - 15.3.2 Partially Observed Markov Source 568
 - 15.3.3 Structural Results for Systems with Control 571
 - 15.3.4 Case with Noisy Channels and Noiseless Feedback ... 571
 - 15.4 Controlled Markov Optimality Analysis: Existence of
Optimal Zero-Delay Quantizers and a Topology on Quantizers.. 572
 - 15.5 Infinite Horizon Setup and Optimality of Stationary
Codes and Near Optimality of Finite Memory Codes 575
 - 15.5.1 Finite Markov Sources 575
 - 15.5.2 Linear Markov Sources 579
 - 15.6 Multi-Terminal (Decentralized) Setting 582
 - 15.6.1 Memoryless Sources 582
 - 15.6.2 Markov Sources and a Counterexample Under
Signaling 583
 - 15.7 Appendix: Proofs 585
 - 15.7.1 Proof of Theorem 15.3.1 585
 - 15.7.2 Proof of Theorem 15.3.2 587
 - 15.7.3 Proof of Theorem 15.3.6 589

- 15.7.4 Proof of Theorem 15.4.2 590
- 15.8 Concluding Remarks..... 594
- 15.9 Bibliographic Notes..... 594
- 16 Coding and Control for Linear Gaussian Systems:
Separation, Optimality and Linearity 597**
- 16.1 Simultaneous Design of Coding and Control Over
Discrete Channels: Optimality of Separation of Coding
and Control..... 598
 - 16.1.1 Optimal LQG Coding and Control for the
Fully Observed Case: Optimality of Predictive
Quantizers Separating Quantization and Control..... 599
 - 16.1.2 Partially Observed Setup: Separation of
Estimation, Quantization and Control 608
- 16.2 The Gaussian Channel Case: Optimality of Linear
Coding and Control Policies 611
 - 16.2.1 Optimality of Linear Coding Policies Over a
Gaussian Channel with Matching Between
the Source and the Channel 612
 - 16.2.2 Optimization of Coding and Control for
Linear Systems Controlled Over Noisy Channels..... 615
 - 16.2.3 An Entropy-Power Based Bound for Ergodic
Cost Under General Noisy Channels and
Tightness for the Gaussian Setup 619
 - 16.2.4 Multi-Dimensional System Case and a
Dynamic Program for Optimal Linear Policies 620
 - 16.2.5 Structural Results for Optimal Encoders for
Partially Observed LQG Problems 621
 - 16.2.6 Revisiting Stabilization: Sufficiency of
Shannon Capacity Bounds for Gaussian Channels..... 621
- 16.3 General Information Theoretic Bounds and Methods..... 624
 - 16.3.1 Information Theoretic Methods and Design
for Upper and Lower Bounds 624
 - 16.3.2 Looseness of Information Theoretic
Bounds for Multi-Terminal Models and
Counterexamples on Suboptimality of Linear
Policies 626
- 16.4 An Asymptotically Rate-Optimal Two-Part Stationary
Coding Scheme 631
- 16.5 Appendix: Proofs 635
 - 16.5.1 Proof of Theorem 16.1.2 635
 - 16.5.2 Proof of Theorem 16.2.5 638
 - 16.5.3 Proof of Theorem 16.2.6 639
- 16.6 Concluding Remarks..... 642
- 16.7 Bibliographic Notes..... 642

17 Optimal Coding Under Small Estimation Error and Observability Criteria 645

17.1 Introduction 646

17.2 Coding for Deterministic Dynamical Systems in the Small Error Regime 647

 17.2.1 The Problem 647

 17.2.2 Eventual Almost Sure Estimation 649

 17.2.3 Asymptotic Almost Sure Estimation 649

 17.2.4 Asymptotic Estimation in Expectation 651

17.3 Coding for Stochastic Dynamical Systems in the Small Error Regime 652

 17.3.1 A Dynamical Systems Approach 653

 17.3.2 An Information-Theoretic and a Probability-Theoretic Approach 656

 17.3.3 An Achievability Result: Coding of Deterministic Dynamical Systems Over Noisy Communication Channels 657

17.4 Discussion, Comparison and a Review of Related Results in the Literature 658

 17.4.1 Deterministic Model 658

 17.4.2 Stochastic Model 660

17.5 Concluding Remarks 661

17.6 Bibliographic Notes 661

Part IV Information and Signaling Games

18 Signaling Games: Equilibria Under Criteria or Prior Mismatch Between an Encoder and a Decoder 665

18.1 Introduction to Signaling Games: Nash and Stackelberg Solution Concepts, and Some Basic Results 666

18.2 Quadratic Signaling Games 671

 18.2.1 Quadratic Signaling Game with Costless Communication (Cheap Talk) 671

 18.2.2 Multi-Dimensional Cheap Talk 681

 18.2.3 Quadratic Cheap Talk: Stackelberg Equilibria (Bayesian Persuasion) 683

 18.2.4 Quadratic Signaling Game: Noisy Channel Case 684

18.3 Signaling Games with Subjective Priors Between an Encoder and a Decoder 685

 18.3.1 The Commitment Setup with Quadratic Costs: Stackelberg Equilibria 686

 18.3.2 The Non-cooperative/No-commitment Setup (Nash Equilibria) 695

 18.3.3 Signaling Under Subjective Priors Without Channel Noise: Cheap Talk Revisited 696

18.4 Concluding Remarks..... 697

18.5 Bibliographic Notes..... 698

19 Two Classes of Signaling Games: Reputation and Information Privacy Games 701

19.1 Reputation Games as Signaling Games..... 702

 19.1.1 The Model..... 702

 19.1.2 Optimal Strategies and Equilibrium Behavior..... 704

 19.1.3 Undiscounted Time-Average Payoff Case and an Upper Payoff Bound for the Arbitrarily Patient Long-lived Player..... 711

19.2 Information Privacy Games 714

 19.2.1 Nash Equilibria 716

 19.2.2 Stackelberg Equilibria 721

 19.2.3 The MMSE Gaussian Information Bottleneck Problem 723

19.3 Concluding Remarks..... 729

19.4 Bibliographic Notes..... 730

20 Zero-Sum Information Games: Jamming and Witsenhausen’s Counterexample as a Game 733

20.1 Witsenhausen’s Counterexample as a Zero-Sum Signaling Game 733

 20.1.1 Zero-Sum Signaling Games 734

 20.1.2 Witsenhausen’s Counterexample as a Zero-Sum Signaling Game and Optimality of Linear Policies 734

20.2 Jamming Games 736

 20.2.1 Equilibrium Solutions 739

 20.2.2 The Case with a One-Way Channel [86] and a Discussion..... 742

20.3 Appendix: Proofs 743

 20.3.1 Proof of Theorem 20.2.1 743

20.4 Concluding Remarks..... 748

20.5 Bibliographic Notes..... 748

Part V Reinforcement Learning Under Full, Partial, Decentralized, and Networked Information

21 Reinforcement Learning for Fully Observable Models: Convergence to Near Optimality under Weak Feller Continuity 753

21.1 Introduction 753

21.2 Classical Q-Learning for Finite Models and Full State Information..... 754

 21.2.1 Classical Q-Learning for Finite Fully Observed Models 755

- 21.3 Near Optimal Planning for Continuous Space MDPs:
Near Optimality of Finite Models under Weak Feller
Continuity 759
 - 21.3.1 Finite Action Approximation to MDPs 759
 - 21.3.2 Finite State Approximation to MDPs 762
- 21.4 Reinforcement Learning for Continuous Space Models
under weak Feller Continuity: Quantized Q-Learning 764
- 21.5 Concluding Remarks..... 768
- 21.6 Bibliographic Notes..... 768
- 22 Reinforcement Learning for Partially Observable Models 769**
 - 22.1 Introduction 769
 - 22.2 Reinforcement Learning for the Partially Observed
Setup I: Near Optimality of Quantized Policies Under
Weak Feller Property of Non-linear Filters 770
 - 22.3 Reinforcement Learning for the Partially Observed
Setup II: Near-Optimality of Finite Window Policies 771
 - 22.3.1 Near-Optimality of Finite Window Policies
Under Filter Stability 771
 - 22.3.2 An Alternative Finite Window Belief-MDP
Reduction and Its Approximation 771
 - 22.3.3 Q-Learning Convergence of Finite Window Policies .. 775
 - 22.4 Concluding Remarks..... 780
 - 22.5 Bibliographic Notes..... 780
- 23 Multi-Agent Reinforcement Learning with General
Information Structures: Convergence to Equilibria 783**
 - 23.1 Introduction 783
 - 23.2 A General Paradigm for Multi-Agent Learning: Two
Time Scales and a Markov Chain over Play Path Graphs 785
 - 23.3 Learning with Local Information for Stochastic Teams
and Weakly Acyclic Games with Deterministic Baseline
Policies 787
 - 23.3.1 Stochastic Teams and Weakly Acyclic Games
with Local Action and Global State Information 787
 - 23.3.2 Weakly Acyclic Games 789
 - 23.4 Satisficing Paths and Independent Multi-Agent
Reinforcement Learning with Possibly Randomized Policies 795
 - 23.5 Subjective Reinforcement Learning for Mean-Field
Problems with General Information Structures..... 798
 - 23.5.1 N -Player Weakly Coupled (Mean-Field) Games..... 798
 - 23.5.2 On the Observation Channel 802
 - 23.5.3 Stationary Equilibrium Policies: Existence Results 802
 - 23.5.4 Existence of Stationary Equilibrium Under
Global State Observability 803

23.5.5	Existence of Stationary Equilibrium Under Mean-Field State Observability	804
23.5.6	Convergence of Subjective Single-Agent Learning Under Stationary Policies	807
23.5.7	Existence of Subjective Equilibrium Under Global or Mean-Field State Observability	809
23.5.8	Equilibria Under Compressed Mean-Field State and Local State Information	810
23.6	Subjective Satisficing, Policy Update Dynamics and Subjective Q-Learning	810
23.6.1	Naively Learned Subjective Functions and ϵ -Satisficing	812
23.6.2	Quantization of the Policy Space	815
23.6.3	Learning Algorithm and Convergence to Subjective Q-Learning Equilibria	818
23.7	Concluding Remarks	822
23.8	Bibliographic Notes	822
24	Reinforcement Learning for Networked Control and Optimal Quantization	825
24.1	Introduction	826
24.2	Zero-Delay Source-Channel Coding as an MDP	827
24.2.1	Topology on Quantizers	829
24.2.2	Predictor and Filter Stability and Unique Ergodicity Under an Exploration Policy	830
24.3	Algorithms	831
24.3.1	Quantizing π_t	832
24.3.2	Quantized Q-Learning for Source-Channel Coding ...	832
24.3.3	Connection to the Average Cost Problem	833
24.3.4	Simulations	834
24.4	Concluding Remarks	835
24.4.1	Controlled Models	836
24.5	Bibliographic Notes	836
A	Topological Notions and Optimization	837
A.1	Sets	837
A.2	Linear, Metric, Hilbert, Banach and Standard Borel Spaces	838
A.3	Matrices	841
A.4	Convex Sets and Functionals	842
A.5	Optimization of Functionals	843
A.6	Contraction Mappings and Fixed-Point Theorems	844
B	Probability Theory and Stochastic Processes	847
B.1	Probability	847
B.1.1	Measurable Spaces	847

- B.1.2 Integration 848
- B.1.3 Probability Spaces and Random Variables 850
- B.2 Conditional Expectation and Estimation 851
- B.3 Stochastic Processes 853

- C On Spaces of Probability Measures** 855
- C.1 Convergence of Sequences of Probability Measures 855
- C.2 Some Measurability Results on Spaces of Probability Measures 857
- C.3 A Generalized Dominated Convergence Theorem 858
- C.4 The w - s Topology 858
- C.5 Lusin’s Theorem 859

- D Markov Chains, Martingales and Ergodic Processes** 861
- D.1 Markov Chains 861
- D.2 Discrete-Time Martingales 864
- D.3 Stochastic Stability of Dynamical Systems and Random
Processes 865
- D.3.1 Stationary, Ergodic, and Asymptotically Mean
Stationary Processes 865

- E Markov Decision Theory and Optimality of Markov Policies** 867
- E.1 Controlled Markov Models 867
- E.2 Fully Observed Markov Control Problem Model 867
- E.2.1 Classes of Control Policies 868
- E.2.2 Optimality of Markov Policies and
Elimination of Irrelevant Information 869
- E.2.3 Markov Decision Processes (MDPs) and
Optimality of Markov Policies 869
- E.2.4 Dynamic Programming and Measurable
Selection Criteria 869
- E.2.5 Discounted Cost Criterion 871
- E.3 Partially Observable MDPs (POMDPs) 873
- E.4 Kalman Filter and the Linear Quadratic Gaussian
Optimal Control Problem 875

- F Relaxed Control Topologies** 879
- F.1 Young Topology on Control Policies 879
- F.2 Borkar (Weak*) Topology on Control Policies 880
- F.3 Some Properties of Young and Borkar Topologies 881

- G Stochastic Stability of Markov Chains and Random-Time
State-Dependent Stochastic Drift** 883
- G.1 Stochastic Stability and Drift Criteria 883
- G.1.1 One-Stage Foster-Lyapunov Drift Criteria 884
- G.1.2 State-Dependent Drift Criteria 884

G.1.3	Random-Time State-Dependent Stochastic Drift Criteria	885
H	Topological and Metric Entropy in Dynamical Systems	887
	References	891
	Index	925

Acronyms and Notations

$\text{trace}(A)$	Trace of a square matrix A
$\det(A)$	Determinant of a square matrix A
1_E	Indicator function for event E
\mathbb{X}	A generic vector space
$\langle x, y \rangle$	Inner product between x and y on a Hilbert space
$\mathcal{B}(\mathbb{X})$	The Borel σ -field on \mathbb{X}
$\sigma(y)$	σ -field generated by a random variable y
$\mathcal{P}(\mathbb{X})$	Set of probability measures on $\mathcal{B}(\mathbb{X})$
\mathbb{R}	Set of real numbers
\mathbb{R}^n	Vector space of n -dimensional real vectors
\mathbb{Z}	Set of integers
\mathbb{Z}_+	Set of non-negative integers
\mathbb{N}	Set of positive integers
\mathcal{Q}	A space of quantizers
Q	Quantizer or channel depending on context
$\Pi^{comp,i}$	Composite quantization policy for encoder i
Q_t^i	Quantizer used by Agent i at time t
DMi	Decision Maker i
Ai	Agent i
$\underline{\gamma}^i$	Policy of DM i , that is $\{\gamma_t^i, t \geq 0\}$
$\underline{\gamma}$	Ensemble of policies for all decision makers, that is $\{\underline{\gamma}^i\}$
I_t^i or \mathcal{I}_t^i	Information variable at Agent i at time t
$\underline{\eta}$	Information structure inducing map $\{\eta^1, \dots, \eta^N\}$
$\underline{E}_P^\gamma\{\cdot\}$	Expectation under policy $\underline{\gamma}$, with initial condition measure P
E_x	Expectation conditioned on an initial condition realization x , or with respect to a random variable x , depending on the context
$Df(x)$ or $J(f(x))$	Jacobian of f at x
$H(\cdot)$	Discrete entropy
$h(\cdot)$	Differential entropy

$I(\cdot; \cdot)$	Mutual information
$D(P_1 P_2)$	Kullback-Leibler Divergence between P_1 and P_2
$ x $	Euclidean norm of a finite-dimensional real-valued vector x
A' or A^T	transpose of matrix A
$ S $	Cardinality of a set S
$A \setminus B$	Ordered set difference: $\{x : x \in A, x \notin B\}$
$A \Delta B$	$(A \setminus B) \cup (B \setminus A)$
$\ln(x)$ or $\log(x)$	Natural logarithm of positive real x
$\underline{0}$	Zero vector
$\overline{\mathcal{T}}$	Time/stage index set, $\{1, 2, \dots, T\}$ or $\{0, 1, \dots, T - 1\}$
$\mathcal{N}(\mathcal{L})$	Decision maker (DM) index set, $\{1, 2, \dots, N\}$ ($\{1, 2, \dots, L\}$)
$u_{[k,s]}$	action (decision) variables from $t = k$ to $t = s$ for $s > k$, $\{u_k, u_{k+1}, \dots, u_s\}$
\mathbf{u}	Collection of actions in \mathcal{N} (or \mathcal{L}): $\{u^1, u^2, \dots, u^N\}$