

UNIVERSITY OF CRETE
DEPARTMENT OF COMPUTER SCIENCE
FACULTY OF SCIENCES AND ENGINEERING

Study on the Effect of Misinformation in Multi-Agent Systems

by

Constantinos Varsos

PhD Dissertation

Presented

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

Heraklion, February 25, 2022

Acknowledgments

There are many people whom I would like to thank for their contribution and support.

I would first like to thank my supervisor Professor Dimitrios Plexousakis that support me during these years. His eagerness to discuss my concerns, ideas, and thoughts, alongside with his advices provided me with the necessary support to fulfill my research.

Further, I express my appreciation for my advisor Professor Ioannis Tsamardinos for his will to discuss any issue that arose in my endeavor, his insights, and constructive ideas helped me to clarify and crystallize many aspects of my research.

Next, I am indebted to my advisor, Dr. Giorgos Flouris. He had the most decisive contribution in this study; his insightful guidance, inspired advices and constructive tenacity played a major role in this dissertation.

However, a great deal of this work would not happen without the assistance, guidance, and advices of Dr. Marina Bitsaki and Dr. Michail Fasoulakis. Further, I owe them a debt of gratitude for sharing with me some of their deep insights into the game theory.

Additionally, I would like to acknowledge the patient assistance of Christos Papadimitriou. Our discussions helped me to clarify many aspects of Probability Theory. Moreover, I would like to thank Dr. Theodore Patkos and Philippos Goudis for their advices, suggestions and ideas in the field of Knowledge Representation and Logical Programming.

This work was supported by a graduate fellowship from the Institute of Computer Science (ICS) of the Foundation for Research and Technology Hellas (FORTH), and the Special Research Fund Account (SARG) of the University of Crete, which I want to acknowledge for providing financial assistance and technical equipment. Furthermore, I wish to thank the Information Systems Laboratory of ICS that hosted me. In addition, I wish to acknowledge the help and support from the secretariat and the administrators of Computer Science Department and ICS-FORTH.

Nevertheless, a PhD endeavor is hardly ever segregated from everyday life. I feel fortunate for that many friends and colleagues that perseveringly provided me with enduring support, continuous encouragement and relaxation time, that balanced out the mental challenges that I faced towards completing my thesis. I will not mention their names, they know. I owe them a lot.

Finally, this thesis can be dedicated to none other but my beloved family, the warm home that instilled in me the love of knowledge.

The limits of my language mean the limits of my world.
-*Ludwig Wittgenstein: Tractatus Logico-Philosophicus*

Abstract

Industrial processes and research activities quite often involve interactions between self-interested participants. Game theory is a standard tool to analyze and study these interactions, but usually comes along with the assumption that the participants (i.e. agents, players) have a common and correct (albeit not always complete) knowledge with regards to the abstract formulation of the interaction. However, in many real-world situations, it could be the case that (some of) the agents are misinformed with regards to the game that they play, essentially having an *incorrect* understanding of the setting, without being aware of it. This would invalidate the common knowledge assumption. To study this phenomenon in this dissertation we establish a new framework.

We initiate our study by presenting a new game-theoretic framework, called *misinformation games*, that provides the formal machinery necessary to study this phenomenon, and present some basic results regarding its properties. Interestingly, the new concept provides new equilibrium concepts, related to the Nash equilibrium. Thereupon, we introduce a new metric, called *Price of Misinformation*, in order to quantify the influence of misinformation in the efficiency of the interaction. Furthermore, we apply our framework in a variety of well-known classes of games.

Afterwards, we expand the misinformation game model, by developing a discrete-time iterative procedure, where in each time step each agent chooses an action according to the (possibly erroneous) game specification that she possesses. Then, the actual payoffs that correspond to the agglomeration of the agents' choices are publicly announced, thus allowing agents to update their information. Consequently, agents may re-evaluate their behaviour in the next time step. We call this process *Adaptation Procedure*, and we provide various results regarding its properties. Further, we present a complete analysis of the behaviour of the agents as their game specifications are updated, and show that this leads to new equilibrium concepts.

Thereafter, we enrich the Adaptation Procedure by incorporating the epistemic view that each agent has regarding the interaction. Towards this direction, we formally define the epistemic perspective of Adaptation Procedure in misinformation games. Namely, we construct a process, called *Epistemic Adaptive Evolution*, where agents revise both their information and their epistemic knowledge according to the game they play. This also provides new equilibrium concepts. With this at hand, we complete our framework, through which we can study the phenomenon of agent interaction with incorrect information.

Evidently, in several cases in our model, it is necessary to compute several equilibrium

concepts. For that, we introduce a novel online learning algorithm. Specifically, we propose a novel variant of the multiplicative weights update method using best-response strategies, that guarantees last-iterate convergence for zero-sum games with a unique Nash equilibrium.

Next, we consider the case of misinformation games where the misinformation is due to random noise that additively distorts the payoff matrices of the agents (e.g., due to communication errors). We call this setting *noisy games*. We analyze the general properties of two-players noisy games and we derive theoretical formulas which determine the probability that the noise will significantly affect the strategic behaviour of the agents, based on the noise intensity and pattern.

Following the analysis and study of interaction from the perspective of the participants, we approach the problem from the perspective of the game’s designer. In particular, we introduce a novel approach for Coordination mechanisms in games, based on the idea of misinforming agents about the game formulation, in order to steer them towards a behaviour that leads to an improved outcome in terms of social welfare. We propose a mechanism that provides the agents with the right incentives to adopt a socially optimal behaviour by misinforming them.

Keywords: Misinformation games, Adaptation Procedure, Epistemic Adaptive Evolution, natural misinformed equilibrium, stable misinformed equilibrium, Epistemic games, epistemic natural equilibrium, stable epistemic natural equilibrium.

Supervisor: Dimitrios Plexousakis
Professor
Computer Science Department
University of Crete

Περίληψη

Συχνά οι βιομηχανικές διεργασίες και οι ερευνητικές δραστηριότητες περιλαμβάνουν αλληλεπιδράσεις μεταξύ συμμετεχόντων (ήτοι πράκτορες, παίκτες). Η Θεωρία Παιγνίων αποτελεί, συχνά, το εργαλείο με το οποίο μελετάμε τέτοιου είδους αλληλεπιδράσεις· συνήθως περιλαμβάνοντας την υπόθεση ότι οι πράκτορες έχουν κοινή και σωστή (χωρίς να είναι πάντοτε πλήρης) πληροφορία σχετικά με την δραστηριότητα στην οποία συμμετέχουν. Ωστόσο, στον πραγματικό κόσμο οι πράκτορες συχνά ενεργούν έχοντας λανθασμένη πληροφόρηση σχετικά με την αλληλεπίδραση και τις συνθήκες της, χωρίς να το γνωρίζουν· ακυρώνοντας κατ' αυτόν τον τρόπο την υπόθεση της κοινής και σωστής πληροφόρησης. Για να μελετήσουμε αυτό το φαινόμενο, στην παρούσα διατριβή θεμελιώνουμε μια καινούργια δομή.

Αρχικά, εισάγουμε στην δομή μας μια καινούργια κλάση παιγνίων, την οποία ονομάζουμε *παίγνια εσφαλμένης πληροφορίας*, η οποία μας παρέχει όλα τα αναγκαία θεωρητικά εργαλεία ώστε να μελετήσουμε το φαινόμενο της αλληλεπίδρασης υπό εσφαλμένη πληροφορία. Αξιοσημείωτο είναι ότι αυτή η κλάση παιγνίων παρέχει καινούργια σημεία ισορροπίας, τα οποία βασίζονται στα σημεία ισορροπίας Nash. Παράλληλα, παρουσιάζουμε βασικές ιδιότητες της δομής αυτής και την εφαρμόζουμε σε διάφορες κλασσικές κλάσεις παιγνίων. Ταυτοχρόνως, ορίζουμε μια νέα μετρική η οποία ποσοτικοποιεί την επίδραση της εσφαλμένης πληροφορία στην απόδοση της αλληλεπίδρασης.

Εν συνεχεία, επεκτείνουμε το μοντέλο των παιγνίων εσφαλμένης πληροφορίας, αναπτύσσοντας μια επαναληπτική διαδικασία διακριτού χρόνου, όπου σε κάθε χρονικό βήμα κάθε πράκτορας διαλέγει μια ενέργεια σύμφωνα με τις (πιθανώς εσφαλμένες) προδιαγραφές της αλληλεπίδρασης που κατέχει. Κατόπιν, ανακοινώνονται δημόσια οι πραγματικές ανταμοιβές που λαμβάνουν οι πράκτορες, σύμφωνα με τον συνδυασμό των επιλογών που έχουν κάνει· ενημερώνοντας κατ' αυτόν τον τρόπο την πληροφορία που κατέχουν. Ονομάζουμε αυτή την διαδικασία *Διαδικασία Προσαρμογής* και παρουσιάζουμε θεωρητικά αποτελέσματα σχετικά με τις ιδιότητες της. Περαιτέρω, αναλύουμε την συμπεριφοράς των πρακτόρων καθώς η πληροφορία που κατέχουν ενημερώνεται, και αποδεικνύουμε ότι αυτή οδηγεί σε καινούργια σημεία ισορροπίας.

Έπειτα, διευρύνουμε την Διαδικασία Προσαρμογής ενσωματώνοντας την γνωσιολογική θεώρηση που έχει ο κάθε πράκτορας σχετικά με την αλληλεπίδραση στην οποία συμμετέχει. Προς αυτή την κατεύθυνση, ορίζουμε τυπικά την γνωσιολογική θεώρηση της Διαδικασίας Προσαρμογής στα παίγνια εσφαλμένης πληροφορίας. Συγκεκριμένα, κατασκευάζουμε μια διαδικασία, την οποία ονομάζουμε *Γνωσιολογική Προσαρμοστική Εξέλιξη*, κατά την οποία οι πράκτορες αναθεωρούν τόσο την πληροφορία που κατέχουν όσο και την γνωσιολογική πληροφορία σχετικά με την αλληλεπίδραση στην οποία συμμετέχουν. Όπως και στα προηγούμενα σκέλη του μοντέλου μας,

αναδύονται καινούργια σημεία ισορροπίας. Με αυτό το μέρος της μελέτης, ολοκληρώνουμε την δομή μέσω της οποίας μπορούμε να μελετήσουμε το φαινόμενο της αλληλεπίδρασης πρακτόρων με εσφαλμένη πληροφορία.

Γίνεται άμεσα αντιληπτό ότι είναι σημαντικός ο υπολογισμός σημείων ισορροπίας (π.χ. Nash) σε διάφορα σημεία του μοντέλου μας. Για τον λόγο αυτό, αναπτύσσουμε έναν καινούργιο on-line αλγόριθμο μάθησης. Πιο συγκεκριμένα, προτείνουμε μια παραλλαγή της multiplicative weights update μεθόδου, χρησιμοποιώντας best-response στρατηγικές, η οποία εγγυάται κατά σημείο σύγκλιση για παίγνια μηδενικού αθροίσματος με μοναδικό σημείο ισορροπίας Nash.

Ως συνέχεια της μελέτης του φαινομένου της επίδρασης της εσφαλμένης πληροφορίας σε αλληλεπιδράσεις πολλών πρακτόρων, αναλύουμε την επιρροή της δομής της εσφαλμένης πληροφορίας στην αλληλεπίδραση. Συγκεκριμένα, εισάγουμε μια νέα οικογένεια παιγνίων εσφαλμένης πληροφορίας, την οποία ονομάζουμε *παίγνια θορύβου*, όπου η εσφαλμένη πληροφορία εξαρτάται από τυχαίο θόρυβο, ο οποίος προστίθεται στις τιμές των ανταμοιβών. Κατόπιν, αναλύουμε τις γενικές ιδιότητες των παιγνίων θορύβου δύο παιχτών και εξάγουμε θεωρητικά αποτελέσματα ως προς την επίδραση του θορύβου στην στρατηγική συμπεριφορά των πρακτόρων.

Τέλος, η μελέτη ολοκληρώνεται με την ανάλυση της αλληλεπίδρασης από την σκοπιά του σχεδιαστή της. Πιο συγκεκριμένα, παρουσιάζουμε μια νέα προσέγγιση στην περιοχή των Coordination mechanisms, βασιζόμενοι στην ιδέα ότι παρέχοντας εσφαλμένη πληροφορία στους πράκτορες, σχετικά με τις προδιαγραφές της αλληλεπίδρασης, μπορούμε να τους κατευθύνουμε σε συμπεριφορές που βελτιώνουν την απόδοση της αλληλεπίδρασης, με όρους συλλογικής πρόνοιας. Προτείνουμε έναν μηχανισμό που παρέχει στους πράκτορες τα κατάλληλα κίνητρα ώστε να υιοθετήσουν μια συλλογικά καλύτερη συμπεριφορά.

Λέξεις κλειδιά: Παίγνια εσφαλμένης πληροφορίας, Διαδικασία Προσαρμογής, Γνωσιολογική Προσαρμοστική Εξέλιξη, φυσικό σημείο ισορροπίας εσφαλμένης πληροφορίας, ευσταθές φυσικό σημείο ισορροπίας εσφαλμένης πληροφορίας, Γνωσιολογικά Παίγνια, γνωσιολογικό φυσικό σημείο ισορροπίας, ευσταθές γνωσιολογικό φυσικό σημείο ισορροπίας.

Επόπτης: Πλεξουσάκης Δημήτρης
Καθηγητής
Τμήμα Επιστήμης Υπολογιστών
Πανεπιστήμιο Κρήτης

Contents

Acknowledgments	iii
Abstract	v
Περίληψη (Abstract in Greek)	vii
Contents	ix
List of Figures	xiii
List of Tables	xvii
1 Introduction	1
1.1 Modeling misinformation	2
1.2 Research Questions	4
1.3 Outline and Contribution	4
2 Related Work	9
2.1 Misperception in games	9
2.2 Games with distorted payoff matrices	10
2.3 Quantifying socially sub-optimal behaviour	11
2.4 Iterative interactions	12
2.5 Dynamics in games	14
2.6 Structure of misinformation	16
2.7 Coordination mechanism	17
3 Theoretical Background	19
3.1 Normal-form games	19
3.2 Zero-sum games and approximate equilibria	21
3.3 Learning Dynamics	22
3.4 Load-balancing games	24
3.5 Congestion games	25
4 Misinformation Games	27
4.1 Introduction	27
4.2 Normal-form games	28
4.2.1 Equilibria for Misinformation normal-form games	30
4.2.2 Price of Misinformation	32
4.3 Misinformation in load balancing games	33
4.3.1 Equilibria in Misinformation load balancing games	34
4.3.2 Price of Misinformation in load balancing games	35
4.4 Misinformation in non-atomic congestion games	38

5	Epistemic Adaptive Evolution	41
5.1	Introduction	41
5.2	The Adaptation Procedure	42
5.2.1	Informal Description	42
5.2.2	Formal Definition	42
5.2.3	Stabilisation of the Adaptation Procedure	45
5.3	Adaptation Procedure: Visualisations	46
5.3.1	Adaptation Procedure Graph (AP-Graph)	46
5.3.2	Adaptation Procedure Induced Graph (API-Graph)	46
5.4	Adaptation Procedure: Properties	48
5.4.1	General Properties	48
5.4.2	Termination, and existence of smes	48
5.4.3	Estimating the length of the Adaptation Procedure	51
5.5	The Epistemic Adaptive Evolution	53
5.5.1	Informal Description	53
5.5.2	Preliminaries	55
5.5.3	Players' Epistemic game	56
5.5.4	Epistemic Update	57
5.5.5	Epistemic game playing	62
6	Learning Dynamics	65
6.1	Introduction	65
6.1.1	Informal Description	66
6.1.2	A Revealing Example	67
6.2	Forward Looking Best-Response Multiplicative Weights Update Method	67
6.2.1	Definition of the Dynamics	67
6.2.2	Main Results	69
6.3	Numerical Experiments	77
7	Noisy Games	89
7.1	Introduction	89
7.2	Noisy Games	90
7.2.1	Basic Definitions	90
7.2.2	Strategies, strategy profiles and equilibria in misinformation games	91
7.2.3	Behavioural Consistency and ε -closeness	92
7.3	Probabilities for Behavioural Consistency	94
7.3.1	Determining equilibrium strategies	94
7.3.2	Misinformation games	99
7.3.3	Probabilities	102
7.4	Results for Noisy games	107
7.4.1	Effect of modifying tolerance (ε)	108

7.4.2	Effect of changing the game (G^0) and the mean (M)	113
7.4.3	Effect of modifying noise intensity (D)	117
7.5	Discussion and Experiments	124
7.5.1	Theoretical and Experimental Computation of the Probability that a Game is (Inverse-) ε -misinformed	125
7.5.2	Optimal strategy profiles in terms of efficiency	129
7.5.3	PoM vs PoA	130
8	Mechanism Design	133
8.1	Introduction	133
8.2	The waterfilling algorithm	134
8.3	Coordination mechanisms with misinformation	135
9	Conclusions	139
9.1	Research path	139
9.2	Synopsis of Contributions	140
9.2.1	Misinformation Games	140
9.2.2	Adaptation Procedure	140
9.2.3	Epistemic Adaptive Evolution	140
9.2.4	Learning Dynamics	141
9.2.5	Noisy Games	141
9.2.6	Mechanism Design	142
9.3	Directions for Future Work and Research	142
	Bibliography	145
	Appendices	
A	Publications	157
B	Supplementary material	159
B.1	Basic Probability theory	159
B.2	Kullback-Leibler divergence	160
B.3	Auxiliary definitions and results	160
B.4	Jacobian matrix	162
B.5	Numerical experiments in noisy games	163
B.5.1	Case OP_x^G	164
B.5.2	Case OM_x^G	164
B.5.3	Case ROM_x^G	165
	Index	167

List of Figures

1.1	Games with information assumptions: (a) complete/correct, (b) incomplete/correct.	3
1.2	Schematic representation of two players' games with misinformation.	4
1.3	Graphical representation of misinformation in multi-agent systems.	5
1.4	Road map of the thesis.	7
3.1	Graphical representation for the sequence of iterative approximations of MWU (green), OMWUM (blue) and EG (red) from time step $t - 1$ to time step $t + 1$. The dashed red curves represent the intermediate step of EG.	25
4.1	Load balancing game and misinformed load balancing game: (a) optimal assignment, (b) worst Nash equilibrium allocation, (c-d) worst natural misinformed equilibrium allocation.	36
4.2	Non-atomic congestion game with 2 parallel links.	40
5.1	Schematic representation of the functionality of Adaptation Procedure from time step t to time step $t + 1$, that is $\mathcal{AD}^{t+1}(mG)$	45
5.2	AP-Graph of Adaptation Procedure for misinformation game in Example 5.2, with $\{mG^{(1)}, mG^{2a}\} \in \mathcal{AD}^{(1)}(\{mG\})$ and $\{mG^{(2b)}, mG^{(3c)}\} \in \mathcal{AD}^{(2)}(\{mG\})$	47
5.3	API graph in mG 's games for Example 5.2.	47
5.4	Player q revises his/her beliefs.	60
5.5	Revision for each player when information about position VA is publicly announced.	63
5.6	Schematic representation of the epistemic adaptive evolution with the epistemic evolution and the adaptive update of the agents.	64
6.1	Schematic representation of the convergence path of OMWU (red) and FLBR-MWU (black), starting from the uniform point $(\mathbf{x}^0, \mathbf{y}^0)$	68
6.2	A random realization of the learning dynamics for three variants of MWU.	68
6.3	The D_{KL} between the Nash equilibrium and the FLBR-MWU dynamics for two instances and no rescaling of x-axis (upper panels) and with rescaling (lower panels). The relationship between the number of steps and learning rate is inversely proportional.	79

6.4	The value of the game as a function of the number of steps for the three MWU variants.	80
6.5	The dynamics of the update step per coordinate (solid), as well as the IBR step (dashed) for the row player. For the equilibrium strategy x^* , it holds that $\text{supp}(x^*) = \{1, 2, 4, 5, 7, 9\}$. Note that x^t converges to the same support. . . .	81
6.6	Same as Figure 6.5, but in logarithmic scale.	82
6.7	The dynamics of the update step per coordinate (solid), as well as the IBR step (dashed) for the column player. For the equilibrium strategy y^* , it holds that $\text{supp}(y^*) = \{2, 4, 5, 8, 9, 10\}$	83
6.8	Same as Figure 6.7, but in logarithmic scale.	84
6.9	Boxplots for the number of steps until convergence for various values of ξ and two payoff matrix sizes.	85
6.10	Boxplots for the number of steps until convergence for various payoff matrix sizes under FLBR-MWU (left) and OMWU/OMD (right). The computational gains when FLBR-MWU is used are striking.	86
6.11	KL divergence and l_1 norm difference for $t_{\max} = 5 \times 10^6$ and two values for the learning rate: $\eta = 0.1$ (upper row of panels) and $\eta = 0.02$ (lower row of panels).	86
7.1	Behavioral consistency between interaction with real specifications and interaction with misinformation.	93
7.2	Example showing that $\mathcal{P}[mG : \varepsilon\text{-misinformed}]$ (blue), $\mathcal{P}[mG : \text{inverse-}\varepsilon\text{-misinformed}]$ (orange) (vertical axis) have non-monotonic dependence to noise (horizontal axis).	124
7.3	Test cases.	125
7.4	Monte Carlo (MC) simulation and probabilistic formulas for Prisoner's Dilemma and Win-Win games. Vertical axis: (7.4a) $\mathcal{P}[mG : \varepsilon\text{-misinformed}]$, (7.4b) $\mathcal{P}[mG : \text{inverse-}\varepsilon\text{-misinformed}]$. Horizontal axis: noise intensity d	126
7.5	Monte Carlo (MC) simulation and probabilistic formulas for Matching Pennies. Vertical axis: (7.5a) $\mathcal{P}[mG : \varepsilon\text{-misinformed}]$, (7.5b) $\mathcal{P}[mG : \text{inverse-}\varepsilon\text{-misinformed}]$. Horizontal axis: noise intensity d	127
7.6	Monte Carlo (MC) simulation and probabilistic formulas for Battle of the Sexes. Vertical axis: (??) $\mathcal{P}[mG : \varepsilon\text{-misinformed}]$, (??) $\mathcal{P}[mG : \text{inverse-}\varepsilon\text{-misinformed}]$. Horizontal axis: noise intensity d	128
7.7	Strategy profiles in terms of efficiency.	129
7.8	PoM plane for games in Table 7.3.	131
8.1	Two possible cases of waterfilling with different budget of power.	134
8.2	Coordination mechanism with misinformation.	138
9.1	Top view of the structure of misinformation games model.	139

B.1	Illustration of water-filling algorithm. The height of each patch is given by α_i . The region is flooded to a level $1/nu^*$ which uses a total quantity of water equal to one. The height of the water (shown shaded) above each patch is the optimal value of x_i^*	161
B.2	Cases OP_x^G : (a) Prisoners' Dilemma, (b) Matching Pennies, and (c) Battle of the Sexes.	164
B.3	Cases OM_x^G : (a) Prisoners' Dilemma, (b) Matching Pennies, and (c) Battle of the Sexes.	164
B.4	Cases ROM_x^G : (a) Prisoners' Dilemma, (b) Matching Pennies, and (c) Battle of the Sexes.	165

List of Tables

4.1	Payoff matrices for the PD and mPD.	27
6.1	Statistics on the number of steps till convergence for various values of ξ and n . The maximum number of steps was set to $t_{\max} = 10^6$	78
6.2	Statistics on the number of steps till convergence for various sizes of the game. The maximum number of steps was set to $t_{\max} = 5 \times 10^6$	79
6.3	Quantile statistics on the number of steps till convergence for various values of ξ and n . The maximum number of steps was set to $t_{\max} = 2 \times 10^6$	85
6.4	Statistics on the number of steps till convergence for OWMU, OMD and FLBR- MWU and various payoff matrix sizes. The maximum number of steps was set to $t_{\max} = 5 \times 10^6$	87
7.1	Visualising the cases of Proposition 22, for $x = r$	98
7.2	Scenarios for ε -misinformed and inverse- ε -misinformed	102
7.3	Formulas related to $U(y, x, i)$ for a given $mG \sim G^0 + \mathcal{N}(M, D)$	103
7.4	Various probabilities pertaining to a given $mG \sim G^0 + \mathcal{N}(M, D)$ (see also Proposition 25).	105
7.5	Probabilities for ε -misinformed and inverse- ε -misinformed ($\mathcal{P}_r^{mis} \cdot \mathcal{P}_c^{mis}$ and $\mathcal{P}_r^{inv} \cdot \mathcal{P}_c^{inv}$ respectively – see also Theorems 7, 8)	107
7.6	Effect of tolerance on behavioural consistency (monotonicity)	108
7.7	Minimal and maximal values for the probabilities of mG being (inverse-) ε - misinformed (resulting by multiplying \mathcal{P}_r^{mis} with \mathcal{P}_c^{mis} and \mathcal{P}_r^{inv} with \mathcal{P}_c^{inv} respectively)	112

List of Algorithms

1	Optimistic Multiplicative Weights Update method.	23
2	Extra-gradient method.	24
3	Waterfilling approach algorithm for computing a <i>pure Nash equilibrium</i> in single-commodity non-atomic congestion games with n parallel links and affine latency functions.	135
4	Linear program	136
5	Coordination mechanism algorithm for an abstract partition (k_1, \dots, k_m) of n	137
6	Coordination mechanism algorithm	137

Chapter 1

Introduction

On the turn of the Century, we have witnessed an unprecedented growth in almost any scientific field and discipline. Data sciences, sensing, communications, computations, and automata, are only some of the research areas that have flourished, drastically changing the way our society functions. A prominent role in this change is the way we collect, process and integrate information. The recent progress in robotics, actuation and sensor network make more and more evident the possibility of autonomous agents helping us achieve tasks that are otherwise hazardous or impossible. Thus, unsurprisingly, the growing importance of multi-agent systems has also been heavily acknowledged by the research community.

Specifically, a multi-agent system is a system composed of multiple interacting autonomous, self-interested and intelligent agents[†], and their environment. Typically, agents need to be incentivized to choose a plan of action. Game theory [Nisan et al., 2007a] is a branch of mathematics that models and analyzes the behavior of agents that have preferences over possible outcomes and have to choose actions in order to implement these outcomes, when the success of these outcomes also depends on the actions of other agents. Thus, it is perfectly suited to provide a theoretical foundation for the analysis of a multi-agent system.

In general, game theory analyses many interaction settings considering several assumptions as to agents' perception, from the more strict consideration of common/correct information to the thinner of imperfect/incomplete information. In the case of common/correct information every specification of the interaction is known to the agents, whereas in the case of imperfect/incomplete assumption agents are unsure as to the specifications, they are well-aware of that, and they do their best out of the uncertainty that they have. Hence, in either case agents have at least some common perception (information, uncertainty etc.) about the underlying interaction. For example, uncertainty is modeled in some formal means, usually through a distribution that determines the probability of the different possible outcomes.

As game theory approaches becomes more involved in the development of modeling tools, many issues arise regarding the aptness of such approaches. A reasoning agent is often faced with erroneous and misleading information as regards the state of the world, and the possible

[†]Note that we use the terms “agent” and “player” interchangeably throughout the document.

outcomes of his/her actions (see [Luce and Raiffa, 1957, Bennett, 1980, Feinberg, 2020]). This could happen on purpose (e.g., by deceptive agents communicating wrong information), due to random effects (e.g., noise in the communication channels, or erroneous sensor readings), by design (e.g., when the game designer deceives the players to enforce a socially-optimal behaviour) or due to environmental changes (e.g., when the game changes due to external factors, without players' knowledge). Moreover, this discrepancy may be the result of players' limited awareness, bounded computational capacity, and cognitive restrictions. So, some aspects of the situation, or the modeling and reasoning with regards to the situation leads the players to incorporate only a selection, possibly incorrect, of the real aspects. Thus, they may miss the most crucial specifications of the interaction, and they may interact relying on a restricted and incorrect perception of the game. In a nutshell, players interact strategically having been misinformed about the real situation.

The present dissertation addresses this problem, i.e., the game-theoretic scenarios where players have been misinformed about the state of affairs, without being aware that such a misinformation may exist. To address this problem, we establish a game theoretic model that allows the players to have subjective specifications for their interaction, to interact with other players' different perceptions, to adapt their behaviour according to the information they receive from the environment and from the decisions of other participants, and to agglomerate these inferences to properly adapt their strategic behaviour; while allowing these subjective views to (possibly) differ from reality.

1.1 Modeling misinformation

The modeling of multi-agent systems has been an active field of research for scholars in Game theory. A common assumption in Game theory is that agents have a common (and correct) knowledge with regards to the abstract formulation of the game (players, payoffs, and strategies). This view is called games with complete information, i.e. see Figure 1.1a. Nevertheless, often (some of) the players lack complete information as to the game they participate. However, in many real-world situations it could be the case that (some of) the players have incorrect information with regards to the game that they play. In this dissertation we study the case where players are misinformed, meaning that they have different and/or incorrect knowledge about the rules of interaction.

To study these situations, we relax the assumption that agents know the correct information related to the abstract formulation of the game, and *introduce the formal machinery necessary to study multi-agent interactions where the agents may be misinformed with regards to the game definition (players, strategies, payoff matrices)*. In a nutshell, the actions of the players are dictated according to the subjective views they possess regarding the specifications of the interaction. As these subjective views may differ from player to player, may emerged many interesting or even “unexpected” behaviours. We call such games *misinformation*

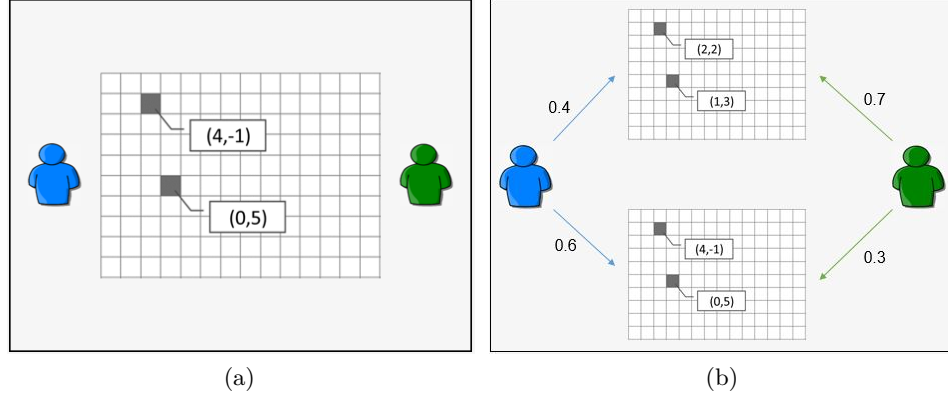


Figure 1.1: Games with information assumptions: (a) complete/correct, (b) incomplete/correct.

games [Varsos et al., 2021].

Obviously, in such settings, game theory and logic dictate the strategic decisions and actions of a player according to his/her view. On the other hand, the feedback, reward or penalty that is received by the players is provisioned according to the real specifications, which may differ from the ones that the players assume, see Figure 1.2.

Scenarios where agents lack complete information about the game specification have also been considered in other contexts, e.g. interactions with incomplete knowledge [Harsanyi, 1967, Zamir, 2009], probabilistic views [Harsanyi, 1967, Zamir, 2009], misspecified views [Luce and Raiffa, 1957] and utility theory [Peterson, 2009]. A popular approach that deals with situations where players lack complete information is that of *Bayesian game* [Harsanyi, 1967, Zamir, 2009], where a critical assumption is that the agents are not sure about the specifications of the game they participate, thus they take a “what if” stance. However, in several cases this typically limits the alternative scenarios considered to a small number, i.e., the agents will associate a positive probability to only a (possible) small number of alternative scenarios (specifically the scenarios that are aware of), see Figure 1.1a-1.1b. Thus, despite the rigorous tools, solid results and clear insights in the multi-agent interactions provided by Bayesian games, the issue of computational and cognitive limitations of the agents is not addressed. In a nutshell, in Bayesian games agents *know that they do not know*, while in *misinformation games they do not know that they do not know*.

On the other hand, utility theory is based on the assumption of rationality and describes all decision outcomes in terms of the utility (or value) placed on them by agents. Within this framework, decisions can be understood in terms of rationally ordered levels of utility attached to different outcomes. Furthermore, uncertainty in utility theory is described by a class of models designed to formalize the manner in which an agent chooses among alternative courses

of action when the consequences of each course of action are not known at the time the choice is made. But the concept of rationality is a great debate until now among utilitarians. Thus, this theory can not approach players with misinformation.*

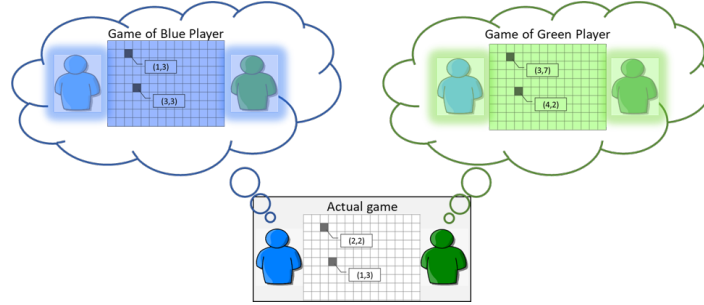


Figure 1.2: Schematic representation of two players' games with misinformation.

1.2 Research Questions

Having at hand the motivation we provide the main research questions that this thesis aims:

1. *How do we model the concept of misinformation formally?* [Chapters 4, 5, 7, 8]
2. *How can we explain and predict, in formal terms, the difference in behavior in the presence of misinformation?* [Chapters 4, 5, 7, 8]
3. *How decisive is the misinformation in various scenarios?* [Chapters 4, 5, 7, 8]
4. *How do we study the above in different classes of games?* [Chapters 4, 7, 8]
5. *How the different types of misinformation (i.e. undeliberate, random, by design) can be modeled and what can we show for each?* [Chapters 7, 8]
6. *What equilibrium points exist in games with misinformation?* [Chapters 4, 5]
7. *How can we analyze the beliefs of misinformed players?* [Chapter 5]
8. *What happens in the misinformed views of the players when they interact with each other?* [Chapter 5]
9. *Can we compute the equilibrium concepts arise from misinformation games in an learning manner?* [Chapter 6]

*In general, game theory (Bayesian games) and utility theory interlinked in several concepts (e.g. utility function), thus the differences between Bayesian games and misinformed games also hold for the framework of utility theory under uncertainty.

1.3 Outline and Contribution

In this thesis, we address the concept of misinformation games from various perspectives, and for different problems that are related to the idea. Each of these different perspectives is considered in its own chapter of this thesis. In more details, this thesis is organized as follows.

In Chapter 2 we present a literature review of the publications related to our study, this review is quite extensive, as our work expands in several areas of game theory, using a range of techniques and ideas.

Thereafter, in Chapter 3 we provide the basic game-theoretic background in order to establish our theory. This chapter is provided to allow a self-contained dissertation, and is meant to contain only the concepts, aspects, techniques and methodologies of game theory that will be useful in the following; therefore, it should not be treated as a complete account of game theory. Supplementary, the reader should look at the appendix sections, where we present concepts that are used in the analysis and lie in mathematical fields other than game theory.

Next, in Chapter 4 (see Misinformation games box in Figure 1.4) we present our game theoretic framework, called misinformation games, that provides the formal machinery necessary to study the phenomenon of misinformation with regards to the abstract formulation of a game (players, payoffs, and strategies). We establish our theory for the case of normal-form games [Shoham and Leyton-Brown, 2008]. As misinformation could play a prominent role in the outcome of the game, without necessarily negative effects, we provide a metric that measures the effect of misinformation on social welfare (compared to the optimum of the actual game), that is called *Price of Misinformation*. Furthermore, we transfuse the concept of misinformation games in the context of two other classes of games, *load balancing games* and *congestion games*, so as to show the applicability of our method. This work was held in collaboration with Dr. Giorgos Flouris, Dr. Marina Bitsaki and Dr. Michail Fasoulakis, resulting in a presentation in the PRICAI 2021 conference, alongside with a publication [Varsos et al., 2021].

In Chapter 5, we first develop and formulate mathematically a procedure for determining the outcome of a sequential interaction between the players in misinformation games. Interestingly, the equilibrium point of this procedure does not coincide with any equilibrium concept in the literature (see Adaptation Procedure box in Figure 1.4). Second, we enhance this approach providing epistemic characteristics to the interaction. Specifically, a player can infer the knowledge and beliefs that each participant possesses, by observing the decisions that have been made (see Epistemic Adaptive Evolution box in Figure 1.4). Further, she/he can take advantage of this *higher-level* information and take advantage of the opponents' misinformation to his/her own benefit. Again, this analysis provides novel solution concepts. This study was held in collaboration with Dr. Giorgos Flouris and Dr. Marina Bitsaki.

Furthermore, in Chapter 6 (see Learning Algorithm box in Figure 1.4) we study mis-

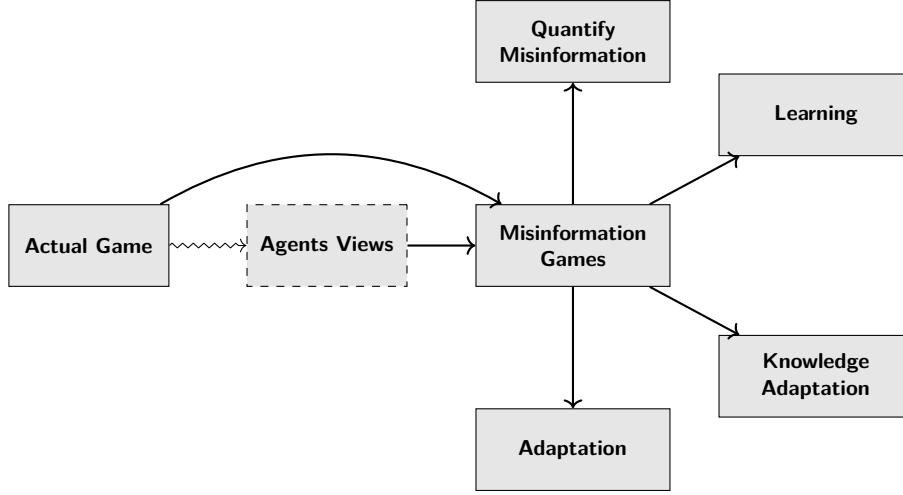


Figure 1.3: Graphical representation of misinformation in multi-agent systems.

information games from a different perspective. Initially, we develop a no-regret learning algorithm in order to compute the equilibrium points of a classical zero-sum game. Towards that direction, we propose a novel variant of *multiplicative weights update method* with forward-looking best-response strategies that guarantees last iteration convergence for zero-sum bimatrix games with unique Nash equilibrium; we call this method Forward Looking Best Response Multiplicative Weights Update method (FLBR-MWU). The proposed algorithm offers substantial gains compared to the state-of-the-art approaches [Mertikopoulos et al., 2019, Daskalakis and Panageas, 2019]. With this at hand, we have an alternative method to compute equilibria concepts that arise in misinformation games, and fulfill the zero-sum condition. This part of the study was accomplished in collaboration with Dr. Michail Fasoulakis, Prof. Vaggelis Markakis and Dr. Yannis Pantazis, resulting in a presentation in the AISTATS 2022 conference, alongside with a publication [Fasoulakis et al., 2021].

Afterwards, we study a special case of misinformation, attributed to noise and signal errors, a situation often occurring in distributed multi-agent systems. Specifically, in distributed multi-agent systems, agents are equipped with an internal logic that allows them to autonomously solve problems of a given nature. However, at deployment time, the precise specification of these problems is often unknown; instead, the details are communicated as needed at operation time, during the so-called “online phase” [Brown et al., 2017]. In such cases, unexpected communication errors, malfunctions in the communication module or noise may cause the agents to operate under a distorted problem specification, leading to unexpected behavior, we address this problem in Chapter 7 (see Noisy games box in Figure 1.4). This part of the study carried in collaboration with Dr. Giorgos Flouris and Dr. Marina Bitsaki, resulting in a manuscript to be submitted soon in IEEE Transactions on Games journal.

Typically, agents need to be incentivized to choose a desirable plan of action. In Chapter 8

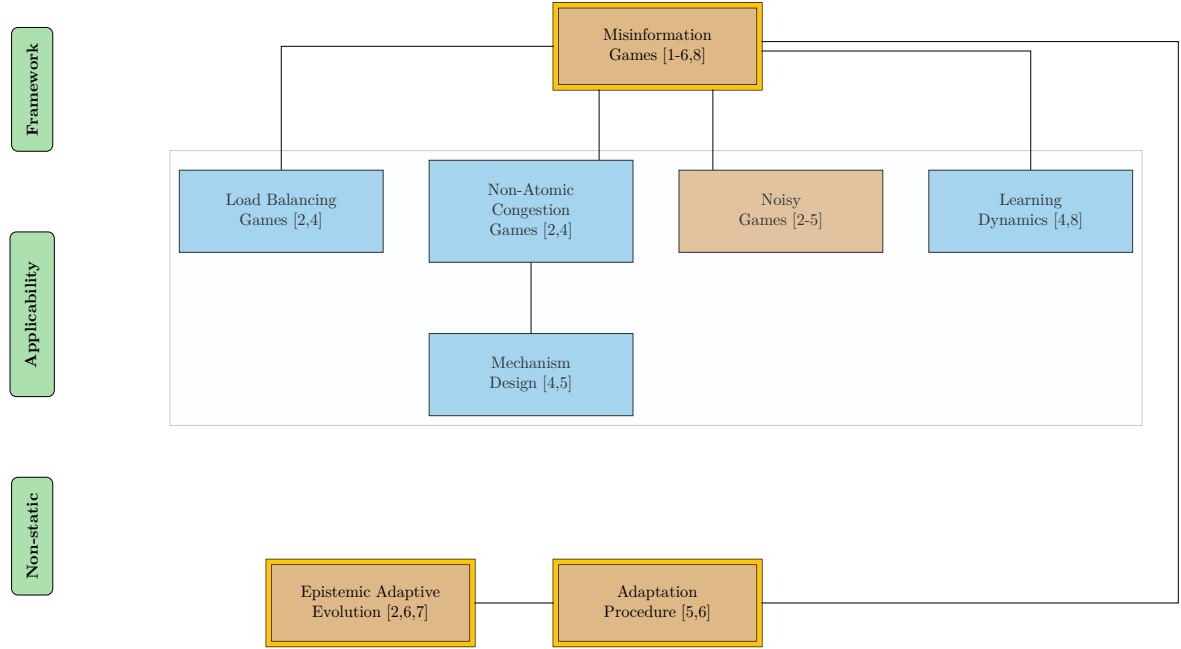


Figure 1.4: Road map of the thesis.

we introduce a novel application of misinformation games for coordination mechanisms in games, based on the idea of *misinforming* the players with regards to the game formulation to lead them to a situation with an improved outcome in terms of social welfare (see Mechanism Design box in Figure 1.4). For this, we study single-commodity non-atomic congestion games with n parallel links and affine cost functions, where the players have a potentially different view of the actual game that is being played. This study held in collaboration with Dr. Michail Fasoulakis, Dr. Giorgos Flouris and Dr. Marina Bitsaki, resulting in a presentation in the ICAART 2022 conference, alongside with a publication [Varsos et al., 2022].

Concluding, in Chapter 9 we provide a summary and conclusions of our study alongside with short-term and long-term future directions.

Finally, the reader should consult Figure 1.4 where we present the interconnection between all different aspects of the thesis. In the brackets inside each box we highlight the number of research questions that are studied in the respective part of the thesis. Blue colored boxes denote the areas that already exist in literature, and in which we contribute, whereas the brown colored boxes represent the new areas that we introduce. Also, the yellow thick boundary indicates our framework.

Chapter 2

Related Work

There are several main strands of work on game theory that related to the concept of misinformation games. These are:

1. Works that study misperception of the real situation.
2. Works that consider various settings and contexts in which the players play under payoff matrices that are different than the actual one.
3. Works that study and quantify the deviations of players' behaviour compared to the socially optimal one, a concept that is very similar to the deviations in players' behaviour due to misinformation.
4. Works that study how the players behave in case of iterative interactions.
5. Works that study online learning.
6. Works that study how the structure of misinformation affects the behaviour of the players.
7. Works that study the Coordination mechanisms.

We analyse these seven strands below, and elaborate on their relationship with misinformation games.

2.1 Misperception in games

Starting from the concept of games with misperceptions (see Chapter 12 in [Luce and Raiffa, 1957]) many studies model subjective knowledge of players with regards to game specifications, leading to the introduction of hypergames (HG) (e.g. [Bennett, 1980, Vane and Lehner, 2002, Sasaki and Kijima, 2012, Kovach et al., 2015, Cho et al., 2019, Bakker et al., 2021]) and games with unawareness (GwU) (e.g. [Copic and Galeotti, 2006, Schipper, 2014, Sasaki, 2017, Schipper, 2017, Feinberg, 2020]), where players may be playing different games. Although

we share motivation with these approaches, there are also some crucial distinctions. First, HG/GwU are behaviour-oriented (*what* the players will play), whereas misinformation games are outcome-oriented. Furthermore, HG focus on perceptual differences among players, and do not model the “actual game”, hence, HG lack grounding to the reality of the modelled situation. In misinformation games we close this gap, modelling also the environment, and allowing differences to also occur between each player and the environment. Moreover, in GwU, though the “actual game” is used as the basis of the models, the analysis based on consistency criteria and belief hierarchies. In misinformation games we do not make such assumptions.

In [Halpern and Rêgo, 2014] authors define the notion of games with awareness based on an extensive-form game; they agglomerate descriptions of reality, changes in players’ awareness and players’ subjective views. Also, they define a generalized Nash equilibrium that is similar with our equilibrium concept. Nevertheless, their analysis is behaviour-oriented. The work in [Feinberg, 2020] incorporates game and unawareness as interrelated objects, whereas in [Copic and Galeotti, 2006] awareness architectures are provided to study players’ limited awareness of strategies. Further, in [Thadden and Zhao, 2014] authors focus on how unawareness affects incentives, whereas [Schipper, 2018] provides a dynamic approach for extensive-form games with unawareness. Moreover, [Ozbay, 2007] proposed a model for games with uncertainty where players may have different awareness regarding a move of nature.

In [Chaib-Draa, 2001, Gharesifard and Cortés, 2011] studied the case where one of the players knows the (mis)perceptions of the opponents. Also, in [Teneketzis and Castañón, 1983] the concept of subjective games is proposed, but without introducing any equilibrium concept. Another approach is given in [Esponda and Pouzo, 2016a] where an equilibrium concept is defined, but has a probabilistic dependence on the actual game specifications.

Further, [Antos and Pfeffer, 2010, Banks et al., 2020, Roponen et al., 2020] the case of uncommon priors was studied, but without addressing the scenario of private priors, which is the case considered in misinformation games.

2.2 Games with distorted payoff matrices

The idea of agents understanding a different payoff matrix than the actual one has been considered for non-atomic routing games [Meir and Parkes, 2015b, Meir and Parkes, 2018], and for normal-form games [Acar and Meir, 2020]. In these studies, the authors suggest that the players play a modified game with cost functions potentially different than the actual game. In our methodology we generalize this idea and propose that each player plays according to the Nash equilibria of his/her own view of the game taking into account the optimal strategies of the other players’ behaviour in this specific view of the game. Thus, the approach of [Meir and Parkes, 2015a] can be considered as a case of misinformation games as it is described in our framework. Moreover, in our more general setting, misinformation is not restricted to

bias, but may have other causes (e.g., noise, deception etc.), or could be deliberate on behalf of the designer.

Although selfish attitude is considered as basic test tube for most of the literature, many modern approaches study the effect of different attitudes (e.g., altruism) in the performance of a multi-agent procedure (see [Chen et al., 2014, Caragiannis et al., 2010b, Brown, 2020]). From a game-theoretic aspect, the differences in agents' attitudes can be described and modelled as a misinformation game, where each agent's payoff is not purely determined by his/her own payoff (as described by the actual game), but is also affected by his/her altruistic motives.

The effect of additional knowledge in agents' behavior in the context of congestion games is studied in [Acemoglu et al., 2018], whereas exogenous distortions in the cost structure are examined in [Balcan et al., 2009]. The effect of information heterogeneity in a congestion game is studied in [Wu et al., 2020]; in this thesis different populations of agents receive a private signal from their traffic information system, while maintaining a belief about the signals received by other populations. In [Bilò et al., 2010], each player is unaware of agents outside his/her social neighborhood, thus his/her individual cost and strategy selection are not affected by them. Moreover, in [Brown et al., 2017] authors studied the effect of communication failures to the solutions (equilibria) of a game.

All the above works essentially study special cases of misinformation. In these works, internal agent characteristics or external conditions (e.g., bias or altruism), cause the agents to play using a payoff matrix that differs from the actual one in certain aspects. However, their settings are very limited in the sense that misinformation in each case has a specific form or affects a specific part of the payoff matrix. Our work provides a unifying, more general framework that can model all different types of misinformation, including the scenarios described in these papers; thus, our work allows a uniform formal description of these settings, and may act as a testbed for comparing their characteristics and uncover commonalities or differences.

Another scenario where the original payoff matrices are modified, can be found in works where tolls, penalties, rewards or other similar methodologies are applied to modify the payoff matrix of players towards some aim (e.g. [Caragiannis et al., 2006, Kleer and Schäfer, 2017b, Kleer and Schäfer, 2017a] etc.). Although in such games the payoff matrix is modified due to tolls or some similar mechanism, the new payoff matrix is the actual one, and is also the one perceived by all players, so these are not misinformation games.

Moreover, there is another stream of studies considering random payoff matrices, (i.e. [Daskalakis et al., 2007, Takahashi, 2008, Stanford, 1996, Rinott and Scarsini, 2000, Bárány et al., 2007, Dresher, 1970]) but authors focus in the distribution of pure Nash equilibria. A tweak of this methodology presented in [Quattropani and Scarsini, 2020] where authors study the distribution of players' average social utility.

2.3 Quantifying socially sub-optimal behaviour

Inspired by the seminal work of [Koutsoupias and Papadimitriou, 1999] and the introduction of the *Price of Anarchy (PoA)* metric, many studies analyze the deviation of the behaviour of a multi-agent system from the optimal one. Initially, *PoA* was applied to traffic networks [Roughgarden and Tardos, 2002, Roughgarden, 2003, Correa et al., 2008], where players' strategies depend on their types in terms of available paths and costs. The Bayesian setting of *PoA* was introduced in [Leme and Tardos, 2010]. The *Price of Risk Aversion* measure was introduced in [Nikolova and Moses, 2015] to deal with uncertainty and risk in selfish routing games with homogeneous risk profiles, an assumption that was dropped in [Cole et al., 2018]. In [Balcan et al., 2009] the term *Price of Uncertainty* was defined to capture the resilience of games to imperfections in agents' dynamics. Additionally, the *deviation ratio* was introduced in [Kleer and Schäfer, 2016] to capture the diversity in agents' decisions. In [Meir and Parkes, 2015b, Meir and Parkes, 2018] the *Biased Price of Anarchy (BPoA)* was introduced to measure the ratio of the equilibrium under biases in knowledge compared to the optimal outcome.

All the above metrics quantify the deviation of a system's behaviour (social welfare of equilibria) compared to the socially optimal one under various settings. The *Price of Misinformation (PoM)* metric introduced in the thesis is similar, capturing the effect of misinformation on social welfare compared to the optimal outcome. Moreover, our work shows that misinformation can be used as a mitigation measure against the sub-optimal behaviour quantified by these studies, as it may lead players to the optimal outcome and, thus, is a valuable tool for mechanism design.

2.4 Iterative interactions

Authors in [Esponda and Pouzo, 2016b] studied asymptotic beliefs and behavior in Markov decision processes, where the players have a prior over a set of possible transition probability functions and update their beliefs using Bayes' rule. Furthermore, in [Esponda and Pouzo, 2019] the authors consider an agent whose uncertainty about the environment is represented by a misspecified model and uses Bayes' rule to update his/her belief about the environment. They focus on the frequency of actions, as opposed to the action itself or the beliefs.

In [Arrow and Green, 1973] the authors provide a learning framework that makes distinctions between objective and subjective games, but they make the assumption that players are completely ignorant with regards to the decisions of their opponents. Further, each player applies Bayes theorem after receiving a new observation and no equilibrium point is considered. Contrary to this study, we allow the existence of more general types of misinformation, any new information is automatically integrated, and we establish the equilibrium concept of *stable misinformed equilibrium*.

Next, in [Spiegler, 2016], a framework is presented where each player makes decisions

under imperfect understanding of correlation structures, and fits his/her personal beliefs to an objective long-run probability distribution. In the same spirit, in [Easley and Kiefer, 1988] examined a procedure where players express beliefs about unknown parameters in terms of distributions. In [Lerer et al., 2019] the authors describe approaches to search in partially observable cooperative games, but they assume perfect knowledge of other players' policies.

Further, in [Jordan, 1991] a class of Bayesian processes for iterated normal form games is studied, where each player knows his/her own payoff function, but is uncertain about the opponents' payoffs. In [Esponda and Pouzo, 2016a] authors provide a time dependent learning mechanism in games where players have subjective views of an objective game. Specifically, at each time step t each player updates his/her beliefs using Bayes rule and the information obtained in all previous time steps. A new equilibrium concept is defined, Berk-Nash equilibrium, where the admissible strategic behaviour derived as a probabilistic distance between the behaviour of the players in their subjective and objective specifications.

In the same manner, in [Ray et al., 2008], behavioral-based model are provided in terms of partial observable Markov decision processes using belief hierarchies, in order to model the ignorance of each player about the opponent. Further, in [Shalizi, 2009] the author gives sufficient conditions for the convergence of the posterior without assuming that the subjective views are part of actual specifications. Also, in [Fudenberg et al., 2016] a complete characterization of the limit behavior of actions in cases with misspecified Bayesian players is provided.

As opposed to the stream of works that rely on probabilistic or Bayesian techniques, in this dissertation we take no probabilistic considerations with regards to the beliefs of the players. Our study focuses on the effect of new information in the misinformed views of the players.

A significant number of works focuses on the limit behaviour of players whose strategic choices are evolving, for various reasons. In [Nyarko, 1991] an example was presented where the agent's actions cycle ad infinitum. Authors in [Romanyuk et al., 2017] based on a continuous time model with beliefs over types, provide characterization of asymptotic behaviour and beliefs. Moreover, in [Heidhues et al., 2017] a convergence analysis was introduced in case where players bias their observations, while in [Heidhues et al., 2018] authors establish convergence of beliefs, and actions, in a misspecified model with endogenous actions.

Besides that, in [Gharesifard and Cortés, 2012] authors introduce a swap learning method in a 1-level HG, where players can change their beliefs according to the information they gain by observing the opponents' actions. Specifically, a player decreases his/her misperception at the cost of potentially incurring inconsistencies in his/her perception. In our framework, we make no considerations as to the revealing inconsistencies. Additionally, authors in [Burkov and Chaib-draa, 2009] provide an approach for learning in adaptive dynamic systems, where a player learns an efficient policy over the opponents' adaptive dynamics, and obtain a lower bound of the utility which is guaranteed to be reached by that approach against any opponent.

Though, they did not consider any concept of incorrect or subjective information.

Further, as strategic reasoning does not always suffice to explain equilibrium behavior, there is a substantial stream of works regarding the epistemic conditions for equilibrium concepts (e.g. [Aumann and Brandenburger, 1995, Bach and Tsakas, 2014] for Nash equilibrium, [Aumann, 1987] for correlated equilibrium). Another equilibrium concept that have been introduced is that of self-confirming solution concept [Fudenberg and Levine, 1993], in which each player's strategy is a myopic best response to his/her beliefs about the play of the opponents and are consistent with what is observed when the player plays the game. Moreover, in [Fudenberg and Levine, 1993] authors provide an epistemic model assuming almost common certainty of payoffs with independent beliefs. Additionally, in [Dekel et al., 1999] authors define the rationalizable self-confirming equilibrium, where players see the realized terminal node at the end of each play of the game in extensive-form games. Towards this direction, authors in [Fudenberg and Kamada, 2015] study the case where there are independent beliefs. Here, the equilibrium concept arises from misinformation game, the *natural misinformed equilibrium*, is just the agglomeration of player's choice, thus may or may not be consistent with the beliefs of each agent. Though *stable natural equilibrium* and *stable epistemic natural misinformed equilibrium* are self-confirming equilibria, they derived as a product of adaption procedure and epistemic adaptive evolution concept respectively.

Considering epistemic approaches in HG, authors in [Sasaki, 2014] introduced a new solution concept such that every player takes a best response to expected choices of the others, every player thinks every player takes a best response to expected choices of the others, and so on. This approach constrained in one-shot HG, thus addressing only one-shot interaction as opposed to our multi-turn processes.

For the case of GwU, in [Heifetz et al., 2008] established, in terms of logic, a state-space model that allows for non-trivial unawareness among several individuals with subjective views and strong properties of knowledge, whereas the model is designed to capture mutual beliefs about unawareness.

In [Meier and Schipper, 2014] authors combine unawareness belief structures and Bayesian games in order to establish Bayesian games with unawareness, applying their idea to strategic games. Thereupon, a methodology to construct unawareness belief structures is given in [Heifetz et al., 2013, Heinsalu, 2011]. Meanwhile, in [Fagin and Halpern, 1987] authors study the lack of logical omniscience (lack of awareness and local reasoning) presenting the Logic of General Awareness and providing the concept of awareness function, which further expanded in [Heifetz et al., 2006, Belardinelli and Rendsvig, 2020]. The main difference of our study with these streams of work is that misinformation is conceptually distinct than unawareness, ignorance, incomplete information or faulty analysis. Namely, misinformation has to do with the lack of conception and knowledge combined. Marginally, if we have only lack of conception we fall in the case of unawareness, while if we have only lack of knowledge misinformation collapses to the other cases.

2.5 Dynamics in games

In terms of games with subjective views, such as HG and GwU, there is no implementation of online algorithms in order to compute equilibria concepts. Nevertheless, in this Section we provide some key works so as the presentation of our work in Chapter 6 to be more complete.

Several approaches have been proposed throughout the past decades, for the development of fast, iterative learning algorithms, starting with *fictitious play* [Robinson, 1951]. Recently, some of the more standard methodologies include the family of no-regret algorithms as well as several classes of first-order methods (see [Cesa-Bianchi and Lugosi, 2006, Hoi et al., 2018] for an overview). To mention a few examples, the important class of *Multiplicative Weights Update* (MWU) method [Littlestone and Warmuth, 1994, Freund and Schapire, 1999], together with Gradient Descent, Mirror Descent, and Extra Gradient methods (for a survey see [Bubeck, 2015] and [Beck, 2017]), all fall within the above approaches. Unfortunately, most of the standard no-regret algorithms exhibit convergence in an average sense, (e.g. [Bailey and Piliouras, 2018] for MWU methods for constrained problems; [Mertikopoulos et al., 2018] on variants of Gradient Descent/Ascent (GDA) methods).

Due to the importance of achieving last-iterate convergence for zero-sum games for applications on learning, such as training GANS, and boosting, [Daskalakis et al., 2018] and [Liang and Stokes, 2019] studied an optimistic variant of GDA (OGDA); proving that OGDA exhibits last-iterate convergence for the unconstrained minmax problem with bilinear functions.

Driven by the progress made by OGDA, [Daskalakis and Panageas, 2019] proposed to study the constrained version of minmax problems, i.e., finding equilibria in zero-sum games, via an optimistic variant of MWU methods, termed *Optimistic Multiplicative Weights Update* OMWU. Their main result is that for games with a unique Nash equilibrium, OMWU converges in the last-iterate sense to the equilibrium. The sequence of approximations in OMWU uses two previous steps in order to compute the next update, where the extra term corrects the behaviour of MWU dynamics and can be seen as a negative momentum. Moreover, the performance of OMWU provides a strengthening to the experimental results in [Syrgkanis et al., 2015], which indicated pointwise convergence of the optimistic version. An alternative view on the convergence behavior of OMWU by studying volume contraction is given in [Cheung and Piliouras, 2020]. Further generalizations of OMWU and more convergence results were obtained in [Lei et al., 2021] for convex-concave landscapes.

As show in [Daskalakis et al., 2018] and [Bailey and Piliouras, 2018] for gradient descent and Follow-The-Regularized-Leader respectively, gradient/mirror descent methods do not suffice to address the bilinear case.

To cope this issue authors in [Mertikopoulos et al., 2019] introduce a gradient-like method, where they first compute an intermediate approximation, by taking a prox step, and then go back to the original state and compute the new state using the former, intermediate,

approximation. This method is known as *Optimistic Mirror Descent* (OMD) (see [Chiang et al., 2012, Rakhlin and Sridharan, 2013]), or *mirror-prox* [Nemirovski, 2004]. A similar approach was introduced by [Gidel et al., 2019] where extra gradients introduced in order to minimize the computational overhead of back-propagation. All these techniques fall under the umbrella of extra-gradient method, a classical method which was introduced in [Korpelevich, 1976] (for more details see [Facchinei and Pang, 2003]).

Several streams of works focus on the convergence of Extra-gradient methods (EG). In [Liang and Stokes, 2019] authors prove linear convergence of the EG with the assumption that the payoff matrix of the bilinear function is square and full rank. Using the same assumption, authors in [Mokhtari et al., 2020] study the EG for the general saddle point problem and provide linear rates of convergence. In the very recent work of [Wei et al., 2021], a new theoretical analysis is provided for both OMWU and OGDA, with the goal of quantifying the progress in the divergence decrease with respect to the number of steps (applicable also to some more general problems).

Finally, authors in [Azizian et al., 2020] provide a similar in spirit approach with FLBR-MWU (Proposition 7), though there are also differences on the dynamics and the step sizes α and η are more constrained than in our setting. Most importantly, the analysis in [Azizian et al., 2020] addresses the unconstrained bilinear case. Even further, they show only local convergence (starting from a point near the fixed point), whereas we do not need such a condition, and their result also holds under certain spectral assumptions.

2.6 Structure of misinformation

As noisy games are a subclass of misinformation games, automatically are related to works in the area of games with misperceptions. Nevertheless, there is no consideration about the structure of misperception this area. As opposed to Bayesian games, where there is a rich literature that studies the influence of the structure of uncertainty in the knowledge of the players to their strategic behaviour. Thus, conceptually we are close to the first group of works. On the other hand though, we also study the effect of distributions in the knowledge of the players as to their strategic behaviour, thus our results can be related to the latter stream of works.

In [Balcan et al., 2009], the authors consider the impact of small fluctuations in the cost functions or in players' perceptions of the cost structure in congestion and load balancing games, and study its effect on players' behaviour. A fluctuation is a departure from the classical viewpoint that treats payoffs as a number; under [Balcan et al., 2009], the payoff is a range of values "close" to the actual payoff. An extension of [Balcan et al., 2009] considered normal-form games, aiming to define a new notion of equilibrium that maximizes the worst case outcome over possible actions by other players [Aghassi and Bertsimas, 2006] in the presence of fluctuations, whereas a further extension (see [Balcan and Braverman, 2017])

studied the robustness of this equilibrium solution, utilizing the notion of approximations of payoffs using a fuzziness to the values of payoff matrices. There are several differences of the concept of fluctuation as compared to misinformation. First, the players are aware of the fluctuations and, thus, take them into account while deciding on their strategic choices. Second, fluctuations affect all players and all payoffs uniformly. Third, fluctuations have a limited effect, whereas the noise considered in our work may have unlimited effect (subject to a certain probability function).

Further, in [Brown et al., 2017] the authors study how resilient is the strategic behaviour of players when an unexpected communication loss occurs, and explore game settings in which communication failures can/cannot cause harm in the strategic behaviour of players. They introduce the notion of *proxy payoffs* in order to funnel communication failures and show that, in several settings, loss of information may cause arbitrary strategic behaviours. Our work has a similar contribution as both works prove that in presence of communication inefficiencies any strategic behaviour is possible. Though, authors in [Brown et al., 2017] focus on how the agents can be provided with policies so as to cope with communication failures, in this study we analyse the impact of disorder in the strategic behaviour of the players. Also, we model communication failure using probabilities, and provide formulas that quantify the probability of arbitrariness in the strategic behaviours, when information is degraded due to noise.

2.7 Coordination mechanism

The idea of designing mechanisms to improve coordination in multi-player systems with selfish players is not new. One approach is to introduce taxes (e.g. [Fleischer et al., 2004, Fotakis and Spirakis, 2008, Caragiannis et al., 2010a]), e.g. in congestion games players pay a toll for every edge they use. Specifically, in [Cole et al., 2003] the authors showed that there exist taxes that reduce the *Price of Anarchy* to 1. There are two major issues for this approach: i) taxes may be very high, and ii) in the case where taxes are part of the cost, then the *Price of Anarchy* is not improved. As another approach, rewards and payments have been used so as to improve coordination [Lavi and Swamy, 2007, Seregina et al., 2017, Turrini, 2016]. An alternative approach was given in [Monderer and Tennenholtz, 2003] that extends the game by adding new strategies for the players, such that all Nash equilibria of the new game involve strategies of the original game exclusively. In that case the *Price of Anarchy* is decreased.

Chapter 3

Theoretical Background

In this chapter we provide all the classical game-theoretic definitions and concepts that are necessary for the chapters to come. Reader, should consider this chapter as the initial point of the thesis and he/she can recur at will.

3.1 Normal-form games

A *game in normal form* is represented by a *payoff matrix* that defines the payoffs of all players for all possible combinations of pure strategies. Formally:

Definition 1. A *normal-form** game G is a tuple $\langle N, S, P \rangle$, where:

- N is the set of the players,
- $S = S_1 \times \dots \times S_{|N|}$, S_i being the set of pure strategies of player $i \in N$,
- $P = (P_1, \dots, P_{|N|})$, $P_i \in \mathbb{R}^{|S_1| \times \dots \times |S_{|N|}|}$ is the payoff matrix of player i .

When player i randomly selects a pure strategy, then he plays a mixed strategy $\sigma_i = (\sigma_{i,1}, \dots, \sigma_{i,|S_i|})$ which is a discrete probability distribution over S_i .

Let the set of all possible mixed strategies σ_i be Σ_i . A strategy profile $\sigma = (\sigma_1, \dots, \sigma_{|N|})$ is an $|N|$ -tuple in $\Sigma = \Sigma_1 \times \dots \times \Sigma_{|N|}$. We denote by σ_{-i} the $|N| - 1$ -tuple strategy profile of all other players except for player i in σ .

We call a *position* a vector of integers that determines a specific item in a payoff matrix. We denote by Pos the set of all positions. Formally (and given the above assumptions on the dimensions of a payoff matrix), $Pos = [|S_1|] \times \dots \times [|S_{|N|}|]$, i.e., a position is a tuple of the form $\vec{v} = (v_1, \dots, v_{|N|}) \in Pos$, where $1 \leq v_i \leq |S_i|$ (for all $i \in [|N|]$).

The payoff function of player i is defined as: $h_i : \Sigma \rightarrow \mathbb{R}$, such that:

$$h_i(\sigma_i, \sigma_{-i}) = \sum_{k \in S_1} \dots \sum_{j \in S_{|N|}} P_i(k, \dots, j) \cdot \sigma_{1,k} \cdot \dots \cdot \sigma_{|N|,j}, \quad (3.1)$$

*In other words, we define normal-form games as finite, noncooperative strategic games. In case, where the games are either non-finite or cooperative we will state it and we will modify our concepts accordingly.

where $P_i(k, l, \dots, j)$ is the payoff of player i in the pure strategy profile (k, l, \dots, j) , and $(k, l, \dots, j) \in \text{Pos}$. In other words, $h_i(\sigma_i, \sigma_{-i})$ represents player's i expected payoff as a function of σ . Further, in case where $|N| = 2$ we have two player normal-form game, called *bimatrix game*. The *Nash equilibrium* in a normal-form game is defined as follows:

Definition 2 (see [Nash, 1951]). *A strategy profile $\sigma^* = (\sigma_1^*, \dots, \sigma_{|N|}^*)$ is a Nash equilibrium, if and only if, for any i and for any $\hat{\sigma}_i \in \Sigma_i$, $h_i(\sigma_i^*, \sigma_{-i}^*) \geq h_i(\hat{\sigma}_i, \sigma_{-i}^*)$.*

In the rest of the dissertation we will denote as $NE(G)$ the set Nash equilibria of a game $G = \langle N, S, P \rangle$ (or simply NE , when G is obvious from the context). This notation will be slightly abused in case we focus only in the payoff matrix of a G and would become $NE(P)$.

Definition 3 (*i*-neighbours). *Given a player i and two strategy profiles $\sigma = (str_1, \dots, str_n)$, $\sigma' = (str'_1, \dots, str'_n)$, the profiles σ, σ' are called *i*-neighbours if and only if $str_i \neq str'_i$ and for all $j \neq i$, $str_j = str'_j$.*

Definition 4 (best response). *Consider a normal-form game $G = \langle N, S, P \rangle$ and a player i , a best response strategy function $BR(i, h, \sigma_{-i}) \subseteq \Sigma_i$ for player i is a strategy of i player against the strategy σ_{-i} of the other players if and only if*

$$BR(i, h, \sigma_{-i}) := \arg \max_{z \in \Sigma_i} h_i(z, \sigma_{-i})$$

Lemma 1. *Consider a normal-form game $G = \langle N, S, P \rangle$, where $N = [|N|]$ and $S = S_1 \times \dots \times S_N$. Consider also some player $x \in N$. We create the game $\tilde{G} = \langle \tilde{N}, \tilde{S}, \tilde{P} \rangle$, such that $\tilde{N} = \{x, \mathbf{y}\}$, $\tilde{S} = \tilde{S}_x \times \tilde{S}_{\mathbf{y}}$, $\tilde{S}_x = S_x$, $\tilde{S}_{\mathbf{y}} = S_1 \times \dots \times S_{x-1} \times S_{x+1} \times \dots \times S_N$, and \tilde{P} is such that for any $s_x \in \tilde{S}_x$, $s_{\mathbf{y}} = (s_1, \dots, s_{x-1}, s_{x+1}, \dots, s_N) \in \tilde{S}_{\mathbf{y}}$ it holds that $\tilde{P}(s_x, s_{\mathbf{y}}) = P(s_1, \dots, s_N)$. Then, for any strategy profile $\sigma = (\sigma_1, \dots, \sigma_N)$, it holds that: $\sigma_x \in BR(x, P, \sigma_{-x})$ if and only if $\sigma_x \in BR(x, \tilde{P}, (\sigma_1, \dots, \sigma_{x-1}, \sigma_{x+1}, \dots, \sigma_N))$.*

Proof. Since $s_{\mathbf{y}} = (s_1, \dots, s_{x-1}, s_{x+1}, \dots, s_N) \in \tilde{S}_{\mathbf{y}}$, then $\sigma_{\mathbf{y}} = (\sigma_1, \dots, \sigma_{x-1}, \sigma_{x+1}, \dots, \sigma_N) \in \tilde{S}_{\mathbf{y}}$. Thus, we have that $BR(x, P, (\sigma_1, \dots, \sigma_{x-1}, \sigma_{x+1}, \dots, \sigma_N)) = \arg \max_{z \in \Sigma_x} h_x(z, \sigma_{-x}) = \arg \max_{z \in \Sigma_x} \tilde{h}_x(z, \sigma_{\mathbf{y}}) = BR(x, \tilde{P}, (\sigma_{\mathbf{y}}))$. \square

A useful concept in the analysis of games is the *support* of a mixed strategy σ_i , consisting of pure strategies with a positive probability under σ_i , i.e., $\text{supp}(\sigma_i) = \{j : \sigma_{i,j} > 0\}$. It is easy to see that at an equilibrium σ^* , any pure strategy s_i , with $i \in \text{supp}(\sigma_i^*)$, is a best-response against $\sigma_{-i}^* \forall i \in [|N|]$.

Definition 5. *We define as social welfare function the function $h(\sigma)$ that measures the players' welfare given a strategy profile σ .*

Among many social welfare functions we use the *utilitarian* or *Benthamite* social welfare function that measures social welfare as the total or sum of individual incomes (rewards, payoffs).

Having at hand the definition of the Nash equilibrium, authors in [Koutsoupias and Papadimitriou, 1999] introduce a metric that measures the efficiency of a game due to selfish behavior of its players in the worst case. Thus, we can compare the behavior of the players in the worst case against the social optimum, based of a social welfare function h . We denote by opt the socially optimal strategy profile, i.e. $opt = \arg \max_{\sigma} h(\sigma)$. In terms of normal-form games this becomes:

Definition 6 (see [Koutsoupias and Papadimitriou, 1999]). *Given a normal-form game G , the Price of Anarchy (PoA) is defined as*

$$PoA = \frac{h(opt)}{\min_{\sigma \in NE} h(\sigma)} \quad (3.2)$$

3.2 Zero-sum games and approximate equilibria

Now, we consider a subclass of normal-form games, the finite two player zero-sum games, with payoff matrix $R \in (0, 1]^{|S| \times |S|}$, where without loss of generality, we assume both players have $|S|$ pure strategies. Thus, by Definition 1 we have the game $G = \langle N, S_1 \times S_2, P = (R, -R) \rangle$, with $|S_1| = |S_2| = S$. Further, we refer to the two players as the *row player* x and the *column player* y respectively. As before, if the row player plays the i -th row and the column player plays the j -th column, then the payoff of the row player is R_{ij} , and the payoff of the column player is $-R_{ij}$. Again mixed strategies are probability distributions (column vectors) on the pure strategies. E.g., a mixed strategy for the row player will be denoted as $\sigma_x = (\sigma_{x_1}, \dots, \sigma_{x_n})$, where σ_{x_i} is the probability of playing the i -th row. For convenience, we will denote the i -th pure strategy of a player by the unit vector e_i , which has probability one in its i -th coordinate and 0 elsewhere.

Given such a strategy profile $\sigma = (\sigma_x, \sigma_y)$, the expected payoff of the row player is $\sigma_x^T R \sigma_y$, whereas for the column player, it is $-\sigma_x^T R \sigma_y$; whereas the Nash equilibrium definition in case of finite two player zero-sum games becomes:

Definition 7. *A strategy profile (σ_x^*, σ_y^*) is a Nash equilibrium in the zero-sum game defined by matrix R , if and only if, for any $i, j \in [|S|]$,*

$$\sigma_x^{*T} R \sigma_y^* \geq e_i^T R y^* \quad \text{and} \quad \sigma_x^{*T} R e_j \geq \sigma_x^{*T} R \sigma_y^*,$$

Furthermore, the payoff of the row player at an equilibrium, $v = \sigma_x^{*T} R \sigma_y^*$, is referred to as the value of the game.

In several cases in game theory the condition of Nash equilibrium is weakened to allow the possibility that a player may have a small incentive to deviate. For that is introduced the concept of ε -Nash equilibrium, that approximately satisfies the condition of the Definition 7. Formally, this concept defined as:

Definition 8. A strategy profile (σ_x^*, σ_y^*) is an ε -Nash equilibrium in the zero-sum game defined by matrix R , if and only if, for any i, j

$$\sigma_x^{*T} R \sigma_y^* + \varepsilon \geq e_i^T R \sigma_y^* \quad \text{and} \quad \sigma_x^{*T} R e_j \geq \sigma_x^{*T} R \sigma_y^* - \varepsilon$$

This solution concept may be preferred to Nash equilibrium due to being easier to compute. In that direction, there are several algorithms (e.g. online, learning, distributed etc.) that rely on the computation ε -Nash. As we will see in Chapter 6, our work focuses on approximating equilibria, and in order to define the relevant notion of approximation, we start with approximate best responses. Given a profile (σ_x, σ_y) , we say that a strategy σ'_x is an ε -best-response strategy to σ_y with $\varepsilon \in [0, 1]$, if it yields a payoff that is at most ε less than the *best-response payoff*. We can define now an approximate Nash equilibrium, as a profile (σ_x, σ_y) where σ_x and σ_y are both approximate best responses to each other. This is precisely the standard notion of additive, approximate equilibria [Nisan et al., 2007b].

Obviously, any zero-Nash equilibrium is an exact equilibrium of the game.

3.3 Learning Dynamics

A central problem in game theory and optimization is computing a pair of probability vectors (x, y) , solving

$$\min_{y \in \mathcal{Y}} \max_{x \in \mathcal{X}} g(x, y) \tag{3.3}$$

where \mathcal{X}, \mathcal{Y} are closed convex sets. In case where g is a bilinear function, that is $g(x, y) = x^T R y$ and $R \in \mathbb{R}^{n \times m}$, \mathcal{X} and \mathcal{Y} are simplex, (3.3) is known as the classical two-player zero-sum game in normal form. Thus, any pair of probability vectors (x, y) corresponding to a strategy profile $(\sigma_x, \sigma_y)^\dagger$. The celebrated minmax theorem provided by von Neumann's [von Neumann, 1928] results in

$$\min_{y \in \mathcal{Y}} \max_{x \in \mathcal{X}} g(x, y) = \max_{x \in \mathcal{X}} \min_{y \in \mathcal{Y}} g(x, y)$$

Hence, all solutions of $\min_{y \in \mathcal{Y}} \max_{x \in \mathcal{X}} x^T R y$ are also solutions to the $\max_{x \in \mathcal{X}} \min_{y \in \mathcal{Y}} x^T R y$, and vice versa[‡].

Interestingly, a solution to (3.3) constitutes an equilibrium in case where $x^T R y$ is the payment of the min player to the max player; when the former selects a distribution y over columns and the latter selects a distribution x over rows of matrix R . This equilibrium point is often called minmax equilibrium, and fulfills the condition of Definition 2.

Meanwhile, (3.3) can be solved using any Linear Programming technique. Fortunately, author in [Adler, 2013] proved that any linear program can be addressed by solving some

[†]Abusing notation, in the part of the study dealing with dynamics we use (x, y) as strategy profiles instead of (σ_x, σ_y) .

[‡]This result is considered as one of the founding stones in the development of game theory.

minmax problem of the form (3.3). Thus, the linkage between the minmax theorem and the Linear Programming paved the way for the development of dynamics for solving minmax optimization problems of the form (3.3).

After arguing about the necessity of learning techniques in game theory, we turn our attention in the no-regret learning algorithms in order to compute Nash-equilibria. Specifically, we focus on multiplicative weights update (MWU) methods that use an exponential multiplication function, as in [Freund and Schapire, 1999]. Below we define one of the popular versions of MWU, that is resulted from the FTRL dynamics (Follow-The-Regularized-Leader), when the regularizer is the negative entropy function, (see [Hoi et al., 2018]). In particular, if (x^t, y^t) is the strategy profile at iteration t , and η is the learning rate parameter, the update rule of the method, for all $i, j \in [n]$ is as follows.

$$x_i^t = x_i^{t-1} \cdot \frac{e^{\eta e_i^T R y^{t-1}}}{\sum_{j=1}^n x_j^{t-1} e^{\eta e_j^T R y^{t-1}}}, \quad y_j^t = y_j^{t-1} \cdot \frac{e^{-\eta e_j^T R^T x^{t-1}}}{\sum_{i=1}^n y_i^{t-1} e^{-\eta e_i^T R^T x^{t-1}}} \quad (3.4)$$

In the sequel, we often provide the optimistic variant proposed by [Daskalakis and Panageas, 2019], which exhibits last-iterate convergence in zero-sum games, unlike the dynamics of (3.4), referred to as the Optimistic Multiplicative Weights Update (OMWU) method. The idea of “optimism” is derived from the notion of *predictable processes* [Rakhlin and Sridharan, 2013]. OMWU takes advantage of a given predictable process that takes into account two previous iterations in order to compute the next update, where the extra term can be seen as a negative momentum, correcting the behavior of MWU dynamics.

The dynamics of OMWU are described below for all $i, j \in [n]$.

$$\begin{aligned} x_i^t &= x_i^{t-1} \cdot \frac{e^{\eta e_i^T (2Ry^{t-1} - Ry^{t-2})}}{\sum_{j=1}^n x_j^{t-1} e^{\eta e_j^T (2Ry^{t-1} - Ry^{t-2})}}, \\ y_j^t &= y_j^{t-1} \cdot \frac{e^{\eta e_j^T (-2R^T x^{t-1} + R^T x^{t-2})}}{\sum_{i=1}^n y_i^{t-1} e^{\eta e_i^T (-2R^T x^{t-1} + R^T x^{t-2})}} \end{aligned} \quad (3.5)$$

and provided by Algorithm 1[§].

Next, we replace *optimism* with an *extra-gradient* step which uses the obtained information to “amortize” the next prox step (possibly outside the convex hull of generated states). This methodology initiated in [Korpelevich, 1976].

The main idea of the Extra-gradient (EG) method is to use the gradient at the current point to find a mid-point, and then use the gradient at that mid-point to find the next iterate. To be more precise, given a stepsize $\eta > 0$, the update of EG at step t solves the bilinear problem generating an intermediate, “waiting” point, equation (3.6). Instead of continuing from that intermediate point, the method samples the gradients and goes back to the original

[§] $\langle \mathbf{x}, \mathbf{y} \rangle$ denotes the standard inner product $\mathbf{x}^T \mathbf{y}$ in the Euclidean space.

Algorithm 1 Optimistic Multiplicative Weights Update method.

Input :Stepsize $\eta > 0$,
 Vectors $\mathbf{x}^0, \mathbf{y}^0, \mathbf{x}^1, \mathbf{y}^1 = \mathbf{1}_n/n$.

Output: $(\mathbf{x}^*, \mathbf{y}^*)$.

for $t = 1, 2, \dots$ **do**

$$\begin{aligned} x_i^{t+1} &= x_i^t \cdot \frac{\exp(-2\eta \nabla_{\mathbf{x}} f(\mathbf{x}^t, \mathbf{y}^t) + \eta \nabla_{\mathbf{x}} f(\mathbf{x}^{t-1}, \mathbf{y}^{t-1}))_i}{\langle \mathbf{x}^t, \exp(-2\eta \nabla_{\mathbf{x}} f(\mathbf{x}^t, \mathbf{y}^t) + \eta \nabla_{\mathbf{x}} f(\mathbf{x}^{t-1}, \mathbf{y}^{t-1})) \rangle} \\ y_i^{t+1} &= y_i^t \cdot \frac{\exp(2\eta \nabla_{\mathbf{y}} f(\mathbf{x}^t, \mathbf{y}^t) - \eta \nabla_{\mathbf{y}} f(\mathbf{x}^{t-1}, \mathbf{y}^{t-1}))_i}{\langle \mathbf{y}^t, \exp(2\eta \nabla_{\mathbf{y}} f(\mathbf{x}^t, \mathbf{y}^t) - \eta \nabla_{\mathbf{y}} f(\mathbf{x}^{t-1}, \mathbf{y}^{t-1})) \rangle} \end{aligned}$$

end

point \mathbf{x} in order to generate a new approximation, equation (3.7).

The dynamics of EG are described below for all $i, j \in [n]$.

$$x_i^{t-1/2} = x_i^{t-1} \cdot \frac{e^{\eta e_i^T R y^{t-1}}}{\sum_{j=1}^n x_j^{t-1} e^{\eta e_j^T R y^{t-1}}}, \quad y_j^{t-1/2} = y_j^{t-1} \cdot \frac{e^{-\eta e_j^T R^T x^{t-1}}}{\sum_{i=1}^n y_i^{t-1} e^{-\eta e_i^T R^T x^{t-1}}} \quad (3.6)$$

$$x_i^t = x_i^{t-1} \cdot \frac{e^{\eta e_i^T R y^{t-1/2}}}{\sum_{j=1}^n x_j^{t-1} e^{\eta e_j^T R y^{t-1/2}}}, \quad y_j^t = y_j^{t-1} \cdot \frac{e^{-\eta e_j^T R^T x^{t-1/2}}}{\sum_{i=1}^n y_i^{t-1} e^{-\eta e_i^T R^T x^{t-1/2}}} \quad (3.7)$$

Algorithm 2 Extra-gradient method.

Input :Stepsize $\eta > 0$,
 Vectors $\mathbf{x}^0, \mathbf{y}^0 \in \mathbb{R}^n$.

Output: $(\mathbf{x}^*, \mathbf{y}^*)$.

for $t = 0, 1, \dots$ **do**

$$\begin{aligned} \text{Compute : } x_i^{t+1/2} &= x_i^t \cdot \frac{\exp(-\eta \nabla_{\mathbf{x}} f(\mathbf{x}^t, \mathbf{y}^t))_i}{\langle \mathbf{x}^t, \exp(-\eta \nabla_{\mathbf{x}} f(\mathbf{x}^t, \mathbf{y}^t)) \rangle} \\ y_i^{t+1/2} &= y_i^t \cdot \frac{\exp(\eta \nabla_{\mathbf{y}} f(\mathbf{x}^t, \mathbf{y}^t))_i}{\langle \mathbf{y}^t, \exp(\eta \nabla_{\mathbf{y}} f(\mathbf{x}^t, \mathbf{y}^t)) \rangle} \\ \text{Update : } x_i^{t+1} &= x_i^{t+1/2} \cdot \frac{\exp(-\eta \nabla_{\mathbf{x}} f(\mathbf{x}^{t+1/2}, \mathbf{y}^{t+1/2}))_i}{\langle \mathbf{x}^{t+1/2}, \exp(-\eta \nabla_{\mathbf{x}} f(\mathbf{x}^{t+1/2}, \mathbf{y}^{t+1/2})) \rangle} \\ y_i^{t+1} &= y_i^{t+1/2} \cdot \frac{\exp(\eta \nabla_{\mathbf{y}} f(\mathbf{x}^{t+1/2}, \mathbf{y}^{t+1/2}))_i}{\langle \mathbf{y}^{t+1/2}, \exp(\eta \nabla_{\mathbf{y}} f(\mathbf{x}^{t+1/2}, \mathbf{y}^{t+1/2})) \rangle} \end{aligned}$$

end

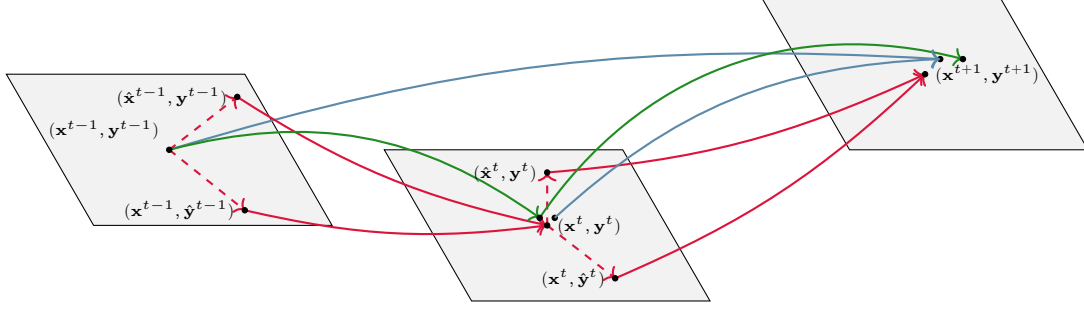


Figure 3.1: Graphical representation for the sequence of iterative approximations of MWU (green), OMWUM (blue) and EG (red) from time step $t - 1$ to time step $t + 1$. The dashed red curves represent the intermediate step of EG.

and provided by Algorithm 2.

3.4 Load-balancing games

In this section, we provide the framework in the case of load balancing games, as defined in [Nisan et al., 2007a], where the tasks selfishly choose to be assigned to machines, in such a way that no task has any incentive to deviate from its machine. Formally:

Definition 9. A load balancing game (LBG) is a tuple $G = \langle k, m, s, w \rangle$, where $k = \{1, \dots, |k|\}$ is the set of tasks, each associated with a weight $w_j \geq 0$, and $m = \{1, \dots, |m|\}$ is the set of machines, each with speed $s_i > 0$.

Here, we consider the case where the tasks play only pure strategies. Under this assumption, the *assignment* of tasks to machines is determined by a mapping $A : k \rightarrow m$ (note that each task is assigned to exactly one machine). The load of machine $i \in m$ under A is defined as $l_i = \sum_{j \in k: A(j)=i} w_j / s_i$. The cost of task j for choosing machine i is $c_j^i = l_i$. Furthermore, the social cost of assignment A is defined as $cost(A) = \max_{i \in m} (l_i)$, in other words the makespan under the assignment A . An assignment A^* is optimal if $cost(A^*) \leq cost(A)$ for all possible assignments A . An assignment A is a pure Nash equilibrium, if and only if, for any j and for any $\hat{i} \in m$, $c_j^{A(j)} \leq c_j^{\hat{i}}$, in other words for any alternative assignment of task j (say to machine \hat{i}) the cost is worse.

3.5 Congestion games

We continue by defining the single-commodity non-atomic congestion game with player-specific costs similar to the approach of [Meir and Parkes, 2015b].

Definition 10. A single-commodity non-atomic congestion game (NACG) with player-specific

costs is a tuple $\Gamma = \langle G, M, l^i, s, t, \mathbf{r} \rangle$, where:

- $G = (V, E)$ is a directed graph,
- M is the set of different types of costs of the players,
- l^i is the set of the non-decreasing, continuous, and non-negative latency cost functions with $l_e^i(x) : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ (one for each edge $e \in E$ and for each cost type $i \in M$ of the players),
- $s \in V$ is the source,
- $t \in V$ is the destination,
- $\mathbf{r} \in \mathbb{R}_{\geq 0}^{|M|}$ is the total mass of flow.

The total mass of flow for players of cost type i is r_i such that $\sum_{i \in M} r_i = r$. We consider that any player of cost type i controls an infinitesimal amount of the flow r_i . In this paper, we assume, without loss of generality, that $r = 1$.

Let P be the set of total paths from s to t , then we define as $g(r_i) \in [0, r]^{|P| \times 1}$ a feasible flow of cost type i of the players routing r_i units of flow on the paths and $g(r) = (g(r_1), \dots, g(r_{|M|}))$ be the tuple of all flows routing $r_1, \dots, r_{|M|}$ units of flow, respectively. We define as $g_p(r_i)$ the flow of players of cost type i that follows the path $p \in P$, $g_p(r)$ the total flow that follows the path $p \in P$ with $\sum_{i \in M} g_p(r_i) = g_p(r)$. The flow of cost type i of the players of the edge $e \in E$ is $g_e(r_i)$ and $g_e(r) = \sum_{i \in M} g_e(r_i)$ is the total flow of the edge $e \in E$.

The cost of following a path p of a player of cost type i is $C_p^i(g_p(r)) = \sum_{e \in p} l_e^i(g_e(r))$ and the total social cost of the flow $g(r)$ is

$$SC(g(r)) = \sum_{i \in M} \sum_{e \in E} g_e(r_i) l_e^i(g_e(r))$$

The socially optimal flow is the feasible flow $g(r)$ such that $SC(g(r))$ is minimum. We now continue with the definition of a *pure Nash Equilibrium*, or in other words *Wardrop Equilibrium*.

Definition 11 (Pure Nash Equilibrium [Nash, 1951]/Wardrop Equilibrium [Wardrop, 1952]). *A pure Nash Equilibrium is a feasible flow $g^*(r)$ such that for any $p, \hat{p} \in P$, and for any $i \in M$,*

$$C_p^i(g_p^*(r)) \leq C_{\hat{p}}^i(g_{\hat{p}}^*(r)),$$

in other words, a flow is an equilibrium if no player has any incentive to deviate from her path.

It has been proved that any non-atomic game with player-specific costs has at least one equilibrium, see Theorem 9 in Appendix B.3.

Chapter 4

Misinformation Games

4.1 Introduction

In this chapter, initially, we provide an intuitive example so that to reveal even better the concept of misinformation games. Consider the classical Prisoner's Dilemma (PD) game [Osborne and Rubinstein, 1994], where two suspects (the players) are being interrogated, having the option to betray the other (B), or stay silent (S). Each of them will get a penalty reduction if he betrays the other, but if they both remain silent, the police can only convict them for lesser charge and not for the principal crime; if they both betray, they will get a reduced penalty for the principal crime. Using classical game theory, this situation is modelled by payoff matrix presented in Table 4.1a, where the only Nash equilibrium is for both players to betray.

Now suppose that the cogent evidence regarding the lesser charge has been obtained in an illegal manner, and thus cannot be used in court. As a result, players' actual payoffs are as shown in Table 4.1b; however, this is not disclosed to the suspects, who still believe that they play under Table 4.1a. This would lead players to betray, although, had they known the truth (Table 4.1b), they also had other options (Nash equilibria), e.g., to both stay silent. We will refer to this game as the *misinformed Prisoner's Dilemma (mPD)* in the rest of this paper.

As shown in Figure 1.2, the main defining characteristic of misinformation games is that agents are unwitting of their misinformation, and will play the game under the misconceived game definition that they have. This essentially means that the assumption of common knowledge is dropped as well.

	S	B		S	B
S	$(-1, -1)$	$(-3, -1/2)$	S	$(0, 0)$	$(-3, -1/2)$
B	$(-1/2, -3)$	$(-2, -2)$	B	$(-1/2, -3)$	$(-2, -2)$

(a) Payoffs (in PD); also, players' view (in mPD). (b) Actual game (in mPD).

Table 4.1: Payoff matrices for the PD and mPD.

Obviously, in such a setting, game theory dictates the *actual player behavior* in her own view, which may be different from the behavior regarding the actual game. On the other hand, *the payoffs received* by the players are the ones provisioned by the actual game, which may differ from the ones they assume.

Therefore we *introduce the formal machinery necessary to study misinformation games*. Specifically, this Chapter is consisted of: i) defining misinformation games and recasting basic game-theoretic concepts without the assumption of common and correct knowledge, ii) introducing a new metric, called the *Price of Misinformation (PoM)*, to quantify the effect of misinformation on the social welfare of players, and iii) applying our ideas to load balancing games and non-atomic congestion games.

4.2 Normal-form games

Misinformation captures the concept that different players may have a different view of the game they play. This leads to the following definition:

Definition 12. A misinformation normal-form game (or simply misinformation game) is a tuple $mG = \langle G^0, G^1, \dots, G^{|N|} \rangle$, where all G^i are normal-form games and G^0 contains $|N|$ players.

In Definition 12, G^0 is called the *actual game* and represents the game that is actually being played, whereas G^i (for $i \in \{1, \dots, |N|\}$) represents the game that player i thinks that is being played (called the *game of player i*). We make no assumptions as to the relation among G^0 and G^i , and allow all types of misinformation (or no misinformation at all) to occur.

We define the following interesting special class of misinformation games:

Definition 13. A misinformation game $mG = \langle G^0, G^1, \dots, G^{|N|} \rangle$ is called *canonical* if and only if:

- For any i , G^0, G^i differ only in their payoffs.
- In any G^i , all players have an equal number of pure strategies.

The first requirement associated with canonical games restricts misinformation to payoffs, and thus avoids the more eccentric cases where misinformation “creates” new, non-existing strategies or players, or cases where players are unaware of certain strategies or players. The second requirement is convenient, since payoff matrices have equal-sized dimensions, an assumption often made (without loss of generality) in standard settings as well.

Although less common, non-canonical misinformation games may occur, e.g., when communication problems, ignorance, or lack of imagination deprives a player from the option to use a viable strategy. However, from a technical perspective, non-canonical misinformation

games are not relevant, because every misinformation game can be transformed into an equivalent game (in terms of its strategic behavior) that is canonical, using the simple process of *inflation* described as follows.

Let mG be a non-canonical misinformation game. To transform it into a canonical misinformation game with the same strategic behavior, we compare G^0 with each G^i ($i > 0$). Then:

1. If G^i does not include a player that appears in G^0 , then we “inflate” G^i by adding this new player, with the strategies that appear in G^0 for this player. We extend the elements of the payoff matrix of G^i to represent the payoffs of the new player, using any fixed constant value. Moreover, the current payoff matrix of G^i is increased by one dimension, by replicating the original payoff matrix as many times as needed (to accommodate the new player’s strategies).
2. If G^i contains an imaginary player not included in G^0 , then we add a new player in G^0 , using the process described in #1 above. In addition, since Definition 12 requires that each player in G^0 is associated with a game in mG , we add a new game in mG , which is a replica of G^0 .
3. If G^i does not contain a certain strategy which appears in G^0 (for a certain player), we just add this new strategy, with payoffs small enough to be dominated by all other strategies.
4. If G^i contains an imaginary strategy that does not appear in G^0 (for a certain player), we inflate G^0 as in #3 above.

Repeating the above process a sufficient (finite) number of times, we will eventually derive a misinformation game that satisfies the first condition of Definition 13 and has the same strategic properties as the original. To satisfy the second condition also, we just inflate the games again according to the largest dimension (number of strategies) of the largest game, using the process described in #3 above. This way, we can always transform a misinformation game into a canonical one with the same strategic properties (the formal proof of this result is quite technical and omitted). Due to this fact, without loss of generality, we will only concern ourselves with canonical misinformation games, thus avoiding the need to study the more eccentric cases.

The definition of misinformed strategies and strategy profiles is straightforward, once noticing that they refer to each player’s own game:

Definition 14. A (pure or mixed) misinformed strategy, $m\sigma_i$ of a player i is a (pure or mixed) strategy of i in the game G^i . We denote the set of all possible misinformed strategies of player i as Σ_i^i .

Definition 15. A misinformed strategy profile of mG is an $|N|$ -tuple of misinformed strategies $m\sigma = (m\sigma_1, \dots, m\sigma_{|N|})$, where $m\sigma_i \in \Sigma_i^i$.

As usual, we denote by $m\sigma_{-i}$ the $|N| - 1$ -tuple strategy profile of all other players except for player i in a misinformed strategy $m\sigma$.

The payoff function h_i of player i under a given profile $m\sigma$ is determined by the payoff matrix of G^0 , and is defined as $h_i : \Sigma_1^1 \times \dots \times \Sigma_{|N|}^{|N|} \rightarrow \mathbb{R}$, such that:

$$h_i(m\sigma_i, m\sigma_{-i}) = \sum_{k \in S_1^1} \dots \sum_{j \in S_{|N|}^{|N|}} P_i^0(k, \dots, j) \cdot m\sigma_{1,k} \cdot \dots \cdot m\sigma_{|N|,j},$$

where $P_i^0(k, l, \dots, j)$ is the payoff of player i in the pure strategy profile (k, l, \dots, j) under the actual game G^0 . Also, S_i^j denotes the set of pure strategies of player i in game G^j .

It is interesting to note that, although each player's strategic decisions are driven by the information in her own game (G^i), the received payoffs are totally dependent on the actual game G^0 , which may be different than G^i . Another important point is that the payoff function would be ill-defined without the assumption that the underlying misinformation game is canonical (because then the strategy profiles of players would have different dimensions).

4.2.1 Equilibria for Misinformation normal-form games

We can now define two alternative notions of equilibria. The first is based on the idea that the players will study their own game, and play one of their Nash strategies, without regards to what other players know or play. Formally:

Definition 16 (Natural misinformed equilibrium). A misinformed strategy, $m\sigma_i$, of player i , is a misinformed equilibrium strategy, if and only if, it is in a Nash equilibrium strategy profile in game G^i . A misinformed strategy profile $m\sigma$ is called a natural misinformed equilibrium if it consists of misinformed equilibrium strategies.

In the following, we denote by $NME(mG)$ (or simply NME , when mG is obvious from the context) the set of natural misinformed equilibria of mG .

The second type of equilibrium is closer to the idea of the Nash equilibrium in normal-form games. It is based on the notion that the players will choose what is best according to their own game, given the actual strategies of the other players (i.e., best response), although such strategies may seem sub-optimal from the player's perspective.

To define this, set h_i^i the payoff function that player i expects to obtain in her game. Formally, f_i^i is defined as $h_i^i : \Sigma_1^i \times \dots \times \Sigma_{|N|}^i \rightarrow \mathbb{R}$, such that:

$$h_i^i(m\sigma_i, m\sigma_{-i}) = \sum_{k \in S_1^i} \dots \sum_{j \in S_{|N|}^i} P_i^i(k, \dots, j) \cdot m\sigma_{1,k} \cdot \dots \cdot m\sigma_{|N|,j}, \quad (4.1)$$

where Σ_j^i denotes the set of all possible mixed strategies of player j according to the game of player i (G^i).

Now we are ready to define the second type of equilibrium as described in the previous paragraph:

Definition 17 (Pseudo misinformed equilibrium). *A misinformed strategy profile $m\sigma^* = (m\sigma_1^*, \dots, m\sigma_{|N|}^*)$ is a pseudo misinformed equilibrium, if and only if, for any i and for any misinformed strategy $m\hat{\sigma}_i$,*

$$h_i^i(m\sigma_i^*, m\sigma_{-i}^*) \geq h_i^i(m\hat{\sigma}_i, m\sigma_{-i}^*)$$

Computing misinformed equilibria is straightforward given the respective results in game theory, as one needs to simply compute the equilibria of each G^i and compose the equilibrium strategies of player i in G^i for all i . However, for pseudo misinformed equilibria, things are not as obvious. The following notion will be helpful in this respect:

Definition 18. *Consider the canonical misinformation game $mG = \langle G^0, G^1, \dots, G^{|N|} \rangle$. Then, we call the merged game of mG the game $\tilde{G} = \langle N, S, P \rangle$, such that N is the set of players in G^0 , S is the set of strategies in G^0 and $P = (P_1^1, \dots, P_{|N|}^{|N|})$ (where P_i^i represents the payoffs of player i in game G^i).*

Essentially, the merged game is created by the agglomeration of the different G^i , such that the payoffs of player i are taken from G^i . We can easily note that the pseudo misinformed equilibria of a misinformed game mG coincide with the Nash equilibria of its merged game. This is immediate from the fact that the computation of a pseudo misinformed equilibrium in Definition 17 depends only on the strategies of each player in her own game:

Proposition 1. *Consider a canonical misinformation game mG and its merged game \tilde{G} . Then a strategy profile is a pseudo misinformed equilibrium of mG , if and only if, it is a Nash equilibrium of \tilde{G} .*

An important corollary of Proposition 1 (and [Nash, 1951]) is that a pseudo misinformed equilibrium always exists. The same is of course true for the natural misinformed equilibrium (because each game G^i has at least one Nash equilibrium, by [Nash, 1951]):

Proposition 2. *Any canonical misinformation game has at least one natural misinformed equilibrium, and at least one pseudo misinformed equilibrium.*

Clearly, the misinformed equilibria and the pseudo misinformed equilibria of a misinformation game need not coincide. The following example shows this:

Example 4.1. *Consider the canonical misinformation game $mG = \langle G^0, G^1, G^2 \rangle$ with payoffs:*

$$(R^0, C^0) := \begin{pmatrix} (2, 1) & (0, 0) \\ (0, 0) & (1, 2) \end{pmatrix}$$

$$(R^1, C^1) := \begin{pmatrix} (0,0) & (2,1) \\ (1,2) & (0,0) \end{pmatrix}, \quad (R^2, C^2) = (R^0, C^0).$$

Note that G^0 (and G^2) is the well-known Battle of the Sexes game, whereas G^1 is the same game whose payoff matrix has been rotated by 90° . The Nash equilibria of G^2 are $((1,0), (1,0))$, $((0,1), (0,1))$ and $((2/3, 1/3), (1/3, 2/3))$, whereas the Nash equilibria of G^1 are $((1,0), (0,1))$, $((0,1), (1,0))$ and $((2/3, 1/3), (1/3, 2/3))$. Thus the equilibrium strategies of row player in G^1 are $A = \{(0,1), (1,0), (2/3, 1/3)\}$, and the equilibrium strategies of column player in G^2 are $B = \{(1,0), (0,1), (2/3, 1/3)\}$. According to Definition 16 the misinformed equilibria of mG are the elements of $A \times B$. On the other hand, mG has one pseudo misinformed equilibrium, namely $((2/3, 1/3), (1/3, 2/3))$, as given by the merged game, whose payoff matrix is (R^1, C^2) . \square

It is also easy to see that the computation of natural/pseudo misinformed equilibria has the same computational complexity as the computation of a Nash equilibrium in a normal-form game. Indeed, the computation of a natural misinformed equilibrium amounts to the computation of the Nash equilibrium of $|N|$ different games (the games of the players, i.e., $G^1, \dots, G^{|N|}$), whereas the computation of the pseudo misinformed equilibrium is performed by computing the Nash equilibrium of the merged game. Thus, using [Chen et al., 2009, Daskalakis et al., 2009]:

Proposition 3. *The computation of a misinformed equilibrium and of a pseudo misinformed equilibrium of a misinformation game is **PPAD**-complete.*

4.2.2 Price of Misinformation

Inspired by Definition 6, we define a metric to measure the effect of misinformation compared to the social optimum, based on a social welfare function h .

Definition 19 (see [Varsos et al., 2021]). *Given a misinformation game mG , the Price of Misinformation (PoM) is defined as:*

$$PoM = \frac{h(opt)}{\min_{\sigma \in NME} h(\sigma)} \quad (4.2)$$

Using the definition of PoA [Koutsoupias and Papadimitriou, 1999] and (4.2) we derive the following formula that links the two metrics:

$$\frac{PoM}{PoA} = \frac{\min_{\sigma \in NE} h(\sigma)}{\min_{\sigma \in NME} h(\sigma)} \quad (4.3)$$

for NE the Nash equilibria of G^0 and mE the misinformed equilibria of mG .

Given an $mG = \langle G^0, G^1, \dots, G^{|N|} \rangle$, interesting results can be derived by comparing the optimal Nash equilibrium of G^0 with the optimal natural misinformed equilibrium of mG (or, equivalently, through equation 4.3, by comparing PoA of G^0 with PoM of mG). If $PoM < PoA$, then misinformation has a beneficial effect on social welfare, as the players are inclined (due to their misinformation) to choose socially better strategies. On the other hand, if $PoM > PoA$, then misinformation leads to a worse outcome, from the perspective of social welfare.

The following proposition essentially says that misinformation can steer the players to any desired behaviour. In other words, regardless of the actual game G , we can use misinformation to make the players behave as if they were playing any other game G' , i.e., enforce any desired behaviour:

Proposition 4. *For any $G = \langle N, S, P \rangle$, $G' = \langle N, S, P' \rangle$, we have:*

1. *There is a misinformation game $mG = \langle G^0, G^1, \dots, G^{|N|} \rangle$ such that $G^0 = G$ and the set of natural misinformed equilibrium strategies of player i in mG are identical to the Nash equilibrium strategies of i in G' , $\forall i \in N$.*
2. *There is a misinformation game $mG = \langle G^0, G^1, \dots, G^{|N|} \rangle$ such that $G^0 = G$ and the set of pseudo misinformed equilibria of mG are identical to the Nash equilibria of G' .*

Proof. To show both cases, take $mG = \langle G, G', \dots, G' \rangle$. □

Proposition 4 shows the power of misinformation as a tool for mechanism design. Similarly, we can show that we can use misinformation to enforce any given behaviour to the players of a game, including the socially optimal one. The proof is easy, using Proposition 4, by showing that there exists a normal-form game G' whose only Nash equilibrium is the desired one:

Proposition 5. *For any normal-form game G and strategy profile σ :*

1. *There is a misinformation game $mG = \langle G^0, G^1, \dots, G^{|N|} \rangle$ such that $G^0 = G$ and the only natural misinformed equilibrium of mG is σ .*
2. *There is a misinformation game $mG = \langle G^0, G^1, \dots, G^{|N|} \rangle$ such that $G^0 = G$ and the only pseudo misinformed equilibrium of mG is σ .*

Corollary 1. *For every normal-form game G there is a misinformation game $mG = \langle G^0, G^1, \dots, G^{|N|} \rangle$ such that $G^0 = G$ and $PoM = 1$.*

The above results show that, given sufficient misinformation, anything is possible in terms of improving (or deteriorating) the social welfare. An interesting follow-up question is to explore the limitations of mechanism design using misinformation, if we impose restrictions on the type or amount of misinformation possible.

4.3 Misinformation in load balancing games

Introducing misinformation in LBGs follows similar patterns as in Section 4.2:

Definition 20. A misinformation LBG is a tuple $mG = \langle G^0, G^1, \dots, G^{|k|} \rangle$, where all G^j are LBGs and G^0 contains $|k|$ tasks.

As before, G^0 is called the *actual LBG*, whereas G^j is the LBG that task j understands. The notion of canonical misinformation games applies here as well:

Definition 21. A misinformation LBG $mG = \langle G^0, G^1, \dots, G^{|k|} \rangle$ is called *canonical*, if and only if, for any j , G^0, G^j differ only with regards to the weights of the tasks and the speeds of the machines.

All misinformation games can be transformed into canonical through the process of inflation described in Subsection 4.2: we compare each G^j with G^0 , and, when tasks are missing, we add tasks with 0 weight; when machines are missing, we add machines with a speed small enough so that a task does not have any incentive to move to the new machines, even if all tasks are assigned to the slowest machine. Therefore, in the following, we only consider canonical misinformation LBGs.

Like in the standard case, a misinformed assignment mA is a mapping of tasks to machines $mA : k \rightarrow m$, where any task j chooses a machine according to its game G^j . Given a specific misinformed assignment mA , the *actual load of a machine i* is $l_i^0 = \sum_{j \in k: i=mA(j)} w_j^0 / s_i^0$, whereas the *perceived load of a machine i for task h* is $l_i^h = \sum_{j \in k: i=mA(j)} w_j^h / s_i^h$. The *actual cost* of task j for choosing machine i is $c_j^{i,0} = l_i^0$, whereas the *perceived cost* is $c_j^{i,j} = l_i^j$. Similarly, the actual social cost of mA is $cost(mA) = \max_{i \in m} (l_i^0)$.

4.3.1 Equilibria in Misinformation load balancing games

As mentioned above, equilibria are achieved when tasks have no incentive to change their assignments. As with normal-form games, this can be formalised in two ways: either the tasks choose the Nash equilibrium assignments in their own game (without regards to what other tasks do), or they choose to adapt their behaviour based on the other assignments. Formally:

Definition 22. A *misinformed task assignment* $mA(j)$ of task j is a pure misinformed equilibrium task assignment, if and only if it is a pure Nash equilibrium assignment for game G^j . A *misinformed assignment* mA is called a *pure misinformed equilibrium assignment* if and only if it consists of pure misinformed equilibrium task assignments.

Definition 23. A *misinformed assignment* mA is a *pure pseudo misinformed equilibrium assignment*, if and only if, $\forall j \in k \ \forall i \in m^j : c_j^{mA(j),j} \leq c_j^{i,j}$.

As each G^j is an LBG, the existence of a pure Nash equilibrium assignment in every G^j is warranted by the results of [Nisan et al., 2007a, Rosenthal, 1973, Nash, 1951], thus a

misinformed equilibrium assignment in misinformation LBGs always exists. Moreover, using complexity results for standard LBGs [Nisan et al., 2007a], we can show the following:

Proposition 6. *Consider a misinformation LBG mG with k tasks, such that each G^j has m identical machines. Then, the computational complexity of computing a misinformed equilibrium assignment in mG is $O(k^2 \log k)$.*

Proof. On identical machines we can transform any assignment A into a pure Nash equilibrium in time $O(k \log k)$ [Nisan et al., 2007a]. To find a misinformed equilibrium, we repeat this once for each G^j ($j > 0$), which requires $O(k^2 \log k)$ time. \square

Unfortunately, unlike pure misinformed equilibria, a pure pseudo misinformed equilibrium is not guaranteed to exist. This is a corollary of the following counter-example:

Example 4.2. *Consider a misinformation LBG with 3 tasks and 2 identical machines with misinformation only in the weights of the tasks. Also assume that task 1 knows the weights $w^1 = (w_1^1 = 1, w_2^1 = 10, w_3^1 = 100)$, task 2 knows the weights $w^2 = (w_1^2 = 100, w_2^2 = 1, w_3^2 = 10)$ and task 3 knows the weights $w^3 = (w_1^3 = 10, w_2^3 = 100, w_3^3 = 1)$. It is easy to see that there is no assignment such that Definition 23 holds, i.e., in any assignment, at least one of the tasks has an incentive to change machine, according to its own game.* \square

As Example 4.2 shows, the existence of a pure pseudo misinformed equilibrium assignment is not warranted. Thus, although a pure Nash equilibrium is guaranteed to exist in LBGs, this is not the case for pure pseudo misinformed equilibria in misinformation LBGs, not even if we assume identical machines and misinformation restricted to weights only. As another corollary of this result, we get that the “merged LBG” (i.e., the counterpart of the merged game – Definition 18) does not exist in the general case. An interesting relevant question (reserved for future work) would be to discover the conditions that allow the existence of a pure pseudo misinformed equilibrium assignment and/or the “merged LBG”.

Note that, since LBGs are special cases of normal-form games, one could transform a misinformed LBG into a misinformed normal-form game, and compute the merged game from there. The above counter-example shows that this process does not always result to a (merged) game that has an LBG counterpart (because if it did, it would have a pure Nash equilibrium, and thus the original misinformation LBG would have a pure pseudo misinformed equilibrium).

4.3.2 Price of Misinformation in load balancing games

The Price of Misinformation (PoM) in misinformation LBGs is defined analogously to Definition 27, but note that here we talk about a minimisation game:

$$PoM = \frac{\max_{A \in NME} \text{cost}(A)}{\text{cost}(A^*)}, \quad (4.4)$$

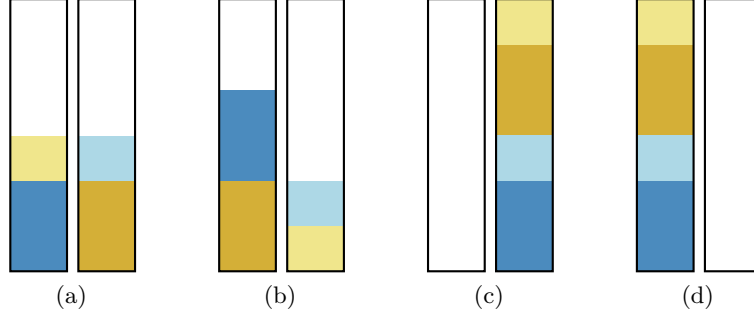


Figure 4.1: Load balancing game and misinformed load balancing game: (a) optimal assignment, (b) worst Nash equilibrium allocation, (c-d) worst natural misinformed equilibrium allocation.

where $\text{cost}(A)$ is the worst cost among the pure misinformed equilibria assignments and $\text{cost}(A^*)$ is the cost of the optimal assignment in the actual game.

The following example is illustrative of the concepts presented in this section:

Example 4.3. Suppose that there are two identical machines with speed $s = 1$ and four tasks with $w_1 = w_2 = 1$ and $w_3 = w_4 = 2$. The optimal assignment maps a task of weight 1 and a task of weight 2 to each of the machines ($A^* = (1, 2, 1, 2)$).

It is obvious that the worst pure Nash equilibrium assignment is $A = (1, 1, 2, 2)$ with $\text{cost}(A) = 4$, Figure 4.1-(b).

Now, consider the misinformation game mG in which tasks have different information on the weights. Let $w^1 = (w_1^1 = 6, w_2^1 = 1, w_3^1 = 2, w_4^1 = 2)$ be the weights in G^1 and $w^j = (w_1^j = 7, w_2^j = 1, w_3^j = 1, w_4^j = 1)$ in G^j , for $j = \{2, 3, 4\}$. The pure Nash equilibrium assignments in each game G^j are $A_1 = (1, 2, 2, 2)$ and $A_2 = (2, 1, 1, 1)$, thus the pure misinformed equilibrium assignments are all combinations aligned with i) task 1 is assigned to a different machine than tasks $\{2, 3, 4\}$ or ii) all tasks are assigned to the same machine. From the above, the worst misinformed equilibrium assignment is derived to be $mA = (1, 1, 1, 1)$ (or $mA = (2, 2, 2, 2)$) with $\text{cost}(mA) = 6$.

On the other hand, the pure pseudo misinformed equilibrium assignments are those in which task 1 is assigned to a different machine than tasks $\{2, 3, 4\}$, Figure 4.1-(c-d). It is interesting that in this example $\text{PoA} = 4/3$ and $\text{PoM} = 2$ implying that misinformation worsens the behaviour of the game. \square

Misinformation, as a tool for mechanism design, is equally strong and flexible for LBGs as for normal-form games. In particular, propositions analogous to Propositions 4, 5 hold:

Proposition 7. For any $G = \langle k, m, s, w \rangle$, $G' = \langle k, m, s', w' \rangle$, we have:

1. There is a misinformation LBG $mG = \langle G^0, G^1, \dots, G^{|N|} \rangle$ such that $G^0 = G$ and the set of pure misinformed equilibrium assignments of player i in mG are identical to the pure Nash equilibrium assignments of i in G' , $i \in N$.
2. There is a misinformation LBG $mG = \langle G^0, G^1, \dots, G^{|N|} \rangle$ such that $G^0 = G$ and the set of pure pseudo misinformed equilibrium assignments of mG are identical to the pure Nash equilibrium assignments of G' .

Proof. To show both cases, take $mG = \langle G, G', \dots, G' \rangle$. □

Proposition 8. For any LBG G and assignment A :

1. There is a misinformation LBG $mG = \langle G^0, G^1, \dots, G^{|N|} \rangle$ such that $G^0 = G$ and the only pure misinformed equilibrium assignment of mG is A .
2. There is a misinformation LBG $mG = \langle G^0, G^1, \dots, G^{|N|} \rangle$ such that $G^0 = G$ and the only pure pseudo misinformed equilibrium assignment of mG is A .

Proof. For each task j , such that $A(j) = i$, we create a game G^j with sufficiently high speed for i and sufficiently low speed for all machines $i' \neq i$. We now create $mG = \langle G, G^1, \dots, G^{|k|} \rangle$. □

Corollary 2. For every LBG G there is a misinformation LBG $mG = \langle G^0, G^1, \dots, G^{|k|} \rangle$ such that $G^0 = G$ and $PoM = 1$.

Due to the special form of LBGs, we can prove various bounds regarding their cost and PoM, based on the task weights and machine speeds. Propositions 9, 10, 11 show some such results:

Proposition 9. Consider a canonical misinformation LBG $mG = \langle G^0, G^1, \dots, G^{|k|} \rangle$, such that $G^0 = \langle k, m, s, w \rangle$ and $s_i > 0$ for all i . Then, for any assignment mA , $cost(mA) \leq \sum_{j=1}^k w_j / \min_i s_i$.

Proof. The worst possible assignment mA^* (from the social cost perspective) is to assign all tasks to the slowest machine, with $cost(mA^*) = \sum_{j=1}^k w_j / \min_i s_i$. Misinformation can achieve this effect, so the result follows. □

Proposition 10. Consider a misinformation LBG $mG = \langle G^0, G^1, \dots, G^{|k|} \rangle$, such that G^0 has m identical machines and finite task weights. Then, the Price of Misinformation is $PoM \leq m$.

Proof. We know that the cost of an optimal assignment $cost(A^*)$ cannot be smaller than the average load over all machines (i.e., $(\sum_{j \in [k]} w_j) / m$). Also, the worst scenario is that all tasks are assigned into one machine, with cost $(\sum_{j \in [k]} w_j)$. Then, using Equation 4.4, we conclude. □

Next, we consider the case of uniformly related machines. We can show the following:

Proposition 11. *Consider a misinformation LBG $mG = \langle G^0, G^1, \dots, G^{|k|} \rangle$, such that $G^0 = \langle k, m, s, w \rangle$ with m uniformly related machines and finite task weights. Then, the Price of Misinformation is*

$$PoM \leq k \cdot \frac{S}{s} \cdot O\left(\frac{\log m}{\log \log m}\right), \quad (4.5)$$

where s is the slowest speed and S is the fastest speed.

Proof. Since there is the case that all tasks be assigned to the slowest machine we have that $\text{cost}(mA) \leq \sum_{i=1}^k w_i/s \leq k \cdot M/s$, where M is the largest weight.

Also, we have that

$$PoM = PoA \cdot \frac{\max_{A \in mA} \text{Cost}(A)}{\max_{B \in NE} \text{Cost}(B)}$$

with A be the worst misinformed equilibrium assignment and B the worst Nash equilibrium assignment. Furthermore, we have that $\max_{B \in NE} \text{Cost}(B) \geq M/S$. Finally, by Chapter 20 of [Nisan et al., 2007a] we have that $PoA \leq O\left(\frac{\log m}{\log \log m}\right)$. \square

4.4 Misinformation in non-atomic congestion games

In this Section, we use an analogous approach as in Section 4.3 to define misinformation games for the case of non-atomic congestion games, where each player has a subjective view about the game he plays, which may be different from the others. Formally:

Definition 24 (Misinformation game [Varsos et al., 2022]). *A misinformation non-atomic congestion game $m\Gamma$ with θ splitting is an $(N+1)$ -tuple $m\Gamma = \langle \Gamma^0, \Gamma^1, \dots, \Gamma^N \rangle$, where N is the number of different views of the game that different players may assume, $\Gamma^0 = \langle G, l, s, t, r \rangle$ is the actual game, $\Gamma^j = \langle G, l^j, s, t, r^j \rangle$ are the different subjective game specifications assumed by the players, of which each player assumes only one, and $\theta = \langle \theta^1, \dots, \theta^N \rangle$, where θ^i is the portion of players that experience view Γ^i .*

Here, we assume that the total mass of flow across all Γ^j (for $j > 0$) is equal to the respective mass in Γ^0 , $r^j = r$. Further, it must hold that $\sum_{i \in [N]} \theta^i = 1$.

Thus, the players have the correct view of the graph and the flow at hand, although they may assume different cost functions. In this case we call the misinformation single commodity non-atomic congestion games as *canonical*. This is analogous to the concept of canonical misinformation games, as defined in Definition 13.

Definition 25 (Misinformed equilibrium strategy). *A misinformed strategy is a flow for portion θ^j , that θ^j with subjective view Γ^j follows in a pure Nash equilibrium strategy of its game view Γ^j .*

Definition 26 (Natural Misinformed equilibrium). *A natural misinformed equilibrium (NME) is a flow f such that each portion θ^j plays a misinformed equilibrium strategy according to its game-specific view j .*

Since any non-atomic congestion game has at least one Nash equilibrium, it is easy to see that any misinformation game of a non-atomic congestion game as defined above has at least one natural misinformed equilibrium.

Having at hand the formal definition of the natural misinformed equilibrium, we measure the deterioration/leverage in efficiency of a non-atomic congestion game due to misinformation. For that, we adapt Definition 27 and the concept of Price of Misinformation (PoM) in the case of non-atomic congestion games becomes:

Definition 27 (Price of Misinformation). *Given a misinformation congestion game, the Price of Misinformation (PoM) is defined as*

$$PoM = \frac{\max_{f_{NME} \in NME} SC(f_{NME})}{SC(f_{opt})}, \quad (4.6)$$

where f_{opt} is the flow that minimizes the Social Cost in the actual game Γ^0 and the nominator is the worst (maximum) value of the Social Cost of the set NME as computed with regards to the actual game.

We can show the following:

Proposition 12. *For every misinformation non-atomic congestion game, we have that:*

$$1 \leq PoM \leq \left(r \cdot \max_{p \in P} C_p(r) \right) / opt \quad (4.7)$$

Proof. In the worst case, all flow will be routed through the most costly routes, which leads to a Social Cost of $r \cdot \max_{p \in P} C_p(r)$. Thus, $PoM \leq \left(r \cdot \max_{p \in P} C_p(r) \right) / opt$. Moreover, $PoM \geq 1$ by definition. \square

Note that when $\Gamma^0 = \Gamma^j$ for any j , then PoM coincides with the PoA . Using the definition of PoA and (4.6), we can derive the following formula that links the two metrics in the general case:

$$PoM = PoA \cdot \left(\frac{\max_{f_{NME} \in NME} SC(f_{NME})}{\max_{f_{NE} \in NE} SC(f_{NE})} \right), \quad (4.8)$$

where NE is the set of all Nash equilibria of the actual game Γ^0 .

Interesting results can be derived by comparing the worst Nash equilibrium of Γ^0 (or PoA of Γ^0) with the worst natural misinformed equilibrium of $m\Gamma$ (or PoM of $m\Gamma$). If

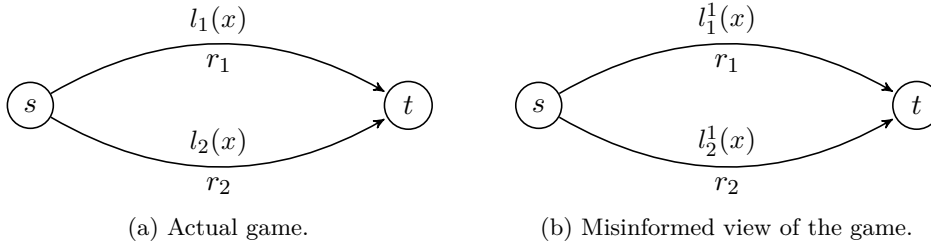


Figure 4.2: Non-atomic congestion game with 2 parallel links.

$PoM < PoA$, then misinformation has a beneficial effect on social welfare, as players are inclined (due to their misinformation) to choose socially better strategies. On the other hand, if $PoM > PoA$, then misinformation leads to a worse outcome from the perspective of social welfare.

A metric similar to PoM was introduced in [Meir and Parkes, 2015b], called the *Biased Price of Anarchy*, which measures the ratio of the equilibrium under biases in knowledge compared to the optimal outcome. In this concept, all players play a game with modified costs and, thus, (possibly) different than the actual costs. In our concept, all players play a game according to the misinformation that they assume that is the same for anyone resulting to different outcomes, so in general the two concepts PoM and *Biased Price of Anarchy* are not equal.

Next, we provide an illustrative example of the above concepts.

Example 4.4. We consider the non-atomic congestion game as depicted in Figure 4.2a (known as *Pigou network* [Pigou, 1920]), with latency functions $l_1(x) = \varepsilon x + 1$, $l_2(x) = x + \varepsilon$, $r = 1$ and $x \in [0, 1]$. It is clear that selfish players in a pure Nash equilibrium will choose all to route through r_2 , resulting to a Social Cost equal to $1 + \varepsilon \approx 1$, as we take arbitrarily small $\varepsilon > 0$. On the other hand, the social optimum can be achieved by allocating the flow as follows: $\approx 1/2$ through route r_1 and $\approx 1/2$ through route r_2 . Thus, the optimal cost is $SC(f_{opt}) \approx 3/4$, and the Price of Anarchy is $PoA \approx 4/3$.

Now, consider the actual game Γ^0 as depicted in Figure 4.2a and the game Γ^1 as depicted in Figure 4.2b. Also, assume the misinformation game $m\Gamma = \langle \Gamma^0, \Gamma^1, \Gamma^2 \rangle$ with θ splitting, where $\Gamma^0 = \langle G, l, s, t, r \rangle$, $\Gamma^1 = \langle G, l^1, s, t, r^1 \rangle$, and $\Gamma^2 = \langle G, l^2, s, t, r^2 \rangle$, and $\theta = \langle \theta^1, \theta^2 \rangle$. Further, $\theta^1 = 2/3$ of the players have the view Γ^1 , and the rest the view Γ^2 .

In this example, $l^2 = l^0 = l$, $l_1^1(x) = \varepsilon x + 1$, and $l_2^1(x) = x + 1$. In Γ^1 the equilibrium is to choose the route r_1 , while in Γ^2 the equilibrium is to route through r_2 . Thus, the natural misinformed equilibrium is that $\approx 2/3$ of the players choose r_1 and the remaining $\approx 1/3$ choose r_2 , leading to $SC_{NME}(f_{NME}) \approx 7/9$ in the $m\Gamma$. So, the Price of Misinformation for this game is $PoM \approx 28/27 < PoA$. Thus, the players improve their behaviour, despite the selfishness and the misinformation.

Chapter 5

Epistemic Adaptive Evolution

5.1 Introduction

In the last chapter we introduce the concept of misinformation games; without considering the *reaction* of the players upon their realization that the received payoffs are different than expected. In other words, our model until now addresses only one-shot games. To cope with this limitation, it is important to consider the reaction of players as they observe unexpected outcomes or (seemingly) irrational decisions made by other players.

To address these issues, initially, we develop an iterative methodology, called the *Adaptation Procedure*, which models the evolution of the strategic behaviour of rational players in a misinformation game, as they obtain new information and update their (erroneous) game specifications. We consider the following setting: time is discrete and players take an action in each period. They update their game specifications according to the payoff they receive. This new information may lead them to a different strategic choice in the next period, so the procedure is iterative and stabilizes when the players do not deviate from their current choices, based on what they know so far. Note that this does not necessarily mean that all players have the correct game specification, i.e., the procedure may stabilize before the players get the chance to update their entire payoff matrix.

Nevertheless, the reactive nature of the players does not mean that they make their decisions in a mechanistic manner only. The most effort for a decision to be made requires a thorough understanding and reasoning regarding the players and their perspective about the interaction. This depends on the information that the player possesses and the information that he/she gains from the opponents, e.g. observing the decisions they made. In the scope of misinformation games, the players have (possible) incorrect subjective information. Hence, each observation (possible) contradicts their beliefs about the undergoing interaction. Inevitably, we need to enrich Adaptation Procedure so that to capture revisions in the beliefs of the players.

For that we establish the subjective epistemic views of the interaction that each player has, and we call it *epistemic misinformation games*. Afterwards we develop an iterative procedure that captures the changes in the epistemic views of the players due to the new information

or unexpected choices, we call this procedure *Epistemic Adaptive Evolution*, and marginally coincides with Adaptation Procedure.

5.2 The Adaptation Procedure

We first provide an informal description of the adaptation process and the related assumptions.

5.2.1 Informal Description

The adaptation begins with a finite misinformation game, say $mG^{(0)}$, where each player has his/her own subjective view of the interaction which may differ from the actual one. As explained in Chapter 4, this will cause each player to employ one of the equilibrium strategies in his/her own game, leading to the emergence of a natural misinformed equilibrium. The payoff received from the players' combined strategic choices will be provided by the actual game, and this may be different from what each player knows in his/her subjective game. Note that we assume that the payoffs received by each player for their strategic choices are publicly announced, thus are common knowledge. As a result, players will update their payoff matrices by replacing the erroneous payoffs with the correct ones just received, leading to a new misinformation game.

It should be noted that the above process is not, in general linear. When the misinformation game has more than one natural misinformed equilibria, and/or when there exist mixed strategic choices in them, each of these choices will be considered in a separate branch of the process. As a result, $mG^{(0)}$ will in fact spawn several new misinformation games, one for each element of the support of the natural misinformed equilibria.

The process continues recursively for each branch, creating new misinformation games. When no new misinformation games are spawned (i.e., all misinformation games spawned will already be in the tree), the Adaptation Procedure *terminates*.

Observe that the Adaptation Procedure produces new games (and thus new *nmes*) in each time step. The *nmes* of the games appearing in the leaves of the recursive tree at the time when the Adaptation Procedure terminates, determine a new equilibrium concept. Specifically, a strategy profile that is an *nme* of one of the leaves is a *stable misinformed equilibrium*, if the players choosing it have no incentive to deviate from it, even in the presence of the updated payoff information.

5.2.2 Formal Definition

Consider a multidimensional matrix A , and a vector $\vec{v} \in Pos$ of matrix A . We denote by $A_{\vec{v}}$ the element of A in position \vec{v} . For example, $A_{(1,2)}$ is the top right element of a 2×2 matrix A . We define the operation of replacement of element $A_{\vec{v}}$ with b as follows.

Definition 28. Consider set F , matrix $A \in F^{n_1 \times n_2 \times \dots \times n_m}$, vector \vec{v} indicating a position in A and some $b \in F$. We denote by $A \oplus_{\vec{v}} b$ the matrix $B \in F^{n_1 \times n_2 \times \dots \times n_m}$, such that $B_{\vec{v}} = b$ and $B_{\vec{u}} = A_{\vec{u}}$ for all $\vec{u} \neq \vec{v}$.

As explained above, the Adaptation Procedure is based on the replacement of an element in the subjective payoff matrices of players, with the respective element in the actual payoff matrix. This is formalised below:

Definition 29. Consider a canonical misinformation game $mG = \langle G^0, G^1, \dots, G^{|N|} \rangle$, where $G^i = \langle N, S, P^i \rangle$ (for $0 \leq i \leq |N|$), and some vector \vec{v} . We define the \vec{v} -update of mG , denoted by $mG_{\vec{v}}$, to be the misinformation game $\langle G^0, G^{1'}, \dots, G^{|N|'} \rangle$, where $G^{i'} = \langle N, S, P^i \oplus_{\vec{v}} P_{\vec{v}}^0 \rangle$, for $1 \leq i \leq |N|$.

Definition 29 tells us how to perform the update process that the Adaptation Procedure requires.

It is easy to show that the following hold:

- $(mG_{\vec{u}_1})_{\vec{u}_1} = mG_{\vec{u}_1}$
- $(mG_{\vec{u}_1})_{\vec{u}_2} = (mG_{\vec{u}_2})_{\vec{u}_1}$

Abusing notation, for a set of positions $X = \{\vec{u}_1, \dots, \vec{u}_k\}$, we denote by mG_X the game $mG_X = (\dots (mG_{\vec{u}_1})_{\vec{u}_2} \dots)_{\vec{u}_k}$. Given the properties above, the notation mG_X is well-defined.

The position where the update takes place (denoted by \vec{v} in Definition 29) is determined by the strategic choices of the players, and can be “extracted” using the following definition:

Definition 30. Consider a strategy profile $\sigma = (\sigma_1, \dots, \sigma_N)$ with $\sigma_i \in \mathbb{R}^{|S_i|}$ and $S_i = \{s_{i1}, \dots, s_{i|S_i|}\}$. The characteristic strategy set of vectors of σ is $\chi(\sigma) = \chi(\text{supp}(\sigma_1)) \times \dots \times \chi(\text{supp}(\sigma_N))$, with $\chi(\text{supp}(\sigma_j)) = \{i | s_{ji} \in \text{supp}(\sigma_j)\}$.

This definition can be clarified with the following example:

Example 5.1. Assume a 4×3 bimatrix game. Then, the characteristic strategy set of vectors of $\sigma = ((1/2, 0, 1/3, 1/6), (0, 0, 1))$ is $\chi(\sigma) = \{(1, 3), (3, 3), (4, 3)\}$. \square

As explained above, the Adaptation Procedure occurs in discrete time steps $t \in \mathbb{N}_0 = \mathbb{N} \cup \{0\}$. It starts from $t = 0$ where player i has the view $G^{i,(0)}$, $\forall i \in [|N|]$, and in each time step t we implement the update operation described in Definition 29 for the vector(s) that correspond to the strategic choices of the players. The following example illustrates this procedure using the above notions, and is also visualised in Figure 5.2:

Example 5.2 (Running Example). Consider the canonical misinformation game $mG^{(t)} = \langle G^0, G^{1,(t)}, G^{2,(t)} \rangle$, where $G^{i,(t)} = \langle \{1, 2\}, S = \{s_1, s_2\}, P^{i,(t)} \rangle$, with $i \in \{1, 2\}$ and

$$P^0 = P^{1,(0)} = \begin{pmatrix} (2, 2) & (0, 3) \\ (3, 0) & (1, 1) \end{pmatrix}, \quad P^{2,(0)} = \begin{pmatrix} (1, 1) & (3, 0) \\ (0, 3) & (2, 2) \end{pmatrix}$$

In $t = 0$, player 1 has equilibrium strategy s_2 in $G^{1,(0)}$, while player 2 has equilibrium strategy s_1 in $G^{2,(0)}$. Thus the nme corresponds to (s_2, s_1) and has strategy profile $((0, 1), (1, 0))$. Using the characteristic strategy vector we take $\chi((0, 1), (1, 0)) = \{(2, 1)\}$.

The update operator gives $mG^{(1)} = \langle G^0, G^{1,(1)}, G^{2,(1)} \rangle$ with the following payoff matrices (note how the bottom-left payoff has been updated):

$$P^{1,(1)} = \begin{pmatrix} (2, 2) & (0, 3) \\ (3, 0) & (1, 1) \end{pmatrix}, \quad P^{2,(1)} = \begin{pmatrix} (1, 1) & (3, 0) \\ (3, 0) & (2, 2) \end{pmatrix}$$

■

The procedure shown in Example 5.2 is formalised as follows, taking into account the fact that the process may branch when $\chi(\sigma)$ is not a singleton set:

Definition 31. For a set M of misinformation games, we set:

$$\mathcal{AD}(M) = \{mG_{\vec{u}} \mid mG \in M, \vec{u} \in \chi(\sigma), \sigma \in NME(mG)\}$$

Thus, we define as Adaptation Procedure the iterative process that

$$\begin{cases} \mathcal{AD}^{(0)}(M) = M \\ \mathcal{AD}^{(t+1)}(M) = \mathcal{AD}^{(t)}(\mathcal{AD}(M)) \end{cases}$$

for $t \in \mathbb{N}_0$.

The functionality of the Adaptation Procedure between two consecutive time steps t and $t + 1$, as provided by Definition 31, is depicted in Figure 5.4.

Note that the Adaptation Procedure is defined over a union of a set of misinformation games. Although our intent is basically to apply it over a single misinformation game, the branching process, along with the recursive nature of the definition, forces us to consider the more general case right from the start. Note also that we will often abuse notation and write $\mathcal{AD}(mG)$ (or $\mathcal{AD}^t(mG)$) instead of $\mathcal{AD}(\{mG\})$ (or $\mathcal{AD}^{(t)}(\{mG\})$).

The following example shows how the Adaptation Procedure of Example 5.2 continues in its second step. Interestingly, $mG^{(1)}$ includes a hybrid natural misinformed equilibrium, thus illustrating the branching process mentioned above.

Example 5.2 (continued). In $t = 1$, player 1 has equilibrium strategy s_2 in $G^{1,(1)}$, while player 2 has a mixed equilibrium strategy (randomizes between s_1 and s_2) in $G^{2,(1)}$. The corresponding nme has strategy profile $((0, 1), (1/3, 2/3))$. Using the characteristic strategy vector we take $\chi(nme) = \{(2, 1), (2, 2)\}$. Notice that, as one player randomized, $\chi(nme)$ has more than one elements, and the Adaptation Procedure branches, resulting to two new misinformation games, say $mG^{(2a)}, mG^{(2b)}$.

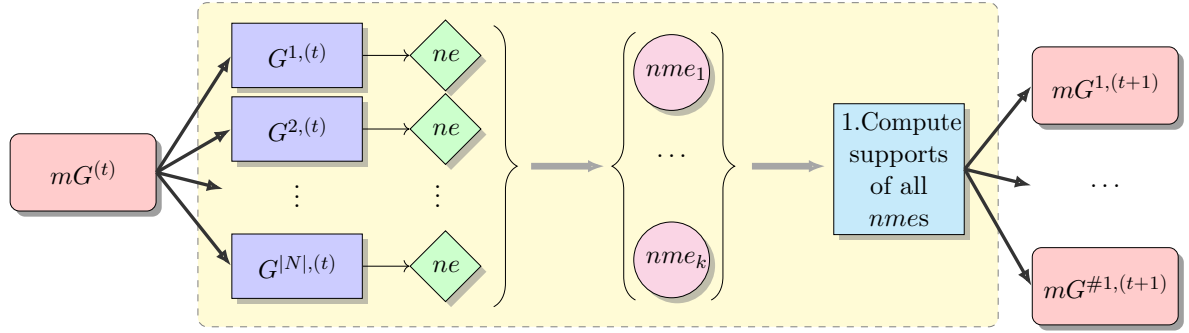


Figure 5.1: Schematic representation of the functionality of Adaptation Procedure from time step t to time step $t+1$, that is $\mathcal{AD}^{t+1}(mG)$.

Let us first consider the element $(2,1)$ of $\chi(nme)$ (which leads to $mG^{(2a)}$). We note that the payoff matrices of $mG^{(1)}$ are already updated with the correct value with respect to the bottom-left element, therefore $mG^{(2a)} = mG^{(1)}$.

Similarly, for the element $(2,2)$ of $\chi(nme)$, we update the bottom-right element of $P^{1,(1)}$ and $P^{2,(1)}$, so $mG^{(1)}$ leads to $mG^{(2b)} = \langle G^0, G^{1,(2b)}, G^{2,(2b)} \rangle$ with payoff matrices:

$$P^{1,(2b)} = \begin{pmatrix} (2,2) & (0,3) \\ (3,0) & (1,1) \end{pmatrix}, \quad P^{1,(2b)} = \begin{pmatrix} (1,1) & (3,0) \\ (3,0) & (1,1) \end{pmatrix}$$

From the above, we conclude that $\mathcal{AD}^{(2)}(\{mG^{(0)}\}) = \{mG^{(1)}, mG^{(2b)}\}$. ■

5.2.3 Stabilisation of the Adaptation Procedure

The following definition determines when the procedure is assumed to have “terminated”; this corresponds to the time point where any further iterations do not provide new information to the players:

Definition 32. We say that the Adaptation Procedure stops at step t , if and only if t is the smallest integer for which

$$\mathcal{AD}^{(t+1)}(M) = \mathcal{AD}^{(t)}(M)$$

for $t \in \mathbb{N}_0$. We call this t the length of the Adaptation Procedure and we denote it as $\mathfrak{L}_{\mathcal{AD}}(M)$.

In other words, the Adaptation Procedure ends at the first time step in which all of the spawned misinformation games already appear in the recursive tree (though not necessarily in the same branch). As we will show later (Proposition 16), all finite misinformation games terminate, so the notion of termination is well-defined (at least for finite games).

To simplify presentation in the following, we will denote by $\mathcal{AD}^*(M)$ the set of all misinformation games created by M through successive applications of the $\mathcal{AD}(\cdot)$ function, i.e., $\mathcal{AD}^*(M) = \bigcup_{t=0}^{\infty} \mathcal{AD}^{(t)}(M)$. Moreover, we will denote by $\mathcal{AD}^\infty(M)$ the misinformation games that $\mathcal{AD}(\cdot)$ produces after its termination point, i.e., $\mathcal{AD}^\infty(M) = \mathcal{AD}^{(t)}(M)$ for $t = \mathfrak{L}_{\mathcal{AD}}(M)$.

Definition 33. Consider a misinformation game mG . Then, σ is a stable misinformed equilibrium (or sme for short) of mG , iff there exists some $\widehat{mG} \in \mathcal{AD}^\infty(\{mG\})$ such that $\sigma \in NME(\widehat{mG})$ and, for all $\vec{v} \in \chi(\sigma)$, $\widehat{mG}_{\vec{v}} = \widehat{mG}$.

We denote by $SME(mG)$ the smes of mG .

Example 5.2 (continued). For $t = 2$, let us first consider $mG^{(2a)}$. As mentioned, $mG^{(2a)} = mG^{(1)}$ so, for reasons analysed in Example 5.2, it branches into $mG^{(3a)} = mG^{(1)}$, $mG^{(3b)} = mG^{(2b)}$.

As regards $mG^{(2b)}$, we note that the nme of $mG^{(2b)}$ is $\sigma = ((0, 1), (1/2, 1/2))$, for which $\chi(\sigma) = \{(2, 1), (2, 2)\}$. Observe that both positions in $\chi(\sigma)$ (namely, $(2, 1)$, $(2, 2)$) are known to the players, i.e., it holds that $P_{\vec{v}}^{i, (2b)} = P_{\vec{v}}^0$, for $i \in \{1, 2\}$, $\vec{v} \in \{(2, 1), (2, 2)\}$. Thus, $mG^{(3c)} = mG^{(2b)}$.

Combining the above, we observe that $\mathcal{AD}^{(3)}(\{mG^{(0)}\}) = \{mG^{(1)}, mG^{(2b)}\} = \mathcal{AD}^{(2)}(\{mG^{(0)}\})$, so the Adaptation Procedure terminates at step 2, i.e., $\mathfrak{L}_{\mathcal{AD}}(\{mG^{(0)}\}) = 2$.

Now let us identify the smes of $mG^{(0)}$. As explained above, and in Example 5.2 $NME(mG^{(1)}) = \{((0, 1), (1/3, 2/3))\}$, $NME(mG^{(2b)}) = \{((0, 1), (1/2, 1/2))\}$. As regards $\sigma_1 = ((0, 1), (1/3, 2/3))$, we note that it is not an sme, because there exists a position $(2, 2) \in \chi(\sigma_1)$, for which $mG_{(2,2)}^{(1)} = mG^{(2b)}$. On the other hand, $\sigma_2 = ((0, 1), (1/2, 1/2))$ is an sme, because, as mentioned above, $P_{\vec{v}}^{i, (2b)} = P_{\vec{v}}^0$, for $i \in \{1, 2\}$, $\vec{v} \in \{(2, 1), (2, 2)\}$, so $mG_{\vec{v}}^{(2b)} = mG^{(2b)}$ for $\vec{v} \in \{(2, 1), (2, 2)\}$. ■

5.3 Adaptation Procedure: Visualisations

5.3.1 Adaptation Procedure Graph (AP-Graph)

The Adaptation Procedure can be visualized in various ways. The most natural one, which is described in this subsection, visualises the Adaptation Procedure as a graph, where nodes are misinformation games, and arrows represent “transitions”, i.e., how each misinformation game was produced by another through the $\mathcal{AD}(\cdot)$ function. Figure 5.2 visualizes the AP-Graph of the Adaptation Procedure of our running example above.

More formally, the AP-Graph of a misinformation game mG is defined as follows:

Definition 34 (AP-Graph). Consider a misinformation game mG . The Adaptation Procedure Graph (AP-Graph for short) of mG is the graph $G(mG) = (V, E)$, where:

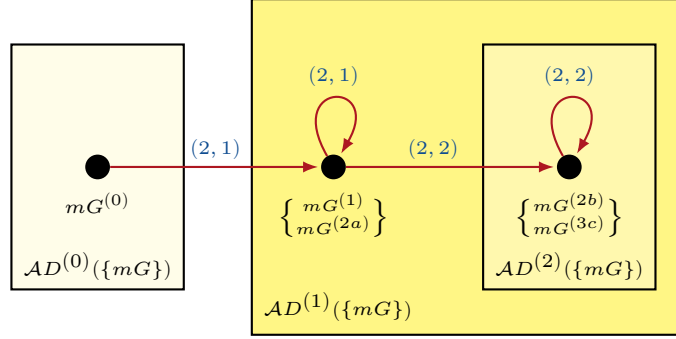


Figure 5.2: AP-Graph of Adaptation Procedure for misinformation game in Example 5.2, with $\{mG^{(1)}, mG^{2a}\} \in \mathcal{AD}^{(1)}(\{mG\})$ and $\{mG^{(2b)}, mG^{(3c)}\} \in \mathcal{AD}^{(2)}(\{mG\})$.

- $V = \mathcal{AD}^*(\{mG\})$
- $E = \{(mG_1, mG_2) \mid mG_1 \in V, mG_2 \in V \text{ and } mG_2 \in \mathcal{AD}(\{mG_1\})\}$

As is obvious by the definition, in the AP-Graph, the nodes are the misinformation games, and two nodes are connected with a directed link if and only if the target node was “produced”, through $\mathcal{AD}(\cdot)$, by the source node.

5.3.2 Adaptation Procedure Induced Graph (API-Graph)

The second representation formalism to be considered is again a graph, whose nodes are the different positions in the payoff matrix of the original game, whereas edges show how the players transition from one position to another during the Adaptation Procedure.

To describe the API-Graph formally, we will need a set enumerating all possible positions in the payoff matrices of a misinformation game. This set is denoted by $S_{\#}(mG)$. Formally, consider a misinformation game $mG = \langle G^0, G^1, \dots, G^{|N|} \rangle$ such that $G^0 = \langle N, S, P \rangle$, $S = S_1 \times \dots \times S_{|N|}$, and suppose that $|S_i| = n_i$. We denote by $S_{\#}(mG)$ the set $S_{\#}(mG) = [n_1] \times \dots \times [n_{|N|}]$. Observe that the elements of $S_{\#}(mG)$ are $|N|$ -tuples of numbers, indicating one position in the payoff matrix of G^i .

The API-Graph is formally defined as follows:

Definition 35 (API-Graph). *Consider a misinformation game mG . The Adaptation Procedure Induced Graph (or API-Graph for short) of mG is the graph $G_*(mG) = (V_*, E_*)$ such that:*

- $V_* = S_{\#}(mG)$
- $E_* = \{(\vec{u}, \vec{v}) \in V_* \times V_* \mid \text{there exists } mG^1 \in \mathcal{AD}^*(\{mG\}) \text{ and } \sigma_1 \in NME(mG^1), \text{ such that } \vec{u} \in \chi(\sigma_1) \text{ and } \vec{v} \in \chi(\sigma_2) \text{ for some } \sigma_2 \in NME(mG^1_{\vec{u}})\}$

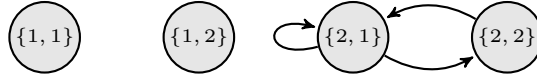


Figure 5.3: API graph in mG 's games for Example 5.2.

As is clear by the definition, to connect two positions \vec{u} , \vec{v} through an edge in the API-Graph, it should be the case that, when updating the position \vec{u} at some step of the Adaptation Procedure, the misinformation game produced has an *nme* whose relevant positions (through the χ function) include \vec{v} . In other words, it links together potential player behaviours: if \vec{u} becomes a possible player behaviour at some step, and \vec{u} points to \vec{v} in the API-Graph, then the next step will include \vec{v} as a possible player behaviour.

Figure 5.3 shows the API-Graph the produced by the Adaptation Procedure in Example 5.2.

5.4 Adaptation Procedure: Properties

5.4.1 General Properties

We will start by providing two useful properties of the Adaptation Procedure:

Proposition 13. *Consider the $AD(M)$, then*

$$AD(M) = \bigcup_{mG \in M} AD(\{mG\})$$

Proof. By definition:

$$\begin{aligned}
 AD(M) &= \{mG_{\vec{u}} \mid mG \in M, \vec{u} \in \chi(\sigma), \sigma \text{ is a } nme \text{ of } mG\} \\
 &= \bigcup_{mG \in M} \{mG_{\vec{u}} \mid \vec{u} \in \chi(\sigma), \sigma \text{ is a } nme \text{ of } mG\} \\
 &= \bigcup_{mG \in M} AD(\{mG\})
 \end{aligned}$$

□

Proposition 14. *Consider the a finite misinformation game mG such that $\{mG, mG'\} = AD(\{mG\})$, with $mG \neq mG'$. Then,*

$$AD(\{mG\}) \subseteq AD(AD(\{mG\}))$$

Proof. Observe that, by Proposition 13,

$$AD(AD(\{mG\})) = AD(\{mG, mG'\}) = AD(\{mG\}) \cup AD(\{mG'\}) \supseteq AD(\{mG\}).$$

□

5.4.2 Termination, and existence of smes

In this Subsection we show that the adaptation procedure will always terminate when mG is finite itself. Firstly, we focus our attention on a single branch of the Adaptation Procedure.

Let us fix some (arbitrary) selection function $\varphi : 2^{S_\#} \setminus \emptyset \mapsto S_\#$, such that $\varphi(X) \in X$ for all $X \in 2^{S_\#} \setminus \emptyset$. A selection function essentially picks a position out of a set of positions, and can be used to determine a branch in the Adaptation Procedure, as the following definition shows:

Definition 36 (*Branch*). Consider a misinformation game mG and some selection function φ . We define the branch determined by the selection function φ as the sequence $(H^t)_{t \geq 0}$ where:

$$\begin{aligned} H^0 &= \emptyset, \\ H^t(mG) &= H^{t-1}(mG) \cup \varphi \left(\{ \chi(\sigma) \mid \sigma \in NME(mG_{H^{t-1}(mG)}) \} \right), \text{ for } t > 0 \end{aligned}$$

Using the notion of the branch, we can show the following, which essentially shows that the notion of the branch, as defined in Definition 36 indeed corresponds to a branch in the adaptation process:

Proposition 15. Consider a finite misinformation game mG . Then, for any branch $(H^t)_{t \geq 0}$, and any $t \geq 0$, there exists some $mG' \in \mathcal{AD}^{(t)}(\{mG\})$ such that $mG' = mG_{H^t(mG)}$.

Proof. Take some branch $(H^t)_{t \geq 0}$. For $t = 0$, $H^0 = \emptyset$ and $mG_{H^0} = mG \in \mathcal{AD}^{(0)}(\{mG\})$, so the result holds.

Assume that the result holds for $t = k$. For $t = k + 1$, we take an $\widehat{mG} \in \mathcal{AD}^{(k+1)}(\{mG\})$. Thus, $\widehat{mG} = \widehat{mG}_b$, for some position $b \in H^{k+1}(mG)$ and $\widehat{mG} \in \mathcal{AD}^{(k)}(\{mG\})$. From the inductive step, $\widehat{mG} = mG_{H^k(mG)}$, thus $\widehat{mG} = (mG_{H^k(mG)})_b = mG_{H^{k+1}(mG)}$. □

Theorem 1. Consider a finite misinformation game mG . Then, for all $mG' \in \mathcal{AD}^{(t)}(\{mG\})$ it holds that, either $mG' \in \mathcal{AD}^{(t-1)}(\{mG\})$, or there exists a branch $(H^t(mG))_{t \geq 0}$ such that $mG' = mG_{H^t(mG)}$ and $|H(mG)^t| = t$.

Proof. For $t = 0$, obviously $mG \in \mathcal{AD}^{(0)}(\{mG\})$, so the result holds. Assume that the result holds for $t = k$. For $t = k + 1$, by definition, $mG' \in \mathcal{AD}(\{mG''\})$ for some $mG'' \in \mathcal{AD}^{(k)}(\{mG\})$. If $mG' = mG''$ the result holds trivially, so let us assume that $mG' \neq mG''$. Note that the inductive hypothesis holds for mG'' , so there are the following two cases with regards to mG'' :

1. If $mG'' \in \mathcal{AD}^{(k-1)}(\{mG\})$, then obviously $mG' \in \mathcal{AD}^{(k)}(\{mG\})$, so the result holds.

2. If there exists a branch $(\hat{H}^t(mG))_{t \geq 0}$ such that $mG'' = mG_{\hat{H}^k(mG)}$ and $|\hat{H}^k(mG)| = k$, then we note that $mG' = mG''_{\vec{v}}$ for some position \vec{v} , i.e., $mG' = (mG_{\hat{H}^k(mG)})_{\vec{v}}$. If $\vec{v} \in \hat{H}^k(mG)$, then $mG' = mG''$, a contradiction by our hypothesis. So $\hat{H}^k(mG) \cup \{\vec{v}\}$ contains $k + 1$ elements. Moreover, there exists some branch $(H^t(mG))_{t \geq 0}$ such that $H^{k+1}(mG) = \hat{H}^k(mG) \cup \{\vec{v}\}$. This concludes the proof. \square

We continue by showing that the Adaptation Procedure will always terminate when mG is finite itself:

Proposition 16. *Consider a finite misinformation game mG . Then $\mathcal{L}_{AD}(mG)$ is finite and $AD^\infty(\{mG\})$ is finite.*

Proof. Observe that as mG is finite, the different positions are finite, so H^t is finite. Thus, from Theorem 1, the Adaptation Procedure terminates. \square

Unfortunately, a similar result cannot be shown for infinite games. The following counterexample proves this fact:

Example 5.3. *Consider $G^0 = \langle N, S, P^0 \rangle$ such that $N = \{r, c\}$, $S_r = \{1, 2, \dots\}$, $S_c = \{1, 2\}$, $S = S_r \times S_c$, and the payoff for a position $(x, y) \in S_r \times S_c$ is computed as follows:*

$$P_{(x,y)}^0 = \left(\frac{1}{x}, y \right)$$

As a result of this definition, the only (pure) nash equilibrium for G^0 is in position $(1, 2)$, where the payoff is $(1, 2)$.

Now consider the misinformation game $mG = \langle G^0, G^r, G^c \rangle$, where $G^r = \langle N, S, P^r \rangle$ such that $P_{(x,y)}^r = (\frac{1}{x} + 1, y)$, and $G^c = G^0$. By the definition of mG we note that player c knows the correct payoffs and will always play strategy 2. On the other hand, player r knows the correct payoffs as far as player c is concerned, but his own payoffs are distorted, and he believes that their actual value is 1 point more than they really are. The key observation here is that, for player r , any of his subjective payoffs is better than any of the actual ones. Therefore, when he learns any position, this position becomes highly unattractive and cannot be selected again.

More formally, we note that $NME(mG) = \{\sigma_0\}$, where $\chi(\sigma_0) = \{(1, 2)\}$. Thus, $AD^{(1)}(\{mG\}) = \{mG^{(1)}\}$, where $mG^{(1)} = mG_{(1,2)}$. It is easy to see that $NME(mG^{(1)}) = \{\sigma_1\}$, where $\chi(\sigma_1) = \{2, 2\}$. Continuing this process, we observe that $AD^{(i)}(\{mG\}) = \{mG^{(i)}\}$, where $mG^{(i)}$ is such that the positions $(1, 2), (2, 2), \dots, (i, 2)$ have been learnt. But the only (pure) nme of $mG^{(i)}$ corresponds to the position $(i + 1, 2)$. As a result, the Adaptation Procedure will continuously lead to the learning of new positions (and, thus, to new misinformation games), which shows that the Adaptation Procedure will not terminate.

Due to this negative result, all subsequent analysis focuses on finite misinformation games.

We will now show that all finite misinformation games have an sme. To start with, we show the following result, which describes a condition sufficient for the existence of an sme:

Lemma 2. *If $mG' \in AD^*(\{mG\})$ and $AD(\{mG'\}) = \{mG'\}$ then $NME(mG') \subseteq SME(mG)$.*

Proof. Suppose that $mG' \in AD^{(t_0)}(\{mG\})$ for some $t_0 \geq 0$. Since $mG' \in AD(\{mG'\})$, it follows that $mG' \in AD^{(t)}(\{mG\})$ for all $t \geq t_0$, thus, $mG' \in AD^\infty(\{mG\})$. Since $AD(\{mG'\}) = \{mG'\}$, it is clear that for all $\sigma \in NME(mG')$ and for all $\vec{v} \in \chi(\sigma)$, it holds that $mG'_{\vec{v}} = mG'$. Now the result is direct from Definition 33. \square

Lemma 3. *Take some finite mG_1, \dots, mG_n such that $mG_{i+1} \in AD(\{mG_i\})$, for $i = 1, \dots, n-1$ and $mG_1 \in AD(\{mG_n\})$. Then, $mG_i = mG_j$ for all i, j .*

Proof. Observe that $mG_i \in AD(AD(\dots AD(\{mG_i\}))) = AD^{(n)}(\{mG_i\}) \forall i \in [n]$. Suppose, for the sake of contradiction, that $mG_i \neq mG_j$ for some i, j , and assume, without loss of generality, that $i < j$. Then, it holds that $mG_j \in AD^{(j-i)}(\{mG_i\})$, i.e., mG_j has resulted from mG_i by updating some (at least 1 and at most $j-i$) elements of the respective payoff matrices of mG_i . But then, we also have that $mG_i \in AD^{(n-j+i)}(\{mG_j\})$ (by the periodic pattern above), so again, mG_i has resulted from mG_j by updating some (at least 1 and at most $n-j+i$) elements of the respective payoff matrices of mG_j . But this is an absurdity, because replacements are cumulative and cannot be “undone” by subsequent ones (see Definitions 29 and 31). \square

Proposition 17. *If mG is finite, then $SME(mG) \neq \emptyset$.*

Proof. Set $S = AD^\infty(\{mG\})$. For any given $mG_1, mG_2 \in S$, we define the relation \mapsto , such that $mG_1 \mapsto mG_2$ if and only if $mG_1 \neq mG_2$ and $mG_2 \in AD(\{mG_1\})$. Now let us suppose, for the sake of contradiction, that mG has no sme. By Lemma 2, it follows that for any $mG' \in S$ there exists some $mG'' \in S$ such that $mG' \mapsto mG''$ (otherwise $SME(mG) \neq \emptyset$ by Lemma 2, which contradicts our hypothesis). Since S is finite (see Proposition 16), there must exist a sequence of $mG_1, \dots, mG_n \in S$, such that $mG_i \mapsto mG_{i+1}$ (for $i = 1, \dots, n-1$) and $mG_n \mapsto mG_1$. Which is an absurdity by the definition of \mapsto and Lemma 4. \square

5.4.3 Estimating the length of the Adaptation Procedure

The only input to the Adaptation Procedure is the original misinformation game, say mG . Therefore, the entire Adaptation Procedure, as well as the determination of the related quantities (e.g., $SME(mG)$, $\mathcal{L}_{AD}(mG)$ etc.) are essentially determined by mG . However, in order to compute these, we need to perform the entire recursive process.

An interesting research question is what can be said about the Adaptation Procedure by just looking at mG , i.e., without necessarily “running” the Adaptation Procedure. In

this subsection, we derive upper bounds for $\mathfrak{L}_{AD}(mG)$ in an “a priori” manner, i.e., by just looking at mG , and without computing the nmes of mG .

To do this, we start by defining the *neighborhood* of a position \vec{v} :

Definition 37. Consider a pair of positions \vec{v}, \vec{u} . We say that \vec{v}, \vec{u} are neighbours if and only if they differ exactly in one element. We define the neighbourhood of \vec{v} as the set $\mathfrak{N}(\vec{v}) = \{\vec{u} \mid \vec{v}, \vec{u} \text{ are neighbours}\}$. Further, we define the k -neighbor of \vec{v} to be the set of vectors that differ from \vec{v} in the k^{th} position only, denoted by $\mathfrak{N}_k(\sigma)$.

Example 5.2 (continued). Consider the game $G^{1,(0)}$ and the strategy profile $\sigma = ((0, 1), (1, 0))$, then the neighborhood of σ is $\mathfrak{N}(\sigma) = \{((1, 0), (1, 0)), ((0, 1), (0, 1))\}$. Also, $\mathfrak{N}_1(\sigma) = \{((0, 1), (0, 1))\}$ and $\mathfrak{N}_2(\sigma) = \{((1, 0), (0, 1))\}$. ■

Our estimate is based on the following idea: consider a pair of neighbouring positions \vec{v}, \vec{u} in a misinformation game mG , and some player x . If the payoff for x that corresponds to \vec{v} is better than the one corresponding to \vec{u} , then the pure strategy profile corresponding to \vec{u} is not an nme. If, additionally, the payoff corresponding to \vec{v} remains better, even after the players learn the actual payoffs associated with \vec{v} , then \vec{u} can never be part of an nme, or learnt by the players. It is thus called an *impossible* position, and will never be visited during the Adaptation Procedure. Formally:

Definition 38. Consider a finite misinformation game $mG = \langle G^0, G^1, \dots, G^N \rangle$, such that the payoff matrix of G^i is P^i . Consider also a position \vec{v} . \vec{v} is called an impossible position for mG iff there exists some $i \in [N]$ and some $\vec{u} \in \mathfrak{N}_i(\vec{v})$ such that:

$$P_{\vec{v}}^i < P_{\vec{u}}^i \text{ and } P_{\vec{v}}^i < P_{\vec{u}}^0$$

A position will be called possible iff it is not impossible.

Using the concept of impossible positions, we can set an upper bound for $\mathfrak{L}_{AD}(mG)$ using the simple idea that, in the worst case scenario, each branch of the Adaptation Procedure will visit, at most, the possible positions.

Proposition 18. Consider the misinformation game mG , some $t > 0$, and some $mG^{(t)}$ such that $mG^{(t)} \in \mathcal{AD}^{(t)}(\{mG\})$. Then, there exists some $mG^{(t-1)} \in \mathcal{AD}^{(t-1)}(\{mG\})$ and some possible position \vec{v} such that $mG_{\vec{v}}^{(t-1)} = mG^{(t)}$.

Proof. By the definition of the Adaptation Procedure, indeed there exists some $mG^{(t-1)} \in \mathcal{AD}^{(t-1)}(\{mG\})$ and some position \vec{v} such that $mG_{\vec{v}}^{(t-1)} = mG^{(t)}$. It remains to show that \vec{v} is possible.

Suppose, for the sake of contradiction, that \vec{v} is impossible. By the construction of \vec{v} , it holds that $\vec{v} \in \chi(\sigma)$ for some $\sigma \in NME(mG^{(t-1)})$. Thus, the respective position will be the best response against σ_{-i} , for all players i , which means that it is preferred over

other positions. However, since \vec{v} is impossible, there exists some $\vec{u} \in NBGi\vec{v}$ such that $P_{\vec{v}}^{i,(t-1)} < P_{\vec{u}}^{i,(t-1)}$ and $P_{\vec{v}}^{i,(t-1)} < P_{\vec{u}}^{0,(t-1)}$. Thus, if the payoffs related to \vec{u} have not been updated in any of the previous steps in the Adaptation Procedure, then \vec{v} cannot be preferred, so σ cannot be an nme of $mG^{(t-1)}$ due to the first relation above; similarly, if the payoffs related to \vec{u} have been updated in some previous step in the Adaptation Procedure, then \vec{v} cannot be preferred, so σ cannot be an nme of $mG^{(t-1)}$ due to the second relation above. Consequently, we reach a contradiction, which shows the result. \square

Proposition 19. *For any misinformation game mG , $\mathfrak{L}_{AD}(mG) \leq K$, where K is the number of possible positions in mG .*

Proof. Given the fact that the same position cannot be “learnt twice”, and Proposition 18, the result is direct. \square

Note that this is another way to prove Proposition 16, as in the worst-case scenario, the number of possible positions are all the possible positions, i.e., $|S_{\#}(mG)|$, so $\mathfrak{L}_{AD}(mG) \leq |S_{\#}(mG)|$.

An interesting observation that is derived from the above proposition is that there is a limit on the portion (percentage) of the actual specifications that players learn. It is not hard to see that this is bounded by the ratio between the number of possible positions and the number of total positions, that is $|S_{\#}(mG)|$, of mG . Thus, Adaptation Procedure could terminate without the players knowing the whole G^0 when impossible positions exist.

5.5 The Epistemic Adaptive Evolution

Until now we have developed a procedure where each player mechanistically absorbs any new information that is publicly announced. Nevertheless, as have been already mentioned, in a multi-turn process players reconsider not only the values of the payoffs they have, but a whole mindset. In this section we develop such a framework.

5.5.1 Informal Description

The input to the epistemic game-playing process is a finite misinformation game $mG^{(0)}$, where each player i receives his/her own game, $G^{i,(0)}$, and initially assumes that this is the game being played. Therefore, he/she will play one of his/her equilibrium strategies. The chosen strategies for each player are then publicly announced, along with the payoffs associated with the emerging strategy profile.

At that point, each player realizes that (some of the) other players play in an unexpected manner. Given the assumption that players are rational, this can only be explained by assuming that there is an error in the payoff matrix of said player. Therefore, the player takes two actions:

- First, he/she updates the payoff matrix in such a way that the other players' choices make sense (i.e., they are equilibrium strategies of the new matrix). Any changes made should be minimal, i.e., the resulting matrix should be one with the above property, but also being “as close as possible” to the one the player originally believed. This is called *epistemic evolution*.
- Second, he/she replaces (in the respective positions in the payoff matrix) the payoffs that were communicated by the environment (i.e., the real payoffs associated with the strategy profile that was played). This is called *adaptive update*. Note that adaptive update is the same as adaptation procedure in Section 5.2.

There are various subtleties in the above analysis, some of them are analogous with that of Section 5.2 while others are new. Namely, the procedure is iterative, i.e., it continues until it “stabilizes” somewhere. Stabilization will occur at a state where all players receive input (strategic choices of opponents and payoffs) that are consistent with what they believe.

Second, the epistemic evolution may lead to a set of possible payoff matrices that are “equally close” to the original one(s). Therefore, the entire process should be general enough to allow for updating sets of payoff matrices.

Third, at each stage, each of the players may have to choose among more than one possible equilibrium strategies. In this case, we assume that all such choices are explored in independent “branches” of the procedure. In other words, the process is not, in general linear.

As in Adaptation Procedure, the Epistemic Adaptive Evolution produces new games in each time step. Here, the concept of natural misinformed equilibrium changes slightly in order to capture the consistency of a player's decision with his/her epistemic view. We call such an equilibrium *epistemic natural equilibrium*. Similarly with Section 5.2 an epistemic natural equilibrium in the time step where Epistemic Adaptive Evolution stabilizes is called *stable epistemic natural equilibrium*.

In conclusion, our approach makes the following assumptions:

1. The input is a canonical misinformation game with N players, each of them having S available strategies.
2. Each player is informed about his/her own game, as dictated by the misinformation game, but he/she has no knowledge of the actual game or the other players' games.
3. In each round, all players assume that all other players have the same game as themselves.
4. In each round, all players assume that the game they have is the correct one.
5. In each round, the players will communicate all Nash equilibrium strategies (of their own game). Each of these strategies is assumed to be communicated independently, i.e., in different branches of the process.

6. In each round, when the players receive unexpected equilibrium strategies from their opponents, they will assume that all players are “right”, so they will update their games in such a way that the equilibrium strategies received make sense. Afterwards, they will receive their payoffs (from the environment, and further update their payoff matrix accordingly.
7. If the above updates can be made in multiple, equally plausible ways (resulting in different, but equally plausible alternative payoff matrices), then the players will make no arbitrary choice; instead, they will keep all such matrices, awaiting future input to disambiguate the various options.

Before present formally our methodology we provide a real-life scenario in order to highlight and explain our setting. Specifically, consider the case where two or more financial groups invest over the same commodities and goods. Each company has his/her own view about the interaction that, possible, differs from the actual specification, and make his/her decisions according to his/her view. Further, in order to decide, each company, assumes: i) that any other has the same view as he/she has, and ii) that this view is the correct one. Though, we can allow the companies to incorporate a Bayesian opponent we restrict our analysis the simpler form of our model. Next, in each round every company plays a Nash equilibrium strategy according to his/her view; the agglomeration of all individual decisions, is the joint decision and is publicly announced with the relevant actual values at the end of each turn.

For the last assumption, consider the case where a company receives an unanticipated decision from a rival company. One way to mitigate this issue is to allow the company to assume that the rival is irrational; this is the case where the company relies only in his/her knowledge, so is highly unsophisticated and non-realistic. A second way is to wipe out any knowledge he/she has and rely on the received information. This produces computational and conceptual issues (e.g. how he/she would produce the necessary information from a single decision?) in our analysis. The only plausible way to cope with an unanticipated behavior is to combine the two previous ways, namely to allow the company to reconsider what he/she knows in order to fit it with the joint decision.

In the rest of this section, we assume fixed mG , as above.

Example 5.3 (Running example). *Consider the canonical misinformation game $mG^{(0)} = \langle G^0, G^{1,(0)}, G^{2,(0)} \rangle$ with $S = S_1 \times S_2$, $|S_1| = |S_2| = 2$ and payoff matrices,*

$$P^0 = \begin{pmatrix} (2, 2) & (0, 3) \\ (3, 0) & (1, 1) \end{pmatrix}, P^{1,(0)} = \begin{pmatrix} (1, 1) & (3, 0) \\ (0, 3) & (2, 2) \end{pmatrix}, P^{2,(0)} = \begin{pmatrix} (2, 1) & (0, 0) \\ (0, 0) & (1, 2) \end{pmatrix}$$

with $nme = \{((1, 0), (1, 0)), ((1, 0), (0, 1)), ((1, 0), (1/3, 2/3))\}$. Notationally, player 1 takes the payoff values by $P_1^{x,(0)}$ and player 2 takes the values $P_2^{x,(0)}$, where $x = \{0, 1, 2\}$. Fur-

ther, the matrix $P_1^{x,(0)}$ has elements $(P_1^{x,(0)})_{ij} = a_{ij}$ and $(P_2^{x,(0)})_{ij} = b_{ij}$. Further, $Pos = \{(1, 1), (1, 2), (2, 1), (2, 2)\}$. ■

5.5.2 Preliminaries

We consider a normal-form game $G = \langle N, S, P \rangle$, with $|S|$ pure strategies per player, hence P is an $|N|$ -dimensional matrix, where each dimension has m elements and each element is a $|N|$ -tuple. We denote by $\overline{\mathfrak{P}}$ all the matrices with this property, i.e.:

$$\overline{\mathfrak{P}} = 2 \left((\mathbb{R}^{|N|})^{|S|} \right)^{|N|}$$

As we focus on the payoff matrices we change slightly the notation of Definition 3.1. Namely, given a position $\vec{v} = (v_1, \dots, v_{|N|}) \in Pos$, the payoff of the player i (according to the matrix P) in the position determined by the strategies $(s_{v_1}, \dots, s_{v_n})$ is $h_i(\vec{v}) = P_i(\vec{v})$.

A *value assignment function* VA is a partial function $VA : Pos \mapsto \mathbb{R}^{|N|}$. Intuitively, VA is a partial specification for a payoff matrix, so $VA(\vec{v})$ returns a vector representing the payoffs of all players for the pure strategy profile represented by \vec{v} , under this partial matrix specification. I.e., if $\vec{v} = (v_1, \dots, v_{|N|})$, then $VA(\vec{v})$ returns the payoffs of the partial matrix specification when the players play $(s_{v_1}, \dots, s_{v_{|N|}})$. If $VA(\vec{v}) = (\rho_1, \dots, \rho_{|N|})$, we say that the *value of* VA for player $i \in N$ is ρ_i . We denote by $VA_i(\vec{v})$ the value of VA for i . We denote by $Dom(VA)$ the domain of VA , i.e., the subset of Pos for which VA is defined. Note that we allow that $Dom(VA) = \emptyset$ in which case we say that we have an *empty assignment function*. The empty assignment function will be denoted by VA_\emptyset .

For two value assignment functions VA_1, VA_2 , such that $Dom(VA_1) \cap Dom(VA_2) = \emptyset$, we denote by $VA_1 \uplus VA_2$ the value assignment function VA such that:

- $Dom(VA) = Dom(VA_1) \cup Dom(VA_2)$
- $VA(\vec{v}) = VA_1(\vec{v})$ for all $\vec{v} \in Dom(VA_1)$
- $VA(\vec{v}) = VA_2(\vec{v})$ for all $\vec{v} \in Dom(VA_2)$

Given a value assignment function VA and a payoff matrix P , we say that P satisfies VA (denoted by $P \vdash VA$) if and only if $P_i(\vec{v}) = VA_i(\vec{v})$ for all $i \in N$, $\vec{v} \in Dom(VA)$. Abusing notation, for a set of payoff matrices \mathfrak{P} , we similarly write $\mathfrak{P} \vdash VA$ (and say that \mathfrak{P} satisfies VA) if and only if $P \vdash VA$ for all $P \in \mathfrak{P}$.

Finally, we will use the symbol $NE(P)$ to denote the set of all Nash equilibria of the payoff matrix P .

5.5.3 Players' Epistemic game

In each step of the epistemic game-playing procedure, each player has a belief about the state of the world (i.e., the game's payoff matrix); however, as already mentioned, such a belief does not necessarily identify some payoff matrix in a unique manner. To capture this intuition, we assume that each player believes that the actual interaction is governed by one of the payoff matrices found in a set, all of them equally plausible. This leads to a generalised definition of a misinformation game, where each player's subjective view is a set of different payoff matrices, rather than a unique one. We call these games *epistemic misinformation games*. Additionally, we incorporate in the epistemic misinformation games a value assignment function, representing the payoffs that the players already know (i.e., have already been communicated by the environment). Formally:

Definition 39. An epistemic misinformation game (or simply epistemic game) \mathcal{E} is a tuple

$$\mathcal{E} = \langle N, S, VA, P^0, \mathfrak{P}_1, \dots, \mathfrak{P}_{|N|} \rangle$$

where:

- $N = \{1, \dots, n\}$ is the set of players.
- $S = S_1 \times \dots \times S_{|N|}$, is the set of strategies, where, for any $i \in N$ it holds that $S_i = \{s_1, \dots, s_{|S|}\}$.
- $VA : Pos \mapsto \mathbb{R}^{|N|}$ is a value assignment function.
- $P^0 \in \overline{\mathfrak{P}}$ is a payoff matrix that satisfies VA , i.e., $P^0 \vdash VA$.
- For all $i \in N$, $\emptyset \subset \mathfrak{P}_i \subseteq \overline{\mathfrak{P}}$ is a non-empty set of payoff matrices that satisfies VA , i.e., $\mathfrak{P}_i \vdash VA$.

Intuitively, \mathfrak{P}_i represents the beliefs of the player q regarding the rules of the interaction. As we deal with canonical games, the set of players and strategies are constant, so, essentially, \mathfrak{P}_i is a shorthand for the set of games $\{\langle N, S, P \rangle \mid P \in \mathfrak{P}_i\}$, all of which are considered as equally plausible to be the actual game, from the perspective of i . The set $\{\langle N, S, P \rangle \mid P \in \mathfrak{P}_i\}$ is called the *subjective epistemic game* of player i , and we denote it by \mathcal{E}_i . Further, in case we refer to specific time step t of a multi-turn interaction we'll denote an epistemic game as $\mathcal{E}^{(t)}$, and the subjective epistemic game player i as $\mathcal{E}_i^{(t)}$.

Example 5.3 (continued). The epistemic misinformation game in time step $t = 0$ is $\mathcal{E}^{(0)} = \langle N, S, VA, P^0, P^{1,(0)}, P^{2,(0)} \rangle$, where $N, S, P^0, P^{1,(0)}$ and $P^{2,(0)}$ provided by the misinformation game $mG^{(0)} = \langle G^0, G^{1,(0)}, G^{2,(0)} \rangle$ and $VA : Pos(P^0) \rightarrow \mathbb{R}^{|N|}$. ■

Next, the equilibrium concept that emerges from an epistemic misinformation game defined as follows:

Definition 40. Consider an epistemic misinformation game $\mathcal{E} = \langle N, S, VA, P^0, \mathfrak{P}_1, \dots, \mathfrak{P}_{|N|} \rangle$. The strategy profile $\sigma = (\sigma_1, \dots, \sigma_{|N|})$ is called an epistemic natural equilibrium (ene) of \mathcal{E} if and only if for all $i \in N$, there exists $P_i \in \mathfrak{P}_i$ and $\sigma' = (\sigma'_1, \dots, \sigma'_{|N|}) \in NE(P_i)$, such that $\sigma_q = \sigma'_q$.

The set of enes of an epistemic misinformation game \mathcal{E} is denoted by $ENE(\mathcal{E})$.

Example 5.3 (continued). In $t = 0$ the interaction we have three enes, that is

$$ENE = \{((1, 0), (1, 0)), ((1, 0), (0, 1)), ((1, 0), (1/3, 2/3))\}.$$

■

5.5.4 Epistemic Update

We observe that a player must epistemically adapt his/her payoff matrix in two distinct ways. First, when a player learns about the behaviour (chosen strategies) of other players, he/she has to adapt his/her beliefs regarding the other players' payoffs, in order for these payoffs to be consistent with the chosen strategies. Secondly, he/she has to update the values of her payoff matrix according to what was communicated to his/her by the environment.

The former update is due to the fact that each player considers that his/her view is also the view of the rest of the players. The latter update is due to the fact that each player trusts its sensory input as regards the payoffs received when a given strategy profile was played.

The two aforementioned operations are applied sequentially, and have a different nature. The first is called *epistemic evolution*, whereas the second is called *adaptive evolution*. We define them in the following subsections.

Epistemic evolution

Given a game $G = \langle N, S, P \rangle$, a player $i \in N$, some strategy $\sigma_i \in \Sigma_i$, and a (possibly empty) value assignment function VA , we denote by:

$$\mathfrak{P}(i, \sigma_i, VA) = \{P \in \overline{\mathfrak{P}} \mid P \vdash VA \text{ and } \exists \sigma_{-i} \in \Sigma_{-i} \text{ such that } (\sigma_i, \sigma_{-i}) \in NE(P)\}$$

In words, $\mathfrak{P}(i, \sigma_i, VA)$ contains all the payoff matrices P , for which: (a) σ_i is a Nash equilibrium strategy of i in P ; and (b) P has the values dictated by VA in all positions where VA is defined.

In the context of epistemic evolution, $\bigcap_{i \in N} \mathfrak{P}(i, \sigma_i, VA)$ contains exactly the matrices that the other players may have, given that σ_i has been played as an equilibrium strategy of i , and that the values determined by VA are common knowledge for all players.

Lemma 4. *For every VA , there exists a strategy profile $\sigma = (\sigma_i, \sigma_{-i})$ such that $\mathfrak{P}(i, \sigma_i, VA) \neq \emptyset$.*

Proof. Given a value assignment function, there exists at least one payoff matrix P for which $P \vdash VA$. Moreover, there exists at least one $\sigma = (\sigma_i, \sigma_{-i}) \in NE(P)$. It follows that $P \in \mathfrak{P}(i, \sigma_i, VA)$, so $\mathfrak{P}(i, \sigma_i, VA) \neq \emptyset$. \square

Next, we define the Hamming distance* between two payoff matrices $P, Q \in \overline{\mathfrak{P}}$ as follows:

$$dist(P, Q) = |\{(i, \vec{v}) \mid P_i(\vec{v}) \neq Q_i(\vec{v})\}|$$

The following can be easily shown:

Lemma 5. *The Hamming distance $dist$ between two matrices is a metric.*

Proof. We observe the following:

- $dist(P, P) = 0, \forall P \in \overline{\mathfrak{P}}$.
- $dist(P, Q) = dist(Q, P), \forall P, Q \in \overline{\mathfrak{P}}$.
- It holds that:

$$\begin{aligned} dist(P, Q) &= |\{(i, \vec{v}) \mid P_i(\vec{v}) \neq Q_i(\vec{v})\}| \\ &= |\{(i, \vec{v}) \mid P_i(\vec{v}) \neq R_i(\vec{v})\} \setminus \{(i, \vec{v}) \mid R_i(\vec{v}) \neq Q_i(\vec{v})\}| \\ &\quad \cup |\{(i, \vec{v}) \mid R_i(\vec{v}) \neq Q_i(\vec{v})\} \setminus \{(i, \vec{v}) \mid P_i(\vec{v}) \neq R_i(\vec{v})\}| \\ &\leq |\{(i, \vec{v}) \mid P_i(\vec{v}) \neq R_i(\vec{v})\} \cup \{(i, \vec{v}) \mid R_i(\vec{v}) \neq Q_i(\vec{v})\}| \\ &\leq |\{(i, \vec{v}) \mid P_i(\vec{v}) \neq R_i(\vec{v})\}| + |\{(i, \vec{v}) \mid R_i(\vec{v}) \neq Q_i(\vec{v})\}| \\ &= dist(P, R) + dist(R, Q), \quad \forall P, Q, R \in \overline{\mathfrak{P}}. \end{aligned}$$

\square

The epistemic revision of a set of matrices (say \mathfrak{P}_1) with another one (say \mathfrak{P}_2) can now be defined as the set of matrices which belong in \mathfrak{P}_2 and are “as close as possible” (based on the metric $dist$, defined above) to some matrix in \mathfrak{P}_1 . This follows the intuition of classical works in belief revision, where the revision of a Knowledge Base K with a formula ϕ should be a Knowledge Base K' that implies ϕ (i.e., the models of K' should be a subset of the models of ϕ), and should be as close as possible to K (i.e., the models of K' should be as close as possible to the models of K) under some given distance metric. The above intuition is expressed formally as follows:

*See Definition 56 in Appendix B.3.

Definition 41. Given two sets of payoff matrices $\mathfrak{P}_1, \mathfrak{P}_2 \subseteq \overline{\mathfrak{P}}$, the epistemic revision of \mathfrak{P}_1 with \mathfrak{P}_2 is defined as follows:

$$\mathfrak{P}_1 \star \mathfrak{P}_2 = \{P \in \mathfrak{P}_2 \mid \exists P_1 \in \mathfrak{P}_1 \text{ such that } \forall P'_1 \in \mathfrak{P}_1, P'_2 \in \mathfrak{P}_2 : \text{dist}(P, P_1) \leq \text{dist}(P'_1, P'_2)\}$$

As explained above, the set of payoff matrices produced from the epistemic revision of \mathfrak{P}_1 with \mathfrak{P}_2 contain the payoff matrices of \mathfrak{P}_2 for which there is a minimal-distance corresponding matrix in \mathfrak{P}_1 . The following can be shown for the \star operator:

Lemma 6. For any $\mathfrak{P}_1, \mathfrak{P}_2 \subseteq \overline{\mathfrak{P}}$, $\mathfrak{P}_1 \star \mathfrak{P}_2 = \emptyset$ if and only if $\mathfrak{P}_2 = \emptyset$.

Proof. If $\mathfrak{P}_1 = \emptyset$, then $\mathfrak{P}_1 \star \mathfrak{P}_2 = \mathfrak{P}_2$ (by definition) so the result is obvious. So let us assume that $\mathfrak{P}_1 \neq \emptyset$.

Assume initially that $\mathfrak{P}_1 \star \mathfrak{P}_2 = \emptyset$, and suppose, for the sake of contradiction, that $\mathfrak{P}_2 \neq \emptyset$. The distance between any two matrices $P_1 \in \mathfrak{P}_1, P_2 \in \mathfrak{P}_2$ is finite, given that the matrices are finite. Therefore, we can always find at least one pair $P_1 \in \mathfrak{P}_1, P_2 \in \mathfrak{P}_2$ such that $\text{dist}(P_1, P_2)$ is minimal among all other such pairs. Therefore, $P_2 \in \mathfrak{P}_1 \star \mathfrak{P}_2$, a contradiction.

For the opposite, note that, if $\mathfrak{P}_2 = \emptyset$, then $\mathfrak{P}_1 \star \mathfrak{P}_2 = \emptyset$ by definition. \square

Lemma 7. For any $\mathfrak{P}_1, \mathfrak{P}_2 \subseteq \overline{\mathfrak{P}}$, the following are equivalent:

1. $\mathfrak{P}_1 \subseteq \mathfrak{P}_2$
2. $\mathfrak{P}_1 \star \mathfrak{P}_2 = \mathfrak{P}_1$

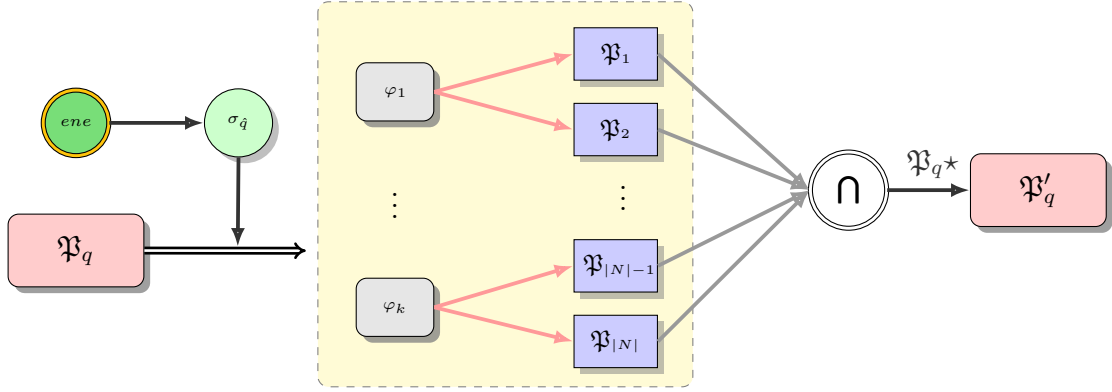
Proof. (1) \Rightarrow (2): Take any $P \in \mathfrak{P}_1$. Then $P \in \mathfrak{P}_2$, so for $P \in \mathfrak{P}_2$ there exists some $P_1 = P \in \mathfrak{P}_1$ for which $\text{dist}(P, P_1) = 0$, thus, minimal. Therefore, $\mathfrak{P}_1 \subseteq \mathfrak{P}_1 \star \mathfrak{P}_2$. Now consider some $P \notin \mathfrak{P}_1$. Then for any $P_1 \in \mathfrak{P}_1$, it holds that $\text{dist}(P, P_1) > 0$, so $P \notin \mathfrak{P}_1 \star \mathfrak{P}_2$. We conclude that $\mathfrak{P}_1 \star \mathfrak{P}_2 = \mathfrak{P}_1$.

(2) \Rightarrow (1): By Definition 41, $\mathfrak{P}_1 \star \mathfrak{P}_2 \subseteq \mathfrak{P}_2$, so the result is obvious. \square

Definition 41 is a prerequisite for the following, which shows how the epistemic misinformation game is epistemically revised based on a strategy profile and a value assignment function. This corresponds to the first type of evolution above, namely the change of a player's view of the world, based on what the others players play:

Definition 42. Consider an epistemic misinformation game $\mathcal{E} = \langle N, S, VA, P^0, \mathfrak{P}_1, \dots, \mathfrak{P}_{|N|} \rangle$, and some strategy profile of P^0 , $\sigma = (\sigma_1, \dots, \sigma_{|N|})$. The epistemic evolution of \mathcal{E} by σ is the epistemic misinformation game $EE(\mathcal{E}, \sigma) = \langle N, S, VA, P^0, \mathfrak{P}'_1, \dots, \mathfrak{P}'_{|N|} \rangle$, where, for all $i \in N$,

$$\mathfrak{P}'_q = \mathfrak{P}_q \star \left(\bigcap_{\bar{i} \in N} \mathfrak{P}(\bar{i}, \sigma_{\bar{i}}, VA) \right).$$

Figure 5.4: Player q revises his/her beliefs.

Obviously, it holds that $\mathfrak{P}'_i \vdash VA$.

The following example shows that the set $\bigcap_{i \in N} \mathfrak{P}(i, \sigma_q, VA)$ may be empty; thus, according to Lemma 6, it could be the case that $\left(\bigcap_{i \in N} \mathfrak{P}(\bar{i}, \sigma_{\bar{i}}, VA)\right) = \emptyset$.

Example 5.4. We will consider a scenario where $|N| = |S| = 2$, and tables:

$$P_1 = \begin{pmatrix} (2,2) & (3,3) \\ (1,1) & (0,0) \end{pmatrix}, \quad P_2 = \begin{pmatrix} (2,2) & (3,3) \\ (1,1) & (4,0) \end{pmatrix}$$

We observe that $NE(P_1) = \{((1,0), (0,1))\}$ and $NE(P_2) = \{((0.5,0.5), (0.5,0.5))\}$.

Now we construct the epistemic game $\mathcal{E} = \langle N, S, VA, P^0, \mathfrak{P}_1, \mathfrak{P}_2 \rangle$ where:

- $N = \{1, 2\}$
- $S = \{s_1, s_2\} \times \{s_1, s_2\}$
- VA is a value assignment function, such that $Dom(VA) = \{(1,1), (1,2), (2,1)\}$ and: $VA(1,1) = (2,2)$, $VA(1,2) = (3,3)$, $VA(2,1) = (1,1)$.
- $P^0 = P_1$
- $\mathfrak{P}_1 = \{P_1\}$
- $\mathfrak{P}_2 = \{P_2\}$

We observe that $P_1 \vdash VA$ and $P_2 \vdash VA$, so \mathcal{E} is indeed an epistemic game under Definition 39.

By the values of $NE(P_1)$, $NE(P_2)$ above, we observe that $ENE(\mathcal{E}) = \{((1,0), (0.5,0.5))\}$. Set $\sigma_1 = (1,0)$, $\sigma_2 = (0.5,0.5)$, $\sigma = (\sigma_1, \sigma_2) \in ENE(\mathcal{E})$, and $\mathfrak{P}_* = \mathfrak{P}(\bar{i}, \sigma_{\bar{i}}, VA)$. We will show that $\mathfrak{P}_* = \emptyset$.

Indeed, suppose that $P \in \mathfrak{P}_*$. It follows that $P \vdash VA$, so P is of the form:

$$P = \begin{pmatrix} (2,2) & (3,3) \\ (1,1) & (a,b) \end{pmatrix}$$

for appropriate $a, b \in \mathbb{R}$. Moreover, there must exist σ'_1, σ'_2 such that $(\sigma'_1, \sigma_2) \in NE(P)$ and $(\sigma_1, \sigma'_2) \in NE(P)$.

Equivalently, there should exist σ'_1, σ'_2 such that:

- $\sigma'_1 \in BR(1, P, \sigma_2)$
- $\sigma_2 \in BR(2, P, \sigma'_1)$
- $\sigma_1 \in BR(1, P, \sigma'_2)$
- $\sigma'_2 \in BR(2, P, \sigma_1)$

We show below that there is no appropriate choice for the value of the parameter a that satisfies the above condition. Indeed:

- If $a \leq 3$ then $BR(1, P, \sigma_2) = \{(1,0)\}$, so $\sigma'_1 = (1,0)$. However, $BR(2, P, \sigma'_1) = \{(0,1)\}$, thus $\sigma_2 \notin BR(2, P, \sigma'_1)$, a contradiction.
- If $a > 3$ then $BR(2, P, \sigma_1) = \{(0,1)\}$, so $\sigma'_2 = (0,1)$. However, $BR(1, P, \sigma'_2) = \{(0,1)\}$, thus $\sigma_1 \notin BR(1, P, \sigma'_2)$, a contradiction.

□

Example 5.3 (continued). Let's analyze the case where the $ENE = ((1,0), (1,0))$ is been publicly announced. This strategy profile is plausible for both players thus they take no further epistemic action.

In case that $ENE = ((1,0), (0,1))$ is been publicly announced. Then \mathfrak{P}'_1 is the set of payoff matrices with elements:

$$(a_{11} \geq a_{21}) \wedge (b_{11} \geq b_{12})$$

For the second player $((1,0), (0,1))$ is plausible, thus \mathfrak{P}'_2 remains unchanged, that is $\mathfrak{P}'_2 = \mathfrak{P}_2$.

In case that $ENE = ((1,0), (1/3, 2/3))$ is been publicly announced. Then \mathfrak{P}'_1 is the set of payoff matrices with elements:

$$(a_{11} \geq a_{21}) \wedge (b_{11} = b_{12})$$

As $(1,0)$ strategy for player 1 is anticipated from player 2 gain \mathfrak{P}'_2 remains unchanged.

In both cases the epistemic evolution $EE(\mathcal{E}^{(0)}, ((1,0), (0,1)))$, provide the epistemic game $\mathcal{E}'^{(0)} = \langle N, S, VA, P^0, \mathfrak{P}'_1, \mathfrak{P}_2 \rangle$. ■

Adaptive evolution

Now let us turn our attention to the second type of evolution, the *adaptive evolution*. The adaptive evolution takes place when the players update their payoff matrices based on the information received by the environment (i.e., the actual payoffs).

We first define the replacement operation for matrices:

Definition 43. Consider a payoff matrix P , and a value assignment function VA . We denote by $P \oplus VA$ the matrix P' , such that, for all $i \in N$, and $\vec{v} \in Pos$:

$$P'_i(\vec{v}) = \begin{cases} VA_i(\vec{v}) & , \text{ when } \vec{v} \in Dom(VA) \\ P_i(\vec{v}) & , \text{ otherwise} \end{cases}$$

The matrix $P \oplus VA$ is called the VA -replacement of P .

We extend Definition 43 to apply for sets of payoff matrices, in the obvious manner:

Definition 44. For a set of payoff matrices \mathfrak{P} , and a value assignment function VA , we set:

$$\mathfrak{P} \oplus VA = \{P \oplus VA \mid P \in \mathfrak{P}\}$$

As before, $\mathfrak{P} \oplus VA$ is called the VA -replacement of \mathfrak{P} .

Observe that $\mathfrak{P} \oplus VA \vdash VA$.

Finally, we apply Definition 44 to define the adaptive evolution of an epistemic misinformation game, when some new payoffs are learnt. The learnt payoffs are determined by the support of some strategy profile $\sigma \in \Sigma$:

Definition 45. Consider an epistemic misinformation game $\mathcal{E} = \langle N, S, VA, P^0, \mathfrak{P}_1, \dots, \mathfrak{P}_{|N|} \rangle$, and some strategy profile of P^0 , $\sigma = (\sigma_1, \dots, \sigma_{|N|})$. Set VA^0 to be the value assignment function such that $Dom(VA^0) = supp(\sigma) \setminus Dom(VA)$, and $VA^0_q(\vec{v}) = P^0_i(\vec{v})$ for all $\vec{v} \in Dom(VA^0)$. The adaptive evolution of \mathcal{E} by σ , denoted by $AD(\mathcal{E}, \sigma)$ is the epistemic misinformation game: $AD(\mathcal{E}, \sigma) = \langle N, S, VA', P^0, \mathfrak{P}'_1, \dots, \mathfrak{P}'_{|N|} \rangle$, where, for any $i \in N$, $\mathfrak{P}'_i = \mathfrak{P} \oplus VA^0$ and $VA' = VA \uplus VA^0$.

Example 5.3 (continued). In case that $ENE = ((1, 0), (0, 1))$ is been publicly announced. Then \mathfrak{P}'_1 is the set of payoff matrices with elements:

$$(a_{11} \geq 3) \wedge (b_{11} \geq 3)$$

For the second player $((1, 0), (0, 1))$ is plausible, thus

$$\mathfrak{P}_2 \oplus VA((1, 2)) = \left(\begin{array}{cc} (2, 1) & (0, 3) \\ (0, 0) & (1, 2) \end{array} \right)$$

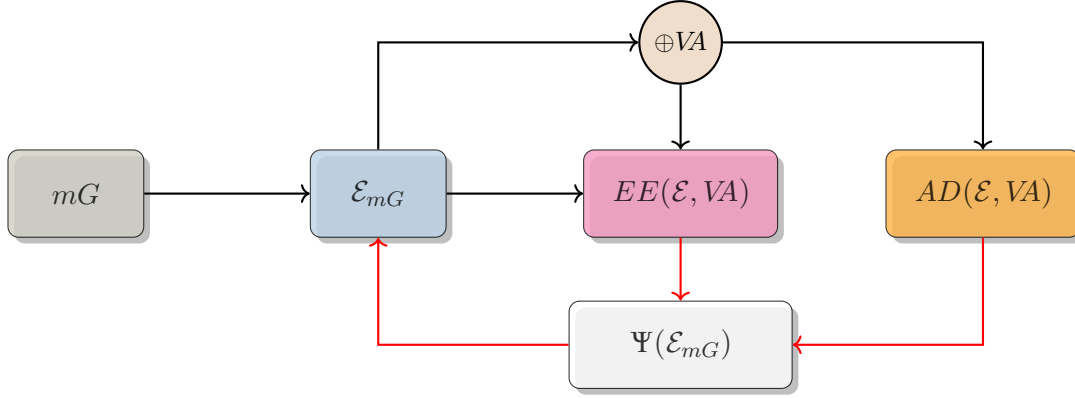


Figure 5.5: Revision for each player when information about position VA is publicly announced.

, where $VA_1((1, 2)) = 0$ and $VA_2((1, 2)) = 3$. The adaptive evolution of epistemic misinformation game \mathcal{E} is $AD(\mathcal{E}'^{(0)}, ((0, 1), (0, 1))) = AD(EE(\mathcal{E}^{(0)}, ((0, 1), (0, 1))))$ and provides the epistemic game $\mathcal{E}^{(1)} = \langle N, S, VA, P^0, \mathfrak{P}'_1, \mathfrak{P}'_2 \rangle$. \square

5.5.5 Epistemic game playing

We now have the necessary tools to define how players play and how their knowledge about the game evolves as time goes by. This evolution happens in turns. In each turn, the ENE is announced, and the players “learn” their new payoff matrix as determined by the EE and AD operations above. An important issue to consider is what happens when the ENE is not unique. In this case, our operators consider all possible $ENEs$ in parallel (independent) branches, thereby exploring all relevant possibilities.

In more details, the interaction starts with one epistemic misinformation game $\mathcal{E} = \langle N, S, VA, P^0, \mathfrak{P}_1, \dots, \mathfrak{P}_{|N|} \rangle$, where $Dom(VA) = \emptyset$ (i.e., $VA = VA_\emptyset$), and each of the players’ subjective view (\mathfrak{P}_i) is a singleton set consisting of the respective player’s subjective game (as determined by the misinformation game provided in the input). Then, the $ENEs$ are computed. For each $\sigma \in ENE(\mathcal{E})$, the operators EE and AD are applied on the epistemic misinformation game (for this σ) to produce the new epistemic game, see Figures 5.5-5.6. As there may be multiple $ENEs$, the original epistemic misinformation game may spawn multiple new ones. The process then starts over. Note that, in subsequent turns, when multiple epistemic misinformation games exist, the process is executed for each one independently.

Let us now formalize the above ideas:

Definition 46. Consider an epistemic misinformation game $\mathcal{E}_{mG} = \langle N, S, VA, P^0, \mathfrak{P}_1, \dots, \mathfrak{P}_{|N|} \rangle$ and some strategy profile of P^0 $\sigma = (\sigma_1, \dots, \sigma_{|N|})$. The epistemic adaptive evolution of \mathcal{E}_{mG} by σ is defined as follows: $\Psi(\mathcal{E}_{mG}, \sigma) = AD(EE(\mathcal{E}_{mG}, \sigma), \sigma)$.

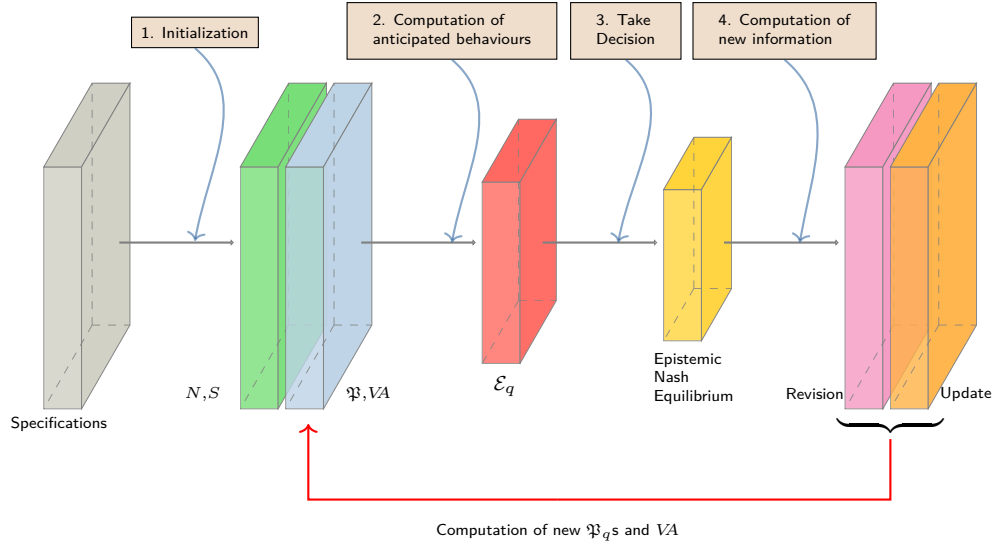


Figure 5.6: Schematic representation of the epistemic adaptive evolution with the epistemic evolution and the adaptive update of the agents.

We can now apply this definition for the generic case, when we may have more than one epistemic misinformation games. In other words, we generalize Definition 46 to be applicable over sets of epistemic misinformation games:

Definition 47. Consider a set of epistemic misinformation games \mathcal{E}_{mG} . The epistemic adaptive evolution of \mathcal{E}_{mG} , denoted by $\Psi(\mathcal{E}_{mG})$ is the set of epistemic misinformation games defined as follows:

$$\Psi(\mathcal{E}_{mG}) = \{\Psi(\mathcal{E}_{mG}, \sigma) \mid \mathcal{E} \in \mathcal{E}_{mG}, \sigma \in ENE(\mathcal{E}_{mG})\}$$

Finally, we define the iterative version of Ψ as follows:

Definition 48. Consider a canonical misinformation game $mG = \langle G^0, G^1, \dots, G^{|N|} \rangle$, where $G^i = \langle N, S, P^i \rangle$ for $i \in [|N|]$. We define the respective epistemic misinformation game $\mathcal{E}_{mG} = \langle N, S, VA_\emptyset, P^0, \{P^1\}, \dots, \{P^{|N|}\} \rangle$. The k -level epistemic adaptive evolution of mG (for $k \geq 0$), denoted by $\Psi^{(k)}(mG)$ is the set of epistemic misinformation games defined as follows:

$$\Psi^{(k)}(mG) = \begin{cases} \{\mathcal{E}_{mG}\} & , \text{ when } k = 0 \\ \Psi(\Psi^{(k-1)}(mG)) & , \text{ when } k > 0 \end{cases}$$

Definition 49. We say that the epistemic sequence of play for mG stabilises after k steps, if and only if $\Psi^{(k+1)}(mG) = \Psi^{(k)}(mG)$, and k is the smallest integer with this property.

A strategy σ is called a stable epistemic Nash equilibrium if and only if it is the ENE at the time when the epistemic sequence of play stabilises. Formally:

Definition 50. *Consider a misinformation game mG , some $k \geq 0$, and a set of epistemic games \mathcal{E}_{mG} such that $\mathcal{E}_{mG} = \Psi^{(k)}(mG) = \Psi^{(k+1)}(mG)$. Then, a strategy profile σ is called a stable epistemic natural equilibrium (SENE) of mG , if and only if there exists some $\mathcal{E} \in \mathcal{E}_{mG}$ such that $\sigma \in \text{ENE}(\mathcal{E})$. The set of all SENE of mG is denoted by $\text{SENE}(mG)$.*

By Definition 50, it is obvious that, if the epistemic procedure for mG stabilises after k steps, then $\text{SENE}(mG) = \{\sigma \mid \sigma \in \text{ENE}(\mathcal{E}) \text{ for some } \mathcal{E} \in \Psi^{(k)}(mG)\}$.

Chapter 6

Learning Dynamics

6.1 Introduction

Having introduced the concept of misinformation games, Chapter 4, and the following Epistemic Adaptive Evolution, Chapter 5 it is evident that the computation of the emerged equilibrium concept is crucial, and especially the computation of Nash and natural misinformed equilibrium. In this chapter we provide an alternative way to compute Nash equilibria in two players' normal-form games. Specifically, we focus on the problem of designing learning algorithms, and we restricted in the class of zero-sum games.

Although one can solve a zero-sum game by centralized linear programming algorithms, the application areas (e.g. boosting and reinforcement learning [Dai et al., 2018], and their relevance in formulating GANs in deep learning [Goodfellow et al., 2014]) highlight the importance of developing fast, iterative learning algorithms, resulting in approximate equilibria.

In the stream of work presented in Section 2.5, we are interested in methods that exhibit *last-iterate* convergence, a property most desirable from an application point of view, meaning that the strategy profile (x^t, y^t) , reached at iteration t of an iterative algorithm, converges to the actual equilibrium as $t \rightarrow \infty$. Unfortunately, many of the methods mentioned above do not satisfy this. No-regret algorithms, like the MWU method, are known to converge only in an average sense, resulting in an ε -Nash equilibrium in expectation (see [Arora et al., 2012]) for $\varepsilon > 0$. In fact, it was shown in [Bailey and Piliouras, 2018] that several MWU variants do not satisfy last-iterate convergence. Similarly, the same can be shown for many descent-based methods (see e.g., [Mertikopoulos et al., 2019]).

Fortunately though, OMWU [Daskalakis and Panageas, 2019] and OMD [Mertikopoulos et al., 2019] variants of MWU have provided positive results and have generated more interest on the behaviour and limitations of such approaches, which is not yet fully understood. Namely, they give rise to further questions, such as:

- (i) *Can we prove last-iterate convergence for other related dynamics?*
- (ii) *Can we establish faster convergence rates?*

These questions are the main focus of this chapter.

Having at hand a learning algorithm that computes Nash equilibria in two players' general zero-sum games, we can implement it in case of two players zero-sum misinformation games in order to compute: i) the natural misinformed equilibria, and ii) the stable misinformed equilibrium.

6.1.1 Informal Description

We introduce a simple yet substantially different variant of Optimistic Mirror Descent method with entropy regularization [Mertikopoulos et al., 2019], for the case of zero-sum games. OMD is an extra gradient method, i.e., it contains an intermediate gradient step before the final update step, and each iteration is characterized by its learning rate parameter, which is the same for both steps (and often the same across all iterations). Our tweak is that the intermediate step uses a different learning rate parameter from the update step in each iteration. In fact, we set this to be sufficiently large, which yields a *game-theoretic interpretation*, namely that we compute (approximate) best response strategies against the profile of the previous iteration, as a look ahead move. Then, during the final update step, we apply multiplicative weights updates by rewarding more the pure strategies that perform better against the best responses that we found in the intermediate step. Consequently, we refer to this OMD variant as Forward-Looking Best-Response - Multiplicative Weights Update (FLBR-MWU) method.

At first sight, this may look counter-intuitive, since learning rates are usually kept small in classic MWU algorithms and, more generally, in any kind of iterative gradient-type optimization algorithms (apart from the notable exception of [Bailey and Piliouras, 2019]). However, our theoretical and experimental study reveal the following promising findings:

- In Subsection 6.2, we investigate theoretically the convergence properties of FLBR-MWU. If η is the standard learning rate parameter used in the update step, and ξ is the corresponding parameter in the intermediate step, then FLBR-MWU exhibits last-iterate convergence for games with a unique equilibrium, when ξ is sufficiently large and $\eta\xi < 1$. Our proof employs a similar methodology to [Daskalakis and Panageas, 2019], adapting convergence tools from the field of dynamical systems. Our method also appears to attain faster convergence, quantified in terms of η , compared to OMD and OMWU. In particular, we prove that the decrease in the divergence from the equilibrium is at least $\Omega(\eta^{1+1/\rho})$ per iteration, for any $\rho > 1$, until we reach an approximate $O(\eta^{1/\rho})$ -equilibrium, by which time, our rule becomes a contraction map (see also Figure 6.2). This improves on the $\Omega(\eta^3)$ bound established for OMWU in [Daskalakis and Panageas, 2019]. Although our bounds do not translate into bounds with respect to time, we suspect a linear convergence rate is highly likely (supported also by our experiments). This has been recently established for OMWU in [Wei et al., 2021], and

is left as an open problem for FLBR-MWU.

- In Subsection 6.3, we perform numerical experiments, using randomly generated data, comparing FLBR-MWU with OMWU*. Our experiments reveal that in practice our method achieves indeed a much faster convergence rate, showing an average speedup by a factor of 10 for small size games and up to hundreds, or even higher, for larger games compared to OMWU.

6.1.2 A Revealing Example

The plot in Figure 6.1 attempts to demonstrate the differences we observed in convergence between the proposed FLBR-MWU and the OMWU dynamics in a qualitative manner. The two phases of the learning dynamics (decrease of divergence, followed by contraction) are highlighted along with the regions of convergence.

A quantitative presentation is shown in the lower plots of Figure 6.2 which depict the convergence behavior of MWU (blue lines), OMWU (red lines) and FLBR-MWU (black lines) for a random realization of a 10×10 payoff matrix with learning rate $\eta = 0.1$. We provide two measures of convergence, the Kullback-Leibler Divergence (D_{KL} , see Appendix B.2) of the Nash equilibrium with respect to the learning dynamics (lower left panel) and the respective l_1 norm difference (lower right panel), which reveal different aspects of the dynamics.

As expected by [Bailey and Piliouras, 2018], MWU fails to converge and a smaller learning rate η would not fix this issue. OMWU does converge but in a very slow pace requiring an enormous number of steps (see also the supplementary material for a longer simulation). On the other hand, FLBR-MWU converges until the machine precision is hit, as revealed by both D_{KL} and l_1 metrics. Indeed, FLBR-MWU is able to escape from the D_{KL} plateau (seen in the lower left panel), where the learning dynamics are moving towards a direction with slow D_{KL} decline and ultimately converges to Nash equilibrium in an oscillatory manner with decreasing amplitude (damped oscillations), as it is evident from the l_1 -norm difference (lower right panel). Moreover, this realization underlines the need for a two-step proof of FLBR-MWU's convergence as in [Daskalakis and Panageas, 2019]. Overall, FLBR-MWU dynamics has more than one order of magnitude faster convergence rate relative to OMWU and furthermore tolerates larger values for the learning rate, thus the speed of equilibrium computation is significantly accelerated.

*We note that for the case of zero-sum games, it has been shown in [Wei et al., 2021] that OMWU can be seen as a variant of OMD with entropy regularization.

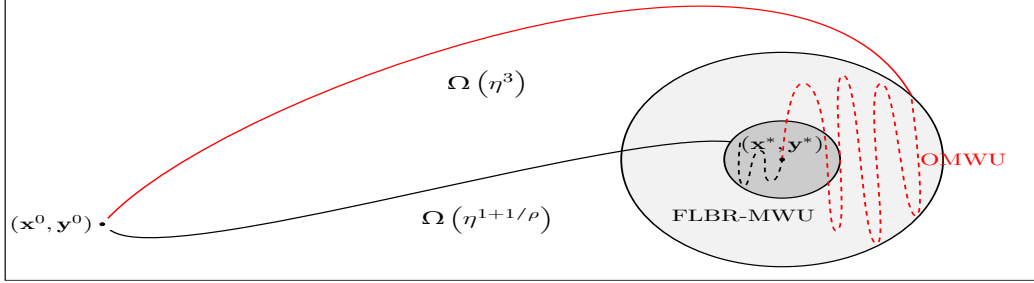


Figure 6.1: Schematic representation of the convergence path of OMWU (red) and FLBR-MWU (black), starting from the uniform point $(\mathbf{x}^0, \mathbf{y}^0)$.

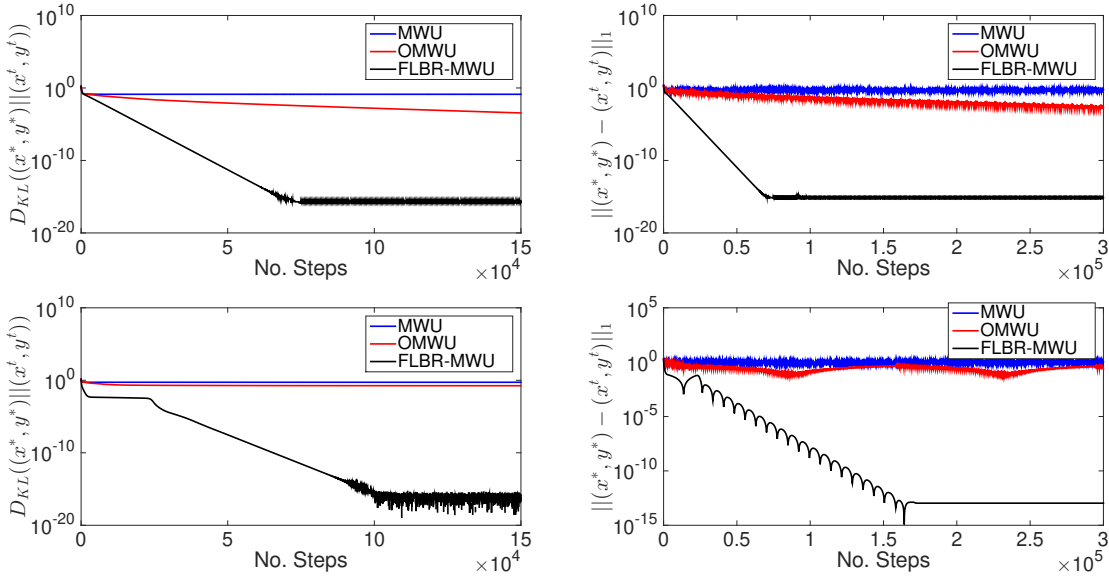


Figure 6.2: A random realization of the learning dynamics for three variants of MWU.

6.2 Forward Looking Best-Response Multiplicative Weights Update Method

6.2.1 Definition of the Dynamics

We now present the method studied in this work, which we refer to as Forward Looking Best-Response Multiplicative Weights Update method (FLBR-MWU). We provide first a short description of the main idea behind the dynamics. This is an extra gradient method and each iteration has an intermediate and a final step. Suppose that starting from some initial

profile, we reach the profile (x^{t-1}, y^{t-1}) by the end of iteration $t - 1$. In the intermediate step of iteration t , we compute a strategy \hat{x}^t for the row player (resp. \hat{y}^t for the column player), which is an approximate best-response strategy to y^{t-1} (resp. to x^{t-1}). This serves as a look ahead step of what would be the currently optimal choices. In the final step of iteration t , we compute the new mixed strategy x^t for the row player, by performing multiplicative weights updates, but after assuming that the opponent was playing \hat{y}^t .

Formally, the first step of the dynamics, denoted as the intermediate best response (IBR) step, is defined below, at iteration t , and for all $i, j \in [n]$, given a non-negative parameter $\xi \in \mathbb{R}^+$ (ξ will be chosen sufficiently large, as will become clear from Lemma 8).

$$\hat{x}_i^t = x_i^{t-1} \cdot \frac{e^{\xi e_i^T R y^{t-1}}}{\sum_{j=1}^n x_j^{t-1} e^{\xi e_j^T R y^{t-1}}}, \quad \hat{y}_j^t = y_j^{t-1} \cdot \frac{e^{-\xi e_j^T R^T x^{t-1}}}{\sum_{i=1}^n y_i^{t-1} e^{-\xi e_i^T R^T x^{t-1}}}. \quad (6.1)$$

The second step, which updates the profile (x^{t-1}, y^{t-1}) to (x^t, y^t) is below, given the learning rate parameter $\eta \in (0, 1)$. We assume that we use the same fixed constants η and ξ in all iterations.

$$x_i^t = x_i^{t-1} \cdot \frac{e^{\eta e_i^T R \hat{y}^t}}{\sum_{j=1}^n x_j^{t-1} e^{\eta e_j^T R \hat{y}^t}}, \quad y_j^t = y_j^{t-1} \cdot \frac{e^{-\eta e_j^T R^T \hat{x}^t}}{\sum_{i=1}^n y_i^{t-1} e^{-\eta e_i^T R^T \hat{x}^t}}. \quad (6.2)$$

Remark 1. By setting $\xi = \eta$ in Equation (6.1) above, the proposed method becomes the same as OMD with entropic regularization [Mertikopoulos et al., 2019], which can also be viewed as OMWU [Wei et al., 2021]. In our method however, η and ξ differ substantially across both theoretical and experimental results.

6.2.2 Main Results

We consider games with a unique Nash equilibrium, as in [Daskalakis and Panageas, 2019], since it has been argued that the set of zero-sum games with non-unique equilibrium has Lebesgue measure equal to zero [van Damme, 1991]. For convenience, we also assume that the initial strategy profile consists of the uniform distribution for each player. However, our results hold for any fully-mixed initial profile, with a non-zero probability to all pure strategies.

The main result of our work is the following theorem.

Theorem 2. Consider a zero-sum game with a unique Nash equilibrium (x^*, y^*) . Starting with the uniform distribution for each player, the FLBR-MWU dynamics attain last-iterate convergence to the Nash equilibrium, i.e., $\lim_{t \rightarrow \infty} (x^t, y^t) = (x^*, y^*)$, for sufficiently small η , and big enough ξ , as long as $\eta\xi < 1$.

The goal of the remaining subsection is to establish the proof of Theorem 2. Towards this, we start with the choice of ξ . The next lemma provides the important observation, that as $\xi \rightarrow \infty$, the strategy \hat{x}^t , computed in the first step of iteration t , becomes a best response against y^{t-1} (analogously for \hat{y}^t).

Lemma 8. *Given any round t of the dynamics, let \hat{x}^t, \hat{y}^t be the strategies produced by the first step of iteration t . As $\xi \rightarrow +\infty$, then \hat{x}^t becomes a best-response strategy against y^{t-1} (similarly for \hat{y}^t against x^{t-1}).*

Proof. Fix t and let us consider the formula that produces the coordinates of \hat{x}^t , given x^{t-1}, y^{t-1} . For simplicity in writing, we drop the superscript $t-1$ and refer to x, y as the strategies of the two players computed at the end of the previous round. Focusing on the row player (the same argument follows for the column player too), we know that

$$\hat{x}_i^t = x_i \cdot \frac{e^{\xi e_i^T R y}}{\sum_{j=1}^n x_j e^{\xi e_j^T R y}}.$$

We want to compute for every i the limit $\lim_{\xi \rightarrow \infty} \hat{x}_i^t$. Using the Taylor expansion for the exponential terms in the above equation, the limit we are interested in can be written as:

$$\lim_{\xi \rightarrow \infty} \hat{x}_i^t = x_i \cdot \lim_{\xi \rightarrow \infty} \lim_{\ell \rightarrow \infty} \frac{\sum_{k=0}^{\ell} \frac{(\xi e_i^T R y)^k}{k!}}{\sum_{j=1}^n x_j \sum_{k=0}^{\ell} \frac{(\xi e_j^T R y)^k}{k!}} = x_i \cdot \lim_{\ell \rightarrow \infty} \lim_{\xi \rightarrow \infty} \frac{\sum_{k=0}^{\ell} (\xi e_i^T R y)^k}{\sum_{j=1}^n x_j \sum_{k=0}^{\ell} (\xi e_j^T R y)^k}.$$

In order to compute first the limit w.r.t. ξ , notice that both the nominator and the denominator can be written as polynomials of ξ where the highest power is ξ^ℓ . To proceed, let also $B(y)$ denote the set of pure best response strategies of the row player against y . We now have:

$$\lim_{\xi \rightarrow \infty} \hat{x}_i^t = x_i \cdot \lim_{\ell \rightarrow \infty} \frac{(e_i^T R y)^\ell}{\sum_{j=1}^n x_j (e_j^T R y)^\ell} = x_i \cdot \lim_{\ell \rightarrow \infty} \frac{(e_i^T R y)^\ell}{\sum_{j \in B(y)} x_j (e_j^T R y)^\ell + \sum_{j \notin B(y)} x_j (e_j^T R y)^\ell}.$$

We can now consider two cases. Suppose first that $i \in B(y)$. For any $j \in [n]$, let $p_j = e_j^T R y$. Since i is a best response pure strategy, we have that $p_i = p_j$ for any $j \in B(y)$, and $p_i > p_j$ for $j \notin B(y)$. We can now conclude that as $\ell \rightarrow \infty$ we have:

$$\lim_{\ell \rightarrow \infty} \hat{x}_i^t = x_i \cdot \lim_{\ell \rightarrow \infty} \frac{1}{\sum_{j \in B(y)} x_j (p_j/p_i)^\ell + \sum_{j \notin B(y)} x_j (p_j/p_i)^\ell} = \frac{x_i}{\sum_{j \in B(y)} x_j},$$

where the last equality above holds because $p_j/p_i < 1$ for any $j \notin B(y)$ and the second sum in the denominator tends to 0. In a similar way we can also show that when $i \notin B(y)$, the limit is zero. Hence, as $\xi \rightarrow \infty$, the strategy \hat{x} contains only best responses of y in its support, and therefore forms a best response too. \square

In the sequel, we assume that ξ has been chosen sufficiently large, so that \hat{x}^t is an ϵ -best response with $\epsilon \rightarrow 0$. For appropriate choices of ξ in practice, we refer to the discussion in Section 6.3.

The proof of Theorem 2 is split into 3 parts. The first part establishes that after a certain number of iterations, the dynamics reach a profile (x^t, y^t) , that is an $O(\eta^{1/\rho})$ -Nash equilibrium with $\rho > 1$. The second part shows that the profile (x^t, y^t) lies within a neighborhood of the actual equilibrium (x^*, y^*) . Finally, the last part shows that the update rule of FLBR-MWU is a contracting map, i.e., once we are within a neighborhood of (x^*, y^*) , the dynamics converge to their fixed point, which directly implies last-iterate convergence. These three parts are established in Theorems 3, 4 and 6 respectively. The structure of the proof is similar to the convergence proof of OMWU in [Daskalakis and Panageas, 2019]. There are however differences in various parts of the analysis. Most importantly, in the first part, we are able to establish a better convergence rate to an approximate equilibrium, whereas OMWU achieves an $\Omega(\eta^3)$ decrease rate. Furthermore, in the third part, the analysis of our Jacobian matrix (proof of Theorem 6) is also different since we are analyzing sufficiently different dynamics.

To proceed with the first part of the proof, we will use the *Kullback-Leibler (KL) divergence* as a measure of progress. The KL divergence quantifies the similarity between two distributions, and here we will consider the divergence between a profile (x^t, y^t) and the equilibrium (x^*, y^*) , which equals:

$$D_{KL}((x^*, y^*) || (x^t, y^t)) = \sum_{i=1}^n x_i^* \ln(x_i^* / x_i^t) + \sum_{j=1}^n y_j^* \ln(y_j^* / y_j^t). \quad (6.3)$$

Note that by the initialization and the definition of the dynamics, $x_i^t > 0$, $y_j^t > 0$ for any given t , and any i, j , so that the logarithmic terms above are well-defined.

Theorem 3. *Consider a zero-sum game with a unique Nash equilibrium (x^*, y^*) . Assume that we run the FLBR-MWU dynamics with the uniform distribution as the initial strategy for both players, and using a sufficiently small η and a big enough ξ . Then, for any $\rho > 1$, the KL divergence $D_{KL}((x^*, y^*) || (x^t, y^t))$ decreases at every iteration with a rate of at least $\Omega(\eta^{1+1/\rho})$, until we reach an $O(\eta^{1/\rho})$ -Nash equilibrium of the game.*

Proof. Let (x^*, y^*) be the Nash equilibrium of the game, and let v be the value of the game, $v = (x^*)^T R y^*$. We take the difference of the KL divergences between two consecutive iterations:

$$\begin{aligned} D_{KL}((x^*, y^*) || (x^t, y^t)) - D_{KL}((x^*, y^*) || (x^{t-1}, y^{t-1})) = \\ - \left(\sum_{i=1}^n x_i^* \ln(x_i^t / x_i^{t-1}) + \sum_{j=1}^n y_j^* \ln(y_j^t / y_j^{t-1}) \right). \end{aligned}$$

We show that this difference is negative and we quantify the decrease in the KL divergence,

till we reach an $O(\eta^{1/\rho})$ -Nash equilibrium. Analytically, we have that

$$\begin{aligned}
& D_{KL}((x^*, y^*) || (x^t, y^t)) - D_{KL}((x^*, y^*) || (x^{t-1}, y^{t-1})) \\
&= -\sum_{i=1}^n x_i^* \ln e^{\eta e_i^T R \hat{y}^t} + \ln \left(\sum_{i=1}^n x_i^{t-1} e^{\eta e_i^T R \hat{y}^t} \right) - \sum_{j=1}^n y_j^* \ln e^{-\eta e_j^T R^T \hat{x}^t} + \ln \left(\sum_{j=1}^n y_j^{t-1} e^{-\eta e_j^T R^T \hat{x}^t} \right) \\
&= -\eta x^{*T} R \hat{y}^t + \eta (y^*)^T R^T \hat{x}^t + \eta (x^{t-1})^T R y^{t-1} - \eta (y^{t-1})^T R^T x^{t-1} \\
&\quad + \ln \left(\sum_{i=1}^n x_i^{t-1} e^{\eta e_i^T R \hat{y}^t - \eta (x^{t-1})^T R y^{t-1}} \right) + \ln \left(\sum_{j=1}^n y_j^{t-1} e^{-\eta e_j^T R^T \hat{x}^t + \eta (y^{t-1})^T R^T x^{t-1}} \right).
\end{aligned}$$

Notice that in the last expression above, the third term $(\eta (x^{t-1})^T R y^{t-1})$ cancels out with the fourth term. Also, since (x^*, y^*) is an equilibrium, it holds that $x^{*T} R \hat{y}^t \geq v$ and $(y^*)^T R^T \hat{x}^t \leq v$. Therefore, the first and second terms also cancel out and yield an upper bound with the two logarithmic terms.

We now apply the Taylor expansion of e^x . For convenience, let $p_i(\eta) = \eta(e_i^T R \hat{y}^t - (x^{t-1})^T R y^{t-1})$, and let $q_j(\eta) = \eta(-e_j^T R^T \hat{x}^t + (y^{t-1})^T R^T x^{t-1})$. The difference of the KL divergences is upper bounded by

$$\begin{aligned}
& \ln \left(1 + \eta((x^{t-1})^T R \hat{y}^t - (x^{t-1})^T R y^{t-1}) + \sum_{i=1}^n x_i^{t-1} \sum_{k=2}^{\infty} \frac{(p_i(\eta))^k}{k!} \right) \\
& + \ln \left(1 + \eta(-(y^{t-1})^T R^T \hat{x}^t + (y^{t-1})^T R^T x^{t-1}) + \sum_{j=1}^n y_j^{t-1} \sum_{k=2}^{\infty} \frac{(q_j(\eta))^k}{k!} \right).
\end{aligned}$$

It is easy to see that $|p_i(\eta)| \leq \eta$ and $|q_j(\eta)| \leq \eta$. This means that for any $k \geq 2$ (i.e., for both odd and even values of k), $(p_i(\eta))^k \leq \eta^k$ and $(q_j(\eta))^k \leq \eta^k$. By using the geometric series, we have that $\sum_{k=2}^{\infty} \frac{(p_i(\eta))^k}{k!} \leq \eta^2/(1-\eta)$, and similarly for the series concerning $q_j(\eta)$. If we also use the inequality $\ln(x) \leq x - 1$, we obtain the following sequence of steps.

$$\begin{aligned}
& D_{KL}((x^*, y^*) || (x^t, y^t)) - D_{KL}((x^*, y^*) || (x^{t-1}, y^{t-1})) \\
& \leq \ln \left(1 + \eta((x^{t-1})^T R \hat{y}^t - (x^{t-1})^T R y^{t-1}) + \frac{\eta^2}{(1-\eta)} \right) \\
& \quad + \ln \left(1 + \eta(-(y^{t-1})^T R^T \hat{x}^t + (y^{t-1})^T R^T x^{t-1}) + \frac{\eta^2}{(1-\eta)} \right) \tag{6.4} \\
& \leq \eta((x^{t-1})^T R \hat{y}^t - (x^{t-1})^T R y^{t-1} + (x^{t-1})^T R y^{t-1} - \hat{x}^t R y^{t-1}) + 2 \frac{\eta^2}{(1-\eta)} \\
& = -\eta(\varepsilon_1 + \varepsilon_2) + 4\eta^2 \leq -\eta(\max\{\varepsilon_1, \varepsilon_2\}) + 4\eta^2,
\end{aligned}$$

where $\varepsilon_1 = \hat{x}^t R y^{t-1} - (x^{t-1})^T R y^{t-1}$, $\varepsilon_2 = (x^{t-1})^T R \hat{y}^t - (x^{t-1})^T R y^{t-1}$, and the last inequality holds because $\eta \leq 1/2$. Let us look now more carefully at ε_1 (an analogous

argument holds for ε_2). The term ε_1 expresses the additional benefit for the row player, if at the profile (x^{t-1}, y^{t-1}) , he deviates to \hat{x}^t . By Lemma 8, we know that as $\xi \rightarrow \infty$, then \hat{x}^t tends to his best response against y^{t-1} . Hence when we select ξ sufficiently large, ε_1 tends to the best possible deviation gain of the row player at the profile (x^{t-1}, y^{t-1}) (resp. for ε_2 and the column player).

To finish the proof, suppose that the profile (x^{t-1}, y^{t-1}) is not an $O(\eta^{1/\rho})$ -Nash equilibrium. Then there exists a deviation that provides additional gain of $\Omega(\eta^{1/\rho})$ to one of the players. This implies that $\max\{\varepsilon_1, \varepsilon_2\} = \Omega(\eta^{1/\rho})$. Hence, by (6.4), and since $\eta < 1$, we can see that as long as we have not reached an $O(\eta^{1/\rho})$ -Nash equilibrium, the KL divergence will keep decreasing by at least $\eta\Omega(\eta^{1/\rho}) - 4\eta^2 = \Omega(\eta^{1+1/\rho})$. As the KL divergence cannot decrease forever, eventually, our dynamics will reach an $O(\eta^{1/\rho})$ -Nash equilibrium. \square

Consider now the first iteration t of the dynamics, where (x^t, y^t) forms an $O(\eta^{1/\rho})$ -Nash equilibrium for some fixed $\rho > 1$. The next step is to show that if we make η small enough, this profile falls within a neighborhood of the equilibrium (x^*, y^*) .

Theorem 4. *Let (x^*, y^*) be the unique Nash equilibrium of the zero-sum game, and let (x^t, y^t) be the first profile reached by the dynamics, that is an $O(\eta^{1/\rho})$ -Nash equilibrium for some $\rho > 1$. Then*

$$\lim_{\eta \rightarrow 0} \|(x^*, y^*) - (x^t, y^t)\|_1 = 0,$$

Proof. The proof is based on the following lemma, shown in [Etesami and Yannakakis, 2010], which we state here for the case of zero-sum games:

Lemma 9. *Consider a zero-sum game given by matrix R with a unique Nash equilibrium (x^*, y^*) , and let $|R|$ be the number of bits needed for the representation of R . There exists a polynomial p such that for every $\delta > 0$, every ε -Nash equilibrium (x, y) satisfies that $|x_i^* - x_i| < \delta$, as long as $\varepsilon \leq 1/2^{p(|R| + \text{size}(\delta))}$, where $\text{size}(\delta) = O(\log(1/\delta))$ is the number of bits needed for representing δ .*

By the assumptions in the statement of Theorem 4, we fix $\varepsilon = c \cdot \eta^{1/\rho}$, for some constant c , so that (x^t, y^t) is an ε -Nash equilibrium. We claim that there exists $\delta(\eta)$ such that ε and $\delta(\eta)$ satisfy the inequality stated in Lemma 9. In particular, by looking more carefully at the desired inequality and solving with respect to δ , one can construct a function $\delta(\eta)$, such that for the given ε we have selected, it holds that

$$\varepsilon \leq 1/2^{p(|R| + \text{size}(\delta(\eta)))} \text{ and } \lim_{\eta \rightarrow 0} \delta(\eta) = 0.$$

Hence, we can now apply Lemma 9 and obtain that for any ε -Nash equilibrium (x, y) we have that $|x_i^* - x_i| \leq \delta(\eta)$ and $|y_i^* - y_i| \leq \delta(\eta)$. The proof now of Theorem 4 is immediate, since $\|(x^*, y^*) - (x^t, y^t)\|_1 = \sum_{i=1}^n |x_i^* - x_i| + \sum_{i=1}^n |y_i^* - y_i| \leq 2n \cdot \delta(\eta)$, which goes to 0 as $\eta \rightarrow 0$. \square

The next and final step of our proof is to show that our dynamics induce a contracting map. An update rule with a fixed point x is called a contraction, if there exists a region U around x , such that for any starting point in U , the rule converges to its fixed point as $t \rightarrow \infty$. In our case, the Nash equilibrium (x^*, y^*) of the game is a fixed point of the FLBR-MWU dynamics and Theorem 4 guarantees that we can reach a neighborhood around (x^*, y^*) . To proceed, we state a sufficient condition for a dynamical system to converge to its fixed point.

Theorem 5 (see [Galor, 2007]). *Let x^* be a fixed point for the dynamical system $x^{(t+1)} = g(x^{(t)})$. If all eigenvalues of the Jacobian matrix of g at x^* have absolute value less than one, then there exists a neighborhood U of x^* such that for all $x \in U$, g converges to x^* , starting from x .*

Using Theorem 5, we show the following theorem.

Theorem 6. *The update rule of FLBR-MWU is a contraction, as long as $\eta\xi < 1$, i.e. $\lim_{t \rightarrow \infty} (x^t, y^t) = (x^*, y^*)$.*

Proof. First we describe first the discrete dynamical system that captures the FLBR-MWU dynamics, and we will prove that for an appropriate norm of the Jacobian matrix of the system, its value is less than one[†]. The update rule φ of FLBR-MWU is

$$\begin{aligned}\varphi(x, y) &= (\varphi_1(x, y), \varphi_2(x, y)), \text{ where} \\ \varphi_{1,i}(x, y) &= (\varphi_1(x, y))_i = x_i \frac{e^{\eta e_i^T R f(x, y)}}{\sum_{\ell} x_{\ell} e^{\eta e_{\ell}^T R f(x, y)}}, \\ \varphi_{2,i}(x, y) &= (\varphi_2(x, y))_i = y_i \frac{e^{-\eta e_i^T R^T h(x, y)}}{\sum_{\ell} y_{\ell} e^{-\eta e_{\ell}^T R^T h(x, y)}},\end{aligned}\tag{6.5}$$

where $f(x, y)$ and $h(x, y)$ are column vectors with $(f(x, y))_i = y_i \frac{e^{-\xi e_i^T R^T x}}{\sum_{\ell} y_{\ell} e^{-\xi e_{\ell}^T R^T x}}$, and $(h(x, y))_i = x_i \frac{e^{\xi e_i^T R y}}{\sum_{\ell} x_{\ell} e^{\xi e_{\ell}^T R y}}$, for all $i \in \{1, \dots, n\}$.

Clearly, the dynamics of FLBR-MWU are captured by $(x^{t+1}, y^{t+1}) = \varphi(x^t, y^t)$. The Jacobian of φ is a $2n \times 2n$ matrix, which can be written in the form of a 2×2 block matrix, as follows:

$$J = \begin{pmatrix} \frac{\partial \varphi_1}{\partial \mathbf{x}} & \frac{\partial \varphi_1}{\partial \mathbf{y}} \\ \frac{\partial \varphi_2}{\partial \mathbf{x}} & \frac{\partial \varphi_2}{\partial \mathbf{y}} \end{pmatrix}\tag{6.6}$$

In order to use Theorem 5 and prove that ϕ is a contraction, we need to argue about the eigenvalues of J at the equilibrium (x^*, y^*) . Towards this, in Appendix B.4, we provide the

[†]Besides [Galor, 2007], readers could advise Chapter 7 [Quarteroni et al., 2006].

exact form of each entry of J at (x^*, y^*) (after some simplification steps by exploiting the fact that (x^*, y^*) is an equilibrium).

We analyze first the eigenvalues that are derived by the rows of J that correspond to $\varphi_{1,i}$ for some $i \notin \text{supp}(x^*)$ and to $\varphi_{2,i}$ for some $i \notin \text{supp}(y^*)$. Let $x^{*T} R y^* = v$ be the value of the game. By referring to Appendix B.4, we have that for any $i \notin \text{supp}(x^*)$:

$$\frac{\partial \varphi_{1,i}}{\partial x_i}(x^*, y^*) = \frac{e^{\eta e_i^T R y^*}}{e^{\eta v}}, \quad \frac{\partial \varphi_{1,i}}{\partial x_j}(x^*, y^*) = 0 \text{ for any } i \neq j, \text{ and } \frac{\partial \varphi_{1,i}}{\partial y_j}(x^*, y^*) = 0, \text{ for any } j.$$

Hence, the i -th row of the upper block of J has only one non-zero entry, namely, the diagonal element, provided that $i \notin \text{supp}(x^*)$. Thus, $\frac{e^{\eta e_i^T R y^*}}{e^{\eta v}}$ is an eigenvalue of J at (x^*, y^*) . We note also that $^\dagger e_i^T R y^* < v$ for $i \notin \text{supp}(x^*)$, hence $|\frac{\partial \varphi_{1,i}}{\partial x_i}(x^*, y^*)| < 1$. Analogously, for $i \notin \text{supp}(y^*)$ we have that $\frac{\partial \varphi_{2,i}}{\partial y_i}(x^*, y^*) = \frac{e^{-\eta e_i^T R^T x^*}}{e^{-\eta v}}$, whereas all other partial derivatives of $\varphi_{2,i}$ are zero. Thus, $\frac{e^{-\eta e_i^T R^T x^*}}{e^{-\eta v}}$ is also an eigenvalue of J , with $|\frac{e^{-\eta e_i^T R^T x^*}}{e^{-\eta v}}| < 1$, since $e_i^T R^T x^* > v$ for $i \notin \text{supp}(y^*)$, by Lemma 18 in Appendix B.3.

We now focus on the rows and columns that correspond to the support of x^* and y^* . We denote this submatrix as \tilde{J} , with $k_1 = |\text{supp}(x^*)|$, $k_2 = |\text{supp}(y^*)|$ and $k = k_1 + k_2$. Thus, $\tilde{J} \in \mathbb{R}^{k \times k}$. It can be seen that J has eigenvalues with absolute value less than one iff the same holds for \tilde{J} as well.

Using equations (B.7) and computing $((\mathbf{1}_{k_1}, \mathbf{0}_{k_2})^T \cdot \tilde{J})_j$ for an arbitrary coordinate j , we end up with the quantity $\sum_i x_i^* \sum_k R_{ik} y_k^* R_{kj}^T - \sum_i x_i^* \sum_k x_k^* \sum_l R_{kl} y_l^* R_{lj}^T$, that equals zero. Thus, $(\mathbf{1}_{k_1}, \mathbf{0}_{k_2})$ is a left eigenvector of \tilde{J} corresponding to the zero eigenvalue. Using the same argumentation we have that $(\mathbf{0}_{k_1}, \mathbf{1}_{k_2})$ is also a left eigenvector of \tilde{J} with eigenvalue zero.

We will make use of the following claim, regarding orthogonal pairs of eigenvectors.

Claim 1. *Consider a matrix $A \in \mathbb{R}^{n \times n}$, an eigenvalue λ and a left eigenvector u^T , corresponding to λ . Then for every right eigenvector v that does not correspond to λ , it holds that $u^T v = 0$.*

The proof of the claim, which is a simple linear algebra exercise, is at the end of this section. From Claim 1, it follows that for any right eigenvector (\tilde{x}, \tilde{y}) corresponding to a nonzero eigenvalue, we have

$$\tilde{\mathbf{x}}^T \mathbf{1}_{k_1} = 0 \text{ and } \tilde{\mathbf{y}}^T \mathbf{1}_{k_2} = 0. \quad (6.7)$$

With that in hand, let us now rewrite \tilde{J} , as $\tilde{J} = J' + A$, where J' is produced by deleting the term $-x_i^*$ (resp. $-y_i^*$) from every element of the upper left (resp. lower right) block of \tilde{J} . I.e., A contains $-x_i^*$ in all entries of the i -th row in the upper left block, and $-y_i^*$ in all entries of the i -th row in the bottom right block. The other two blocks of A contain only

[†]A unique Nash equilibrium of a zero-sum game is also a quasi-strict equilibrium (Theorem 1 in [Norde, 1999]), meaning that strategies that are not in the support of the equilibrium have strictly less payoff than the best-response payoff.

zeros. Using (6.7), we can see that for every non-zero eigenvalue λ of \tilde{J} , that corresponds to a right eigenvector (\tilde{x}, \tilde{y}) , it holds that $A \cdot (\tilde{x}, \tilde{y}) = 0$, thus λ is also an eigenvalue of the matrix J' . By the equations in Appendix B.4, we can write J' as a 2×2 block matrix, as follows.

$$J' = \begin{pmatrix} I_{k_1 \times k_1} + \eta \xi D^{xx} & \eta D^{xy} \\ \eta D^{yx} & I_{k_2 \times k_2} + \eta \xi D^{yy} \end{pmatrix},$$

with

$$D_{ij}^{xx} = -x_i^* \left(\sum_k R_{ik} y_k^* R_{kj}^T - \sum_k x_k^* \sum_l R_{kl} y_l^* R_{lj}^T \right) \text{ with } i, j \in [k_1],$$

$$D_{ij}^{yy} = -y_i^* \left(\sum_k R_{kj} R_{ik}^T x_k^* - \sum_k y_k^* \sum_l R_{kl}^T x_l^* R_{lj} \right) \text{ with } i, j \in [k_2],$$

$$D_{ij}^{yx} = -y_i^* \left(R_{ij}^T - e_j^T R y^* \right) \frac{e^{\xi e_j^T R y^*}}{e^{\xi v}} \text{ with } i \in [k_2], j \in [k_1],$$

$$D_{ij}^{xy} = x_i^* \left(R_{ij} - e_j^T R^T x^* \right) \frac{e^{-\xi e_j^T R^T x^*}}{e^{-\xi v}} \text{ with } i \in [k_1], j \in [k_2].$$

We now consider the diagonal elements of D^{xx} and D^{yy} . For D^{xx} , and for any i , its i -th element along the diagonal is

$$-x_i^* \left(\sum_l R_{il}^2 y_l^* - \sum_k x_k^* \sum_l R_{kl} y_l^* R_{li}^T \right)$$

We establish the following useful property.

Lemma 10. *For any $i \in [k_1]$, $D_{ii}^{xx} < 0$, and for any $j \in [k_2]$, $D_{jj}^{yy} < 0$.*

Proof of Lemma 10. We first prove that for any i , $D_{ii}^{xx} \leq 0$. For the sake of contradiction, assume that there exists an index i , such that $D_{ii}^{xx} > 0$. This means that

$$\sum_l R_{il}^2 y_l^* < \sum_k x_k^* \sum_l R_{kl} y_l^* R_{il}.$$

To proceed, we claim that

$$v \leq \sum_l R_{il} z_l, \tag{6.8}$$

where $z_l = \frac{R_{il} y_l^*}{v}$, and $z = (z_l)_{l \in [n]}$. To see this, it is crucial to notice first that both y^* and z are probability vectors and also that $v = \sum_l R_{il} y_l^*$. Hence, the LHS and the RHS of Equation (6.8) are two different convex combinations of the R_{il} values. To go from the LHS to the RHS, we simply replace y_l^* by z_l . For each R_{il} that is itself less than v , the coefficient y_l^* is replaced by a smaller coefficient, since $z_l < y_l^*$ in this case (by the definition of z_l). On the contrary, for each R_{il} with $R_{il} > v$, it holds that $z_l > y_l^*$ (and we also have $z_l = y_l^*$ when $R_{il} = v$). Hence, we can think of the move from the LHS to the RHS of (6.8), as transferring probability mass

from the lowest valued R_{il} 's to the highest ones. Let Δ be the total amount of probability mass that was transferred. Then $\Delta = \sum_{l: R_{il} < v} (y_l^* - z_l) \geq 0$. Note that it also holds that $\Delta = \sum_{l: R_{il} > v} (z_l - y_l^*)$. If we compare now the LHS with the RHS, the RHS has a deficit of a total value of at most $\Delta \cdot v$ from the terms with $R_{il} < v$, compared to the corresponding terms of the LHS. At the same time, it has a surplus of at least $\Delta \cdot v$ from the terms with $R_{il} > v$. Combining the deficit and the surplus, this proves Equation (6.8).

Using (6.8), we can now obtain the following contradiction:

$$v \leq \sum_l R_{il} z_l < x^{*T} R z = v,$$

where the strict inequality above follows by the condition stated just before Equation (6.8), and the final equality holds since $x^{*T} R e_j = v$ for any $j \in \text{supp}(y^*)$ (and so for any $j \in \text{supp}(z)$).

Thus, we have reached a contradiction, which means that $D_{ii}^{xx} \leq 0$ for every $i \in [k_1]$. In addition, it is not difficult to see that in case $D_{ii}^{xx} = 0$ for some i , the strategy profile (i, y^*) is also a Nash equilibrium. But this would imply that there also exists a pure equilibrium formed by i and its best response, contradicting the fact that we have a unique equilibrium. Hence, D_{ii}^{xx} is strictly negative for every $i \in [k_1]$.

Similarly, the same analysis holds for the matrix D^{yy} , completing the proof of the lemma.

□

To finish the proof, we estimate an upper bound on the p -norm of J' for $p \in \mathbb{N}$. We have that

$$\begin{aligned} \|J'\|_p^p &= \sum_j \left(\sum_i |J'_{ij}|^p \right) \leq k \max_j \left(\sum_i |J'_{ij}|^p \right) \\ &\leq k (|1 + \eta \xi D_{j'j'}^{xx}|^p + \eta^p \xi^p \sum_{\substack{i=1, \\ i \neq j}}^{k_1} |D_{ij'}^{xx}|^p + \eta^p \sum_{i=k_1+1}^{k_2} |D_{ij'}^{yx}|^p) \\ &\leq k (|1 + \eta \xi D_{j'j'}^{xx}|^p + \eta^p \xi^p k_1 + \eta^p k_2), \end{aligned}$$

where j' is the column of J' that achieves the maximum in the above expression, and we assumed without loss of generality that j' belongs to $\{1, \dots, k_1\}$. We can now see that since $D_{j'j'}^{xx}$ is negative, then if $\eta \xi < 1$, and η is sufficiently small, there exists an appropriate p so that $\|J'\|_p^p < 1$. However, it is well known that the maximum absolute value of an eigenvalue of a matrix is bounded by the induced matrix norms, therefore it suffices to check that $\|J'\| < 1$ for some matrix norm, see [Quarteroni et al., 2006]. Thus, the absolute value of the maximum eigenvalue of J' is less than one, and this concludes our proof. □

Proof of Claim 1. Consider two distinct eigenvalues of A λ_1 and λ_2 , such that v is the corresponding to λ_1 left eigenvector, while u is the corresponding to λ_2 right eigenvector [Strang, 2009]. In other words, v is the corresponding to λ_1 right eigenvector for A^T . We observe that, $v^T(A^T u) = (v^T A^T)u = (Av)^T u$. So, $\lambda_1 v^T u = (A^T v)^T u = v^T(Au) = v^T \lambda_2 u = \lambda_2 v^T u$. Thus, $v^T u = 0$. □

6.3 Numerical Experiments

The theoretical results of Subsection 6.2.2 present a clear advantage of FLBR-MWU dynamics over the OMWU approach in terms of convergence speed-up. In this section, we empirically demonstrate these acceleration improvements, along with the assessment of the IBR step in our dynamics, i.e., the step defined by (6.1). Additional supporting figures and an Octave/Matlab implementation of the dynamics are provided in the supplementary material.

Nash equilibrium estimation. In order to make comparisons, we need first to compute the equilibria of the generated instances. Instead of using a linear programming solver, the equilibrium computation is performed using the proposed FLBR-MWU algorithm with $\eta = 0.05$. FLBR-MWU is an iterative approach thus a convergence criterion to ensure that the Nash equilibrium has been reached is required. We propose as a convergence criterion the D_{KL} between the update step and the IBR step of our dynamics: $D_{KL}((x^t, y^t) || (\hat{x}^t, \hat{y}^t))$. This metric is sufficient because the best response strategy at Nash equilibrium is exactly the equilibrium strategy, thus $\lim_{t \rightarrow \infty} D_{KL}((x^t, y^t) || (\hat{x}^t, \hat{y}^t)) = 0$ (for small enough η). We return the solution when the convergence criterion becomes 10^{-15} , which is approximately the machine's arithmetic precision, or when the maximum number of steps, denoted by t_{\max} ,—typically millions of steps—has been reached. In the infrequent latter case (it happened in less than 0.1%), we discard the returned solution.

Effect of the intermediate rate (ξ). In our learning dynamics, the best response strategy is approximated by the softmax function (a.k.a., the normalized exponential function or the Gibbs measure in statistical physics). Sending ξ to infinity, one out of the potentially-many best response strategies is obtained as intermediate dynamics by (6.1). However, ξ should be finite from a practical point of view. Since it appears at the exponentials' argument, very high values of ξ may result in arithmetic imprecision. Therefore, we conducted a numerical study to assess the effect of ξ on the convergence of the algorithm. Table 6.1 presents various statistics about the number of steps required for several values of ξ and for two values of the size of the payoff matrix R , with $\eta = 0.1$. We average over 10^3 repetitions using random payoff matrices whose elements are iid sampled from $\mathcal{U}([0, 1])$. Evidently, as ξ increases, the FLBR-MWU dynamics require fewer steps in order to reach a specific threshold of accuracy (set to 10^{-10} for the D_{KL} between the Nash equilibrium and the FLBR-MWU dynamics). However, the solution occasionally produces 'NaN' for values of ξ above 200 due to overflow in the exponentials[§]. Overall, values between 50 and 100 are a sufficient compromise between the best response approximation and machine precision trade-off. We set $\xi = 100$ in the remaining experiments of this section, even though larger values can be tolerated especially when both $n \gg 1$ and $x_i^*, y_j^* \ll 1$ hold.

[§]Overflow can be easily fixed by subtracting the maximum value but with an increased underflow risk.

Table 6.1: Statistics on the number of steps till convergence for various values of ξ and n . The maximum number of steps was set to $t_{\max} = 10^6$.

Matrix size	Statistic	$\xi = 20$	$\xi = 50$	$\xi = 100$	$\xi = 200$
$n = 10$	Mean	85.9K	52.7K	41.6K	57.6K
	Median	32.0K	22.3K	19.8K	18.0K
	t_{\max} was hit	0.9%	0.0%	0.0%	2.2%
$n = 20$	Mean	352.1K	233.3K	173.6K	141.4K
	Median	225.0K	123.4K	82.2K	65.6K
	t_{\max} was hit	13.5%	5.9%	3.5%	2.5%

Effect of the learning rate (η). The first row of panels in Figure 6.3 shows the D_{KL} between the Nash equilibrium and the FLBR-MWU dynamics for the same payoff matrix instance as in Figure 6.2 and for various values of the learning rate, η . The difference between the left and right panels is that for the right column of panels, the x-axis has been rescaled by multiplying each run with the respective learning rate. A linear scaling is numerically observed showing that the number of steps is effectively of order $O(\eta^{-1})$ for a fixed accuracy level. This inversely-proportional behavior is observed not only during the convergence to the approximate Nash equilibrium, but also during the contraction period. As a rule of thumb, we propose to increase the rate η , because it accelerates the convergence, but with caution since a very large η might result in an oscillatory solution, thus failing to converge (blue line in second row of panels).

Effect of the payoff matrix size (n). The rate of convergence is sensitive to the size of the payoff matrix and the number of steps is expected to substantially increase on average as the size of the game increases. We performed a numerical comparison between FLBR-MWU and OMWU to evaluate the number of steps required to achieve a predefined level of accuracy. Table 6.2 presents statistics on the number of steps for each learning algorithm computed on 100 repetitions using element-wise uniformly-sampled and iid random payoff matrices. The learning rate was set to $\eta = 0.1$. Given that FLBR-MWU requires almost twice as many calculations per iteration, relative to OMWU, it is fair to multiply the number of steps of FLBR-MWU with two and then compare it with the number of steps of OMWU. We observe that FLBR-MWU is approximately 15 times faster on average when $n = 5$. As the size of the payoff matrix increases, the performance gap in convergence rate as measured by the number of steps also increases. Indeed, even for $n = 10$, OMWU requires more than $4.2M$ steps in half of the runs, while the respective number for FLBR-MWU is $16.3K$, implying that FLBR-MWU is 100 times faster than OMWU in the median sense. Larger game sizes make OMWU essentially impractical while FLBR-MWU is still able to converge in less than $5M$ steps.

Next, we demonstrate the properties of the FLBR-MWU algorithm using additional

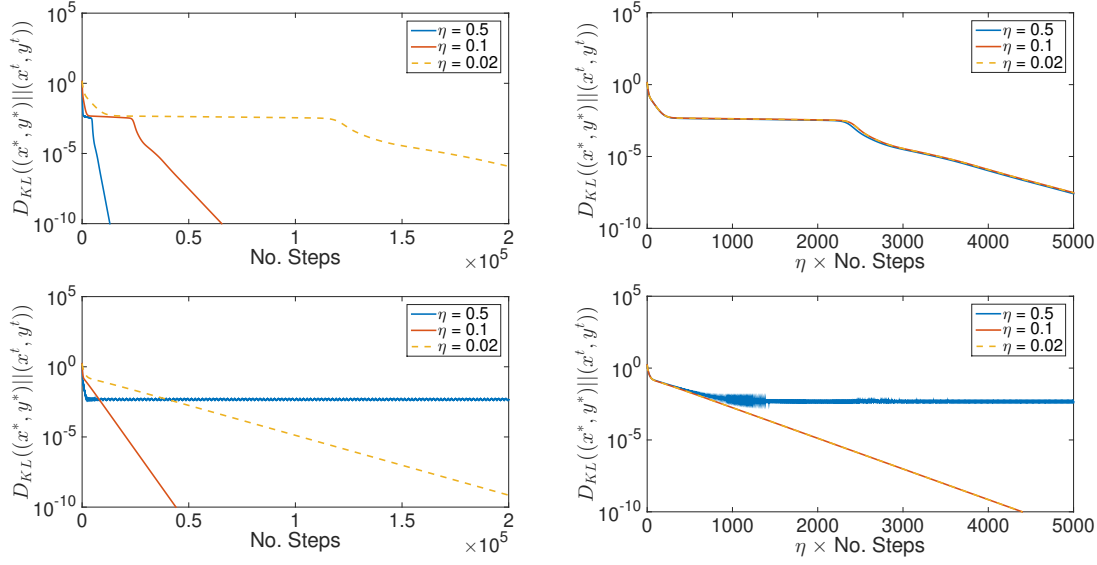


Figure 6.3: The D_{KL} between the Nash equilibrium and the FLBR-MWU dynamics for two instances and no rescaling of x-axis (upper panels) and with rescaling (lower panels). The relationship between the number of steps and learning rate is inversely proportional.

Table 6.2: Statistics on the number of steps till convergence for various sizes of the game. The maximum number of steps was set to $t_{\max} = 5 \times 10^6$.

Learning alg.	Statistic	$n = 5$	$n = 10$	$n = 50$
FLBR-MWU	Mean	33.7K	103.3K	984.9K
	Median	9.8K	16.3K	409.3K
	t_{\max} was hit	0.0%	0.0%	1.0%
OMWU	Mean	1088.8K	3323.2K	5000.0K
	Median	353.8K	4208.1K	5000.0K
	t_{\max} was hit	9.0%	46.0%	100.0%

metrics and perform further comparisons.

Convergence to the value of the game. Figure 6.4 shows the evolution of the current value of the game at each iteration, with the same payoff matrix as that used in the example of Figure 6.2. The current value of the game at iteration t is defined as $v^t = x^{tT} R y^t$, and it serves as another convergence measure to Nash equilibrium. MWU (blue) oscillates around the true value of the game ($v = 0.529677$) without converging, while OMWU (red) oscillates with decreasing amplitude and eventually it converges to the true value. The current game value for the FLBR-MWU dynamics (black) converges much faster requiring only a few thousand

steps.

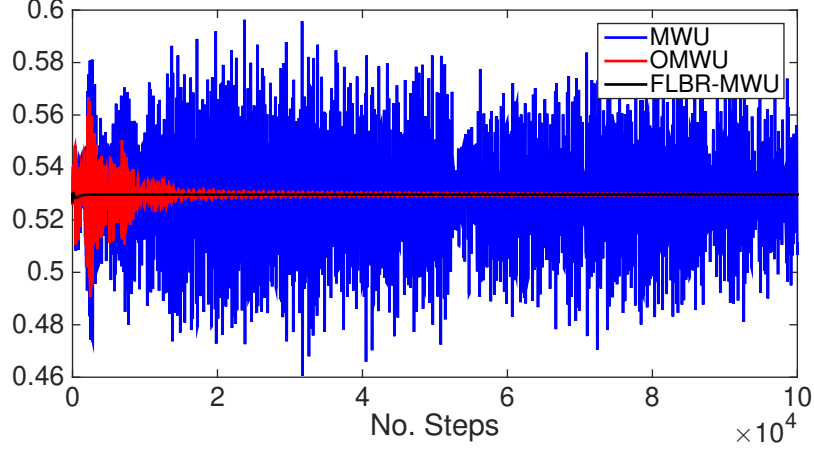


Figure 6.4: The value of the game as a function of the number of steps for the three MWU variants.

Dynamics trajectories. Figures 6.5 and 6.6 show the trajectories of the row player (i.e., x_i^t for $i = 1, \dots, 10$) in linear and log scale, respectively. Similarly, Figures 6.7 and 6.8 show the trajectories of the column player (i.e., y_i^t). Again, the payoff matrix is the same as in Figure 6.2, and the Nash equilibrium is estimated as:

$$\begin{pmatrix} x^{*T} \\ y^{*T} \end{pmatrix} = \begin{pmatrix} 0.126766, 0.276988, 0, 0.22506, 0.081435, 0, 0.191705, 0, 0.098045, 0 \\ 0, 0.058227, 0, 0.298188, 0.213176, 0, 0, 0.283403, 0.000376, 0.146628 \end{pmatrix}.$$

First, we note that for all pure strategies that do not belong to the support of x^* or y^* , the corresponding probabilities in x^t and y^t converge to 0 under FLBR-MWU, after a few thousand steps. Additionally, we observe interesting patterns during the evolution of the learning dynamics in both scales which are intimately connected with the KL divergence trajectory shown in Figure 6.2. Indeed, it is worth looking at y_9^t (log scale; Figure 6.8), which shows the most interesting pattern. Initially it seems that this is not a surviving strategy of the dynamics and its probability decreases for the first 10K steps. However, and, despite its very low value, it recovers to the actual Nash equilibrium value. Similarly, we observe that the non-zero elements of x^t (linear scale; Figure 6.5) are linearly evolving for several thousands of steps. Those changes in the dynamics correspond to the plateau of the KL divergence observed in Figure 6.2. Our explanation of the dynamics trajectories is as follows: starting from the uniform state, the FLBR-MWU algorithm first finds an approximate Nash equilibrium with a value close to the true value of the game but then escapes from it until it eventually converges to the actual Nash equilibrium.

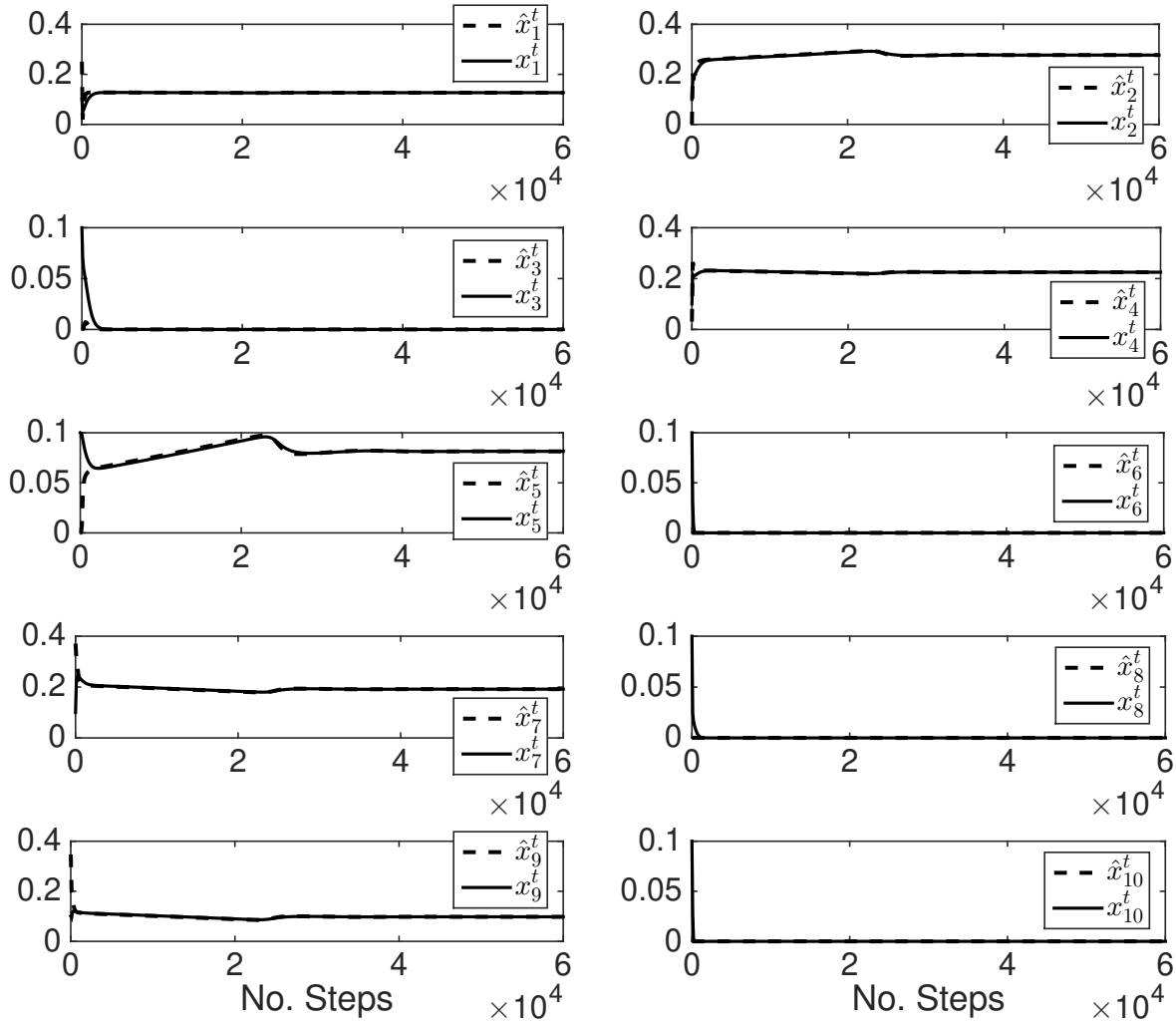


Figure 6.5: The dynamics of the update step per coordinate (solid), as well as the IBR step (dashed) for the row player. For the equilibrium strategy x^* , it holds that $\text{supp}(x^*) = \{1, 2, 4, 5, 7, 9\}$. Note that x^t converges to the same support.

Another interesting observation is that the dynamics of the IBR step (recall Equation (6.2)) drive the FLBR-MWU dynamics in the sense that when the IBR dynamics are above the FLBR-MWU dynamics, then the corresponding probabilities in the update step of FLBR-MWU increase, while the opposite is true when the IBR dynamics are below the FLBR-MWU

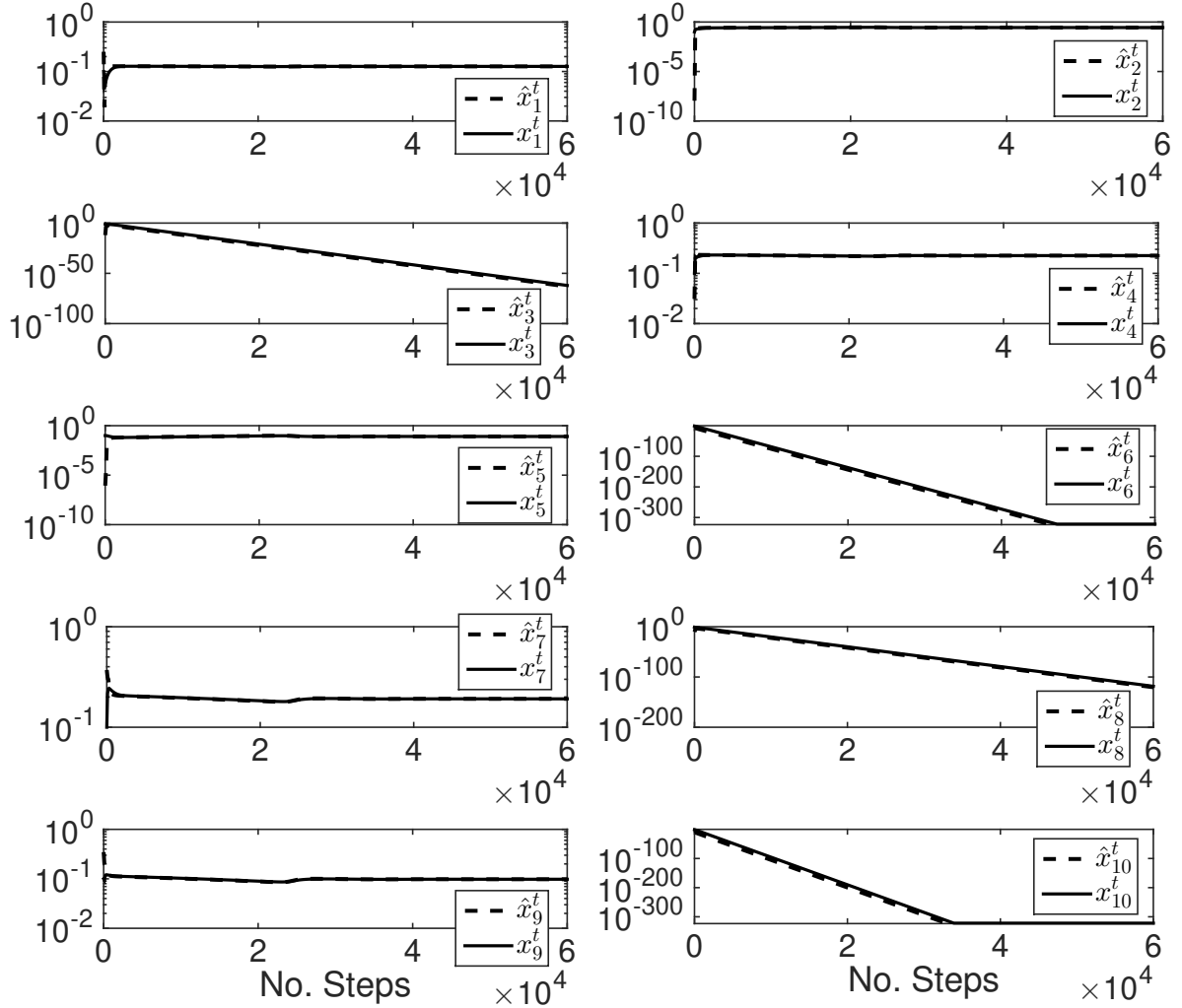


Figure 6.6: Same as Figure 6.5, but in logarithmic scale.

dynamics.

Effect of the intermediate rate (ξ). We present further statistical information on the effect of ξ . Figure 6.9 shows the distribution of the number of steps as a boxplot for $n = 10$ (left) and $n = 20$ (right). The red line in the boxplot corresponds to the median value while the blue box corresponds to the area covered by the 2nd and 3rd quantiles. The distribution of the number of steps till convergence is positively (or right) skewed. Therefore we also report the statistics of the right tail in Table 6.3. The presented results further validate the

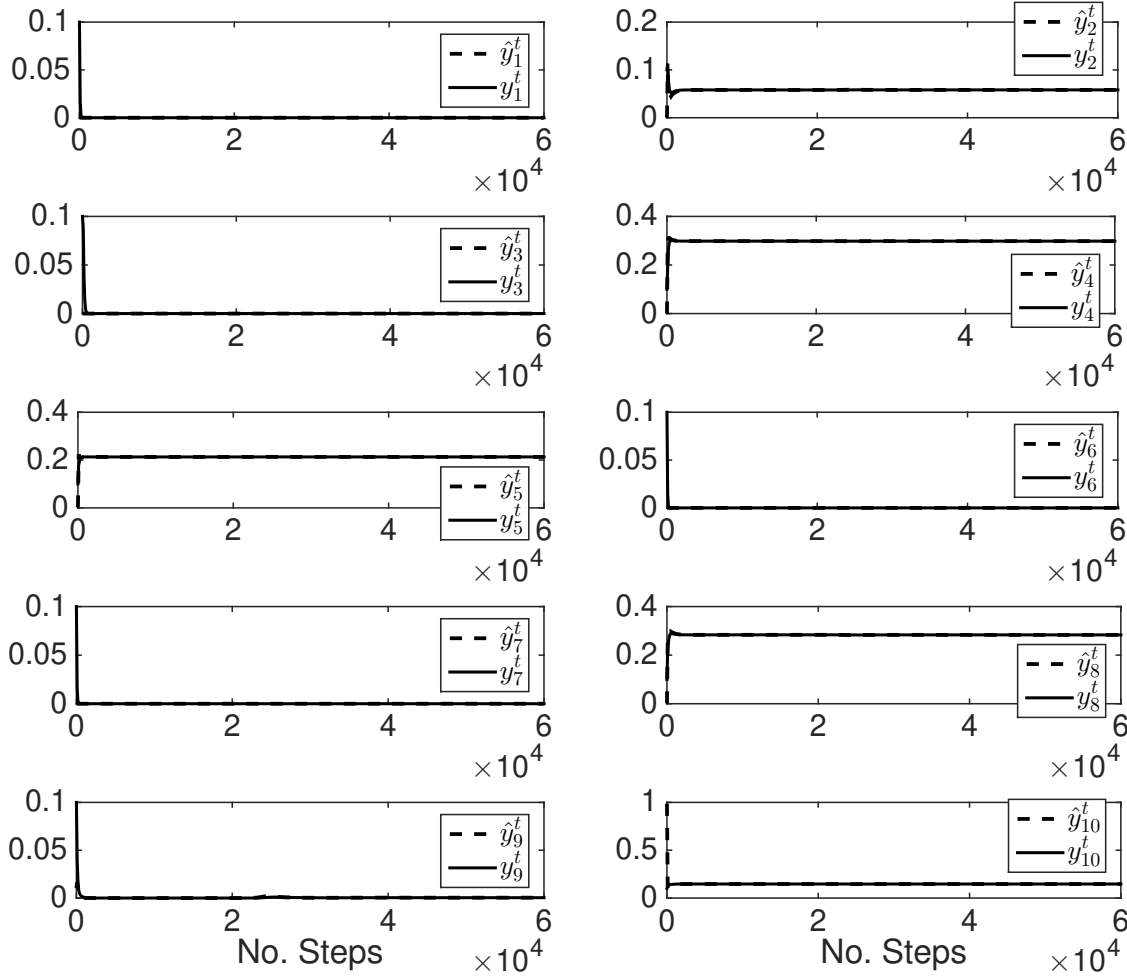


Figure 6.7: The dynamics of the update step per coordinate (solid), as well as the IBR step (dashed) for the column player. For the equilibrium strategy y^* , it holds that $\text{supp}(y^*) = \{2, 4, 5, 8, 9, 10\}$.

suggested value for ξ . We also remark that the product $\eta\xi$ is not always less than 1 in our experiments. Hence, although we needed the condition $\eta\xi < 1$ to prove our theoretical result in Section 6.2, the numerical evidence shows that the product can take values greater than 1 and still attain convergence (however $\eta\xi$ should not become arbitrarily large).

Number of steps. Moving on, we present additional comparisons between FLBR-MWU and OMWU. Figure 6.10 demonstrates the distribution of the number of steps till convergence

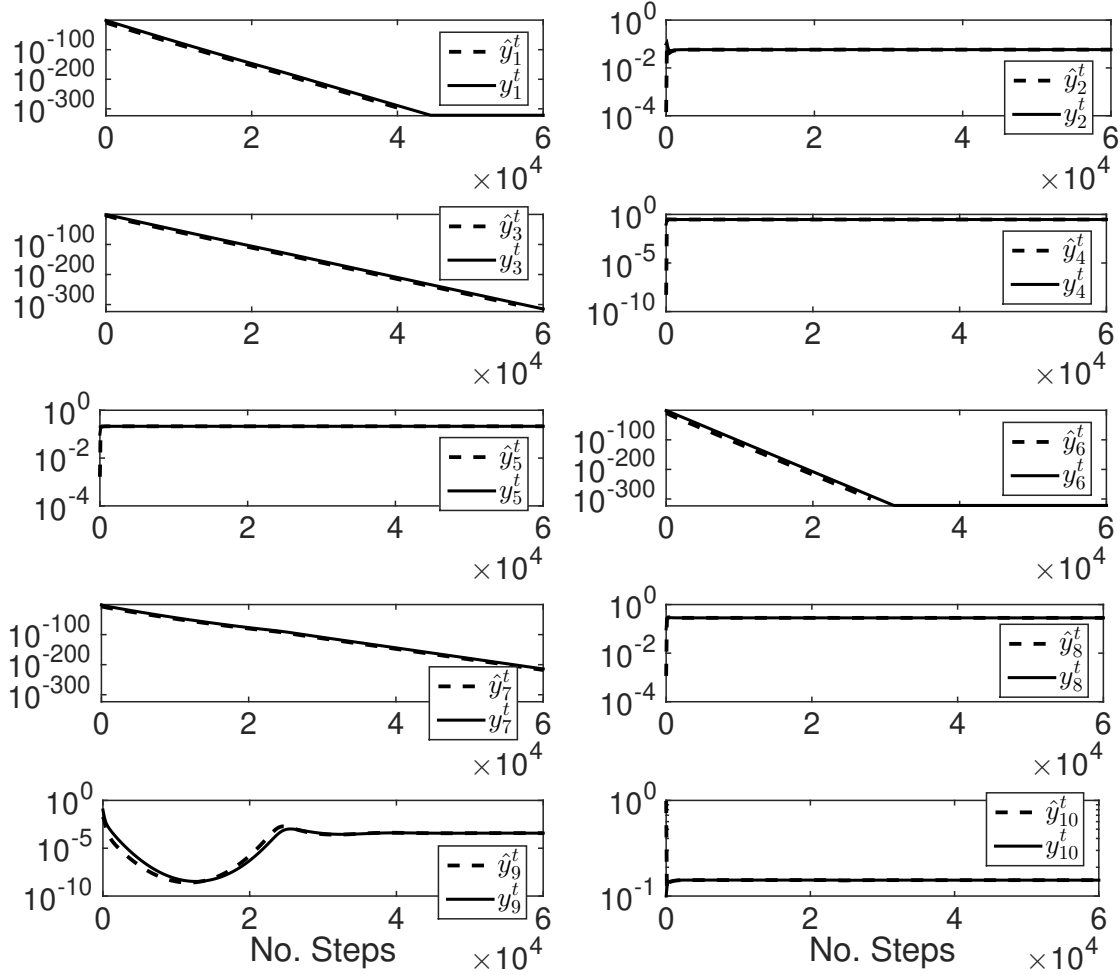


Figure 6.8: Same as Figure 6.7, but in logarithmic scale.

for FLBR-MWU (left) and OMWU (right). Interestingly, the distribution for payoff matrix size $n = 50$ with the FLBR-MWU algorithm is similar to the distribution for $n = 5$ with the OMWU algorithm. The computational gains are expected to be even more dramatic for larger games.

MWU, OMWU, and OMD. We also present a comparison among the MWU, OMWU and OMD dynamics (where for OMD we implemented the version of [Mertikopoulos et al., 2019] with entropy regularization). Figure 6.11 shows the evolution of a long run of 5 million steps and two values for the learning rate, η . We use the same payoff matrix as in Figure 6.2 and recall that the proposed FLBR-MWU method converged after only 100K steps (see

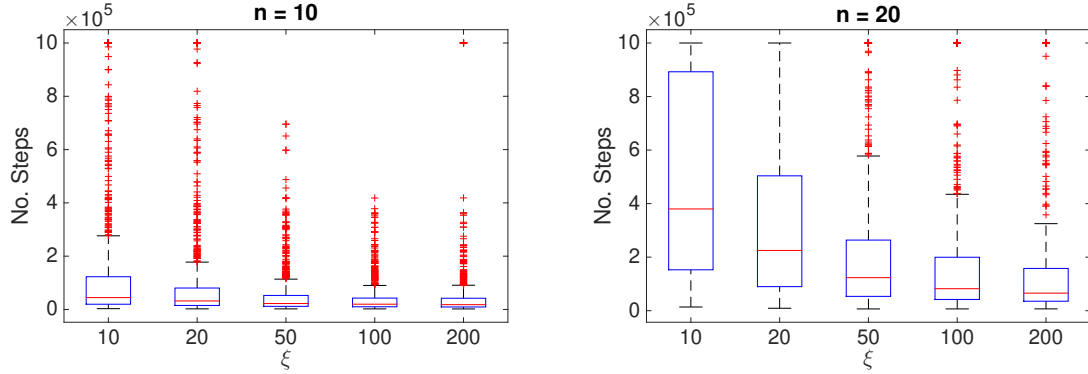


Figure 6.9: Boxplots for the number of steps until convergence for various values of ξ and two payoff matrix sizes.

Table 6.3: Quantile statistics on the number of steps till convergence for various values of ξ and n . The maximum number of steps was set to $t_{\max} = 2 \times 10^6$.

Matrix size	Quantile	$\xi = 10$	$\xi = 20$	$\xi = 50$	$\xi = 100$	$\xi = 200$
$n = 10$	75%	127.1K	83.3K	54.5K	44.1K	43.6K
	90%	346.7K	209.1K	137.9K	111.3K	110.3K
	97.5%	1035.2K	640.3K	322.1K	228.7K	372.8K
$n = 20$	75%	1957.2K	1127.2K	576.9K	441.3K	342.7K
	90%	2000.0K	2000.0K	1644.8K	1076.0K	830.8K
	97.5%	2000.0K	2000.0K	2000.0K	2000.0K	2000.0K

Figure 6.2). It is evident from the KL divergence in Figure 6.11 (leftmost panels) that the OMWU and OMD algorithms have almost the same behavior, as expected by [Wei et al., 2021], and they both converge, but in a very slow pace. The oscillatory behavior is prominent even after a large number of steps, as quantified by the l_1 norm difference (rightmost panels of Figure 6.11).

Finally, we report in Table 6.4 several convergence statistics between OWMU, OMD and FLBR-MWU algorithms with $\eta = 0.1$. This table is an extension of Table 6.2. Once again, the proposed FLBR-MWU algorithm is orders of magnitude faster while the closeness of the statistics between OWMU and OMD reveals the (almost) equivalence between the two algorithms.

As a concluding remark, the reduction of acceleration gain as n increases in our last experiment, it is artificial due to the use of a limit on the number of steps (denoted t_{\max} in the manuscript) which was set to $5M$ steps. Actually, if the allowed number of steps left

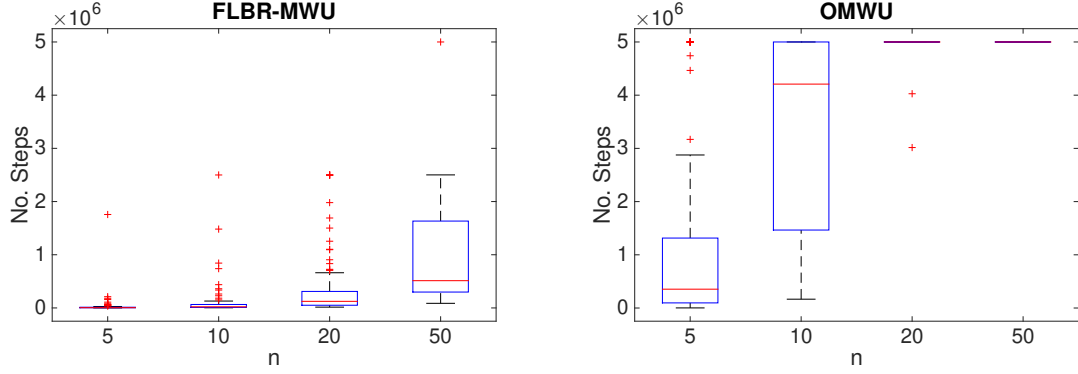


Figure 6.10: Boxplots for the number of steps until convergence for various payoff matrix sizes under FLBR-MWU (left) and OMWU/OMD (right). The computational gains when FLBR-MWU is used are striking.

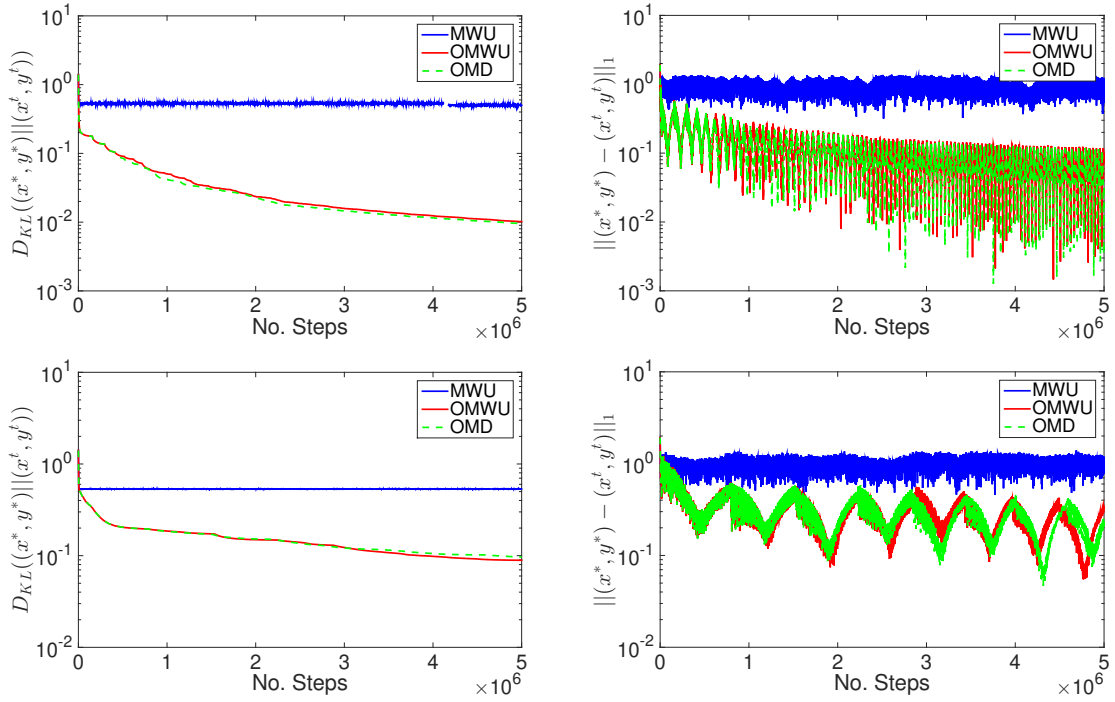


Figure 6.11: KL divergence and l_1 norm difference for $t_{\max} = 5 \times 10^6$ and two values for the learning rate: $\eta = 0.1$ (upper row of panels) and $\eta = 0.02$ (lower row of panels).

Table 6.4: Statistics on the number of steps till convergence for OWMU, OMD and FLBR-MWU and various payoff matrix sizes. The maximum number of steps was set to $t_{\max} = 5 \times 10^6$.

Matrix size	Statistic	$n = 5$	$n = 10$	$n = 20$	$n = 50$
OWMU	Mean	1287.3K	3280.9K	4997.8K	5000.0K
	Median	631.9K	3697.8K	5000.0K	5000.0K
	t_{\max}	12.0	44.0	98.0	100.0
OMD	Mean	1287.6K	3292.9K	4997.8K	5000.0K
	Median	631.9K	3629.1K	5000.0K	5000.0K
	t_{\max}	12.0	44.0	98.0	100.0
FLBR-MWU	Mean	18.8K	45.9K	267.1K	1130.8K
	Median	8.0K	21.4K	64.0K	701.3K
	t_{\max}	0.0	0.0	0.0	2.0

unlimited, then the relative computational gains of FLBR-MWU is even higher for larger n 's.

Chapter 7

Noisy Games

7.1 Introduction

In previous chapters are identified various different causes of misinformation, including deception and misleading reports, human errors, deliberate attempts by the game designer to channel players into different behaviours, erroneous sensor readings and random effects. Here we focus on a special case of misinformation, attributed to noise and signal errors, a situation often occurring in distributed multiagent systems. This class of misinformation games will be called *noisy games*.

Specifically, in distributed multiagent systems, agents are equipped with an internal logic that allows them to autonomously solve problems of a given nature. However, at deployment time, the precise specification of these problems is often unknown; instead, the details are communicated as needed at operation time, during the so-called “online phase” [Brown et al., 2017]. In such cases, unexpected communication errors, malfunctions in the communication module or noise may cause the agents to operate under a distorted problem specification, leading to unexpected behaviour.

For example, consider the scenario where we have two autonomous self-interested agents, already deployed in an unfriendly environment. At some point in time, the human controller asks each of the agents to choose among two actions, also specifying the payoffs for each combination of choices. If the communication goes through as expected, then the behaviour of the agents is predictable by the well-known results of game theory. However, if one (or both) of the agents’ communication module malfunctions, or if there is unexpected noise in the communication channel, the signal may arrive distorted. This could lead agents to receive an erroneous payoff matrix, essentially causing them to believe that they play a game different from the one communicated to them, with unpredictable results (Figure 1.2).

Note that, if, at deployment time, the designer had foreseen the possibility for the agents to receive an erroneous game specification, then the agents would have been programmed to treat all signals as uncertain (i.e., true under a certain probability). In this case, the possibility of error is integrated in the agents’ logic (even when no communication error occurs), and their behaviour can be modelled using the rich results on Bayesian games and games with

incomplete information [Zamir, 2009]. On the other hand, if such a scenario had not been foreseen at deployment time, then the agents will operate under the payoff matrices received, without considering the possibility that the payoff matrices are not the correct ones. This is quite different, as the agents' decisions will be totally misled by the erroneous setting, and will not consider mitigation measures “just in case” the specification that they received is wrong.

The aim of this chapter is to provide the theoretical machinery necessary to study scenarios of this kind. In particular, the main research question to be addressed is:

Given a game and a specific noise pattern affecting the players' perceived payoff matrices, compute the probability that players' behaviour (i.e., chosen strategies) will be as close as possible (in a manner to be formally defined later) to the behaviour that they would have in the absence of noise.

In summary, the main contributions of this chapter are the following:

1. Provide motivation for the need to define misinformation in the context of noisy games, by positioning our work with respect to other similar efforts in the literature, in particular related to games with uncertainty, and games where the players have some kind of misconception related to the game's payoffs (Section 2.6).
2. The definition of a formal model for the description of misinformation in noisy games (Section 7.2).
3. The computation of the probability that the players' behaviour is unaffected by random noise, a feature that we call *behavioural consistency* (Section 7.3).
4. Experimentation to visualize and validate our main results (Section 7.5).

7.2 Noisy Games

7.2.1 Basic Definitions

Noisy games are a special class of misinformation games, where misinformation is due to a random distortion in the original payoff matrix. Formally:

Definition 51 (Noisy game). *A noisy game is a canonical misinformation game $mG = \langle G^0, G^1, \dots, G^{|N|} \rangle$, where $G^i = G^0 + \Delta^i$ for some matrix Δ^i whose elements follow a certain probability distribution.*

Note that the restriction of a noisy game being canonical implies that noise only affects the payoff matrix. In a more general scenario, noise could also affect the number of players and/or the strategies that a player understands (knows) regarding a game. However, as shown

in [Varsos et al., 2021], we do not need to consider this case separately, and we can restrict ourselves to canonical games as above for simplicity.

In this chapter, we concentrate on noisy games whose actual game is a 2×2 bimatrix game, and where each element of Δ^i follows the normal distribution. We call such games *normal noisy games*. Therefore:

Definition 52 (Normal noisy game). *A normal noisy game is a tuple $mG = \langle G^0, G^r, G^c \rangle$, where:*

- G^0, G^r, G^c are 2×2 bimatrix games
- For $x \in \{r, c\}$, $G^x = G^0 + \Delta^x$, where Δ^x is a bimatrix whose elements follow the normal distribution (possibly for a different mean and normal deviation)

Notational conventions and shorthands To avoid confusion caused by the use of multiple indices in subsequent sections, we will use the notation $A[i, j]$ to refer to the element in the i^{th} row and j^{th} column of a matrix A , i.e., if $A = (a_{ij})$, then $A[i, j] = a_{ij}$.

We will use boldface to indicate tables whose elements are all equal to a certain value. For example $[b]_{n \times m}$ represents the $n \times m$ table B , such that $B[i, j] = b$ for all i, j . The $n \times m$ subscript will be omitted when obvious from the context.

For three tables A, M, D of the same dimensions, we write $A \sim \mathcal{N}(M, D)$ to indicate that $A[i, j] \sim \mathcal{N}(M[i, j], D[i, j])$ for all i, j .

We define operators on payoff matrices as follows. Consider a 2×2 bimatrix game $G = \langle N, S, P \rangle$, where $P = (P_r; P_c)$. Then:

- For 2×2 tables M_r, M_c, D_r, D_c , the expression $G \sim \mathcal{N}((M_r; M_c), (D_r; D_c))$ indicates that $P_r \sim \mathcal{N}(M_r, D_r), P_c \sim \mathcal{N}(M_c, D_c)$
- For a 2×2 bimatrix $A = (A_r; A_c)$ and $\lambda \in \mathbb{R}$, the result of the operation $\lambda G + A$ is the 2×2 bimatrix game $G' = \langle N', S', P' \rangle$, where $N' = N, S' = S, P' = \lambda P + A$

7.2.2 Strategies, strategy profiles and equilibria in misinformation games

Here we will use the main definitions of Section 4.2 (strategies, strategy profiles, equilibrium). For simplicity, we will only consider normal noisy games, although it is trivial to extend the definitions for arbitrary misinformation games (see [Varsos et al., 2021] for details). So let us fix some normal noisy game mG .

Notationally, as we consider normal noisy games with two players, we denote the *misinformed strategy profile* for the individual game, and is defined as a pair $\sigma = (\sigma_r, \sigma_c)$, where σ_x is a misinformed strategy of $x \in \{r, c\}$. Due to our assumption of mG being a canonical game, a misinformed strategy (and misinformed strategy profile) is also a strategy (strategy

profile) in G^0 . Due to this fact, we will simply use Σ_x to denote the misinformed strategies of player x in mG , and Σ to denote the misinformed strategy profiles of mG .

As normal noisy games are a subgroup of misinformation games, the decisions of a player are made based on his own payoff matrix (the one in G^x), payoffs are computed on the basis of the actual payoff matrix (the one in G^0). This is reflected in the definition of payoffs, note that this is equivalent with equation (4.1).

Indeed, set $P^0 = (P_r^0; P_c^0)$, $P^r = (P_r^r; P_c^r)$, $P^c = (P_r^c; P_c^c)$ the payoff matrices of G^0, G^r, G^c respectively. Then:

- The *actual payoff function* of player x , under a given strategy profile $\sigma = (\sigma_r, \sigma_c)$, $u_x : \Sigma \rightarrow \mathbb{R}$, is defined as:

$$u_x(\sigma_r, \sigma_c) = \sigma_r^T P_x^0 \sigma_c$$

- The *misinformed payoff function* of player x , under the viewpoint of player y and the strategy profile $\sigma = (\sigma_r, \sigma_c)$, $u_x^y : \Sigma \rightarrow \mathbb{R}$, is defined as:

$$u_x^y(\sigma_r, \sigma_c) = \sigma_r^T P_x^y \sigma_c$$

As in Chapter 4 the actual payoff function represents the payoff that player x will really receive as a response to his strategic choices. On the contrary, the misinformed payoff function represents the payoff of player x , under the (erroneous) view of the game that player y has. Note that the equilibria concepts derived using Definitions 16-17, as Price of Misinformation metric, see Definition 27.

Again, if $PoM = 1$, the players adopt optimal behaviour, due to misinformation. Moreover, interesting results can be derived by comparing the PoA of G^0 with the PoM of mG : if $PoM < PoA$, then misinformation has a beneficial effect on social welfare, as the players are inclined (due to their misinformation) to choose socially better strategies; on the other hand, if $PoM > PoA$, then misinformation leads to a worse outcome, from the perspective of social welfare.

7.2.3 Behavioural Consistency and ε -closeness

Given that the misinformed equilibria of a normal noisy game may be different than the Nash equilibria of the actual game, it makes sense to define a metric to quantify the distance among these equilibria and their respective strategies, essentially measuring the effect of noise on the behaviour of the players:

Definition 53 (ε -closeness). *Let $\sigma = (\sigma_1, \sigma_2)$, $\sigma' = (\sigma'_1, \sigma'_2)$ be two strategies and $\varepsilon \geq 0$. Then we say that σ, σ' are ε -close if and only if $\text{supp}(\sigma) = \text{supp}(\sigma')$ and $\max\{|\sigma_1 - \sigma'_1|, |\sigma_2 - \sigma'_2|\} \leq \varepsilon$. For a strategy σ , the set of strategies that are ε -close to it, is denoted by $Cl^\varepsilon(\sigma)$.*

The above definition applies on strategies in general, and, thus, allows us to apply it also to check ε -closeness among strategies and/or misinformed strategies, as long as they contain the same number of pure strategies. Moreover, although the definition applies for normal noisy games, it is trivial to generalize it for non-normal ones. Finally, note that ε -closeness requires identical supports. This requirement is based on the idea that adding (or removing) a pure strategy to (from) the support of a strategy is considered a major change in the player's behaviour.

We extend Definition 53 to (misinformed and non-misinformed) strategy profiles (and equilibria), in the obvious manner: $\sigma = (\sigma_r, \sigma_c)$ is ε -close to $\sigma' = (\sigma'_r, \sigma'_c)$ if and only if σ_r is ε -close to σ'_r and σ_c is ε -close to σ'_c . We denote by $Cl^\varepsilon(\sigma)$ the strategy profiles that are ε -close to σ . For a set of strategy profiles Σ^* , we set $Cl^\varepsilon(\Sigma^*) = \bigcup_{\sigma \in \Sigma^*} Cl^\varepsilon(\sigma)$, i.e., the strategy profiles that are ε -close to at least one of the strategy profiles in Σ^* .

The definition of ε -closeness gives formal substance to the idea of the behaviour of the players (expressed as an equilibrium) being “similar”: two equilibria that are ε -close are “similar” (and vice-versa), see Figure 7.1. This notion allows us to formally define the behavioural consistency of players in the presence of noise, which amounts to checking whether the equilibria of the noisy game are similar (i.e., ε -close) to the “expected” ones under the actual game. Formally:

Definition 54. *Consider a normal noisy game mG and some tolerance $\varepsilon \geq 0$. Then,*

- *mG is ε -misinformed iff for every natural misinformed equilibrium σ^* of mG , there is a Nash equilibrium σ^0 of G^0 , such that $\sigma^* \in Cl^\varepsilon(\sigma^0)$.*
- *mG is inverse- ε -misinformed iff for every Nash equilibrium σ^0 of G^0 , there is a natural misinformed equilibrium σ^* of mG , such that $\sigma^* \in Cl^\varepsilon(\sigma^0)$.*

The following example will be used as a running example for the rest of this Chapter to illustrate our results.

Example 7.1 (Running example). *We consider two autonomous robotic agents r, c , deployed in a remote environment. At some point in time, the human controller asks each of the agents to choose among two actions s_1, s_2 , also specifying the payoffs for each combination of choices, as shown in matrix P^0 below.*

$$P^0 = \begin{pmatrix} (3, 2) & (0, 0) \\ (0, 0) & (2, 3) \end{pmatrix}$$

The above payoff matrix corresponds to the well-known Battle of the Sexes (BoS) game [Osborne and Rubinstein, 1994], which has 3 Nash equilibria, namely $\sigma_1^0 = ((1, 0), (1, 0))$, $\sigma_2^0 = ((0, 1), (0, 1))$, $\sigma_3^0 = ((3/5, 2/5), (2/5, 3/5))$.

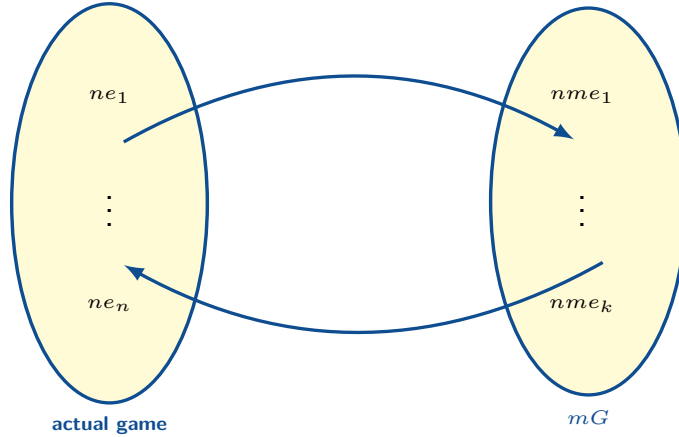


Figure 7.1: Behavioral consistency between interaction with real specifications and interaction with misinformation.

However, one of the components of the central communication module has received damage, unknowingly to the agents or the human controller, causing it to introduce a random noise ($\delta \sim \mathcal{N}(0, 1)$) to each of the values in P^0 during transmission. The above setting can be modelled as a normal noisy game $mG = \langle G^0, G^r, G^c \rangle \sim G^0 + \mathcal{N}(M_y^x, D_y^x)$, where $M_y^x = [\mathbf{0}]_{2 \times 2}$, $D_y^x = [\mathbf{1}]_{2 \times 2}$ for all $x, y \in \{r, c\}$, and the payoff matrix of G^0 is P^0 .

Our objectives are:

1. To compute the probability that the robotic agents will exhibit behavioural consistency (Definition 54), despite the noise caused by the malfunction
2. To determine whether minor inconsequential modifications in P^0 (e.g., multiplication of all its elements by a constant) would modify the above probability, and by how much

The above questions will be addressed and analysed in Sections 7.3 and 7.4 below. ■

7.3 Probabilities for Behavioural Consistency

In this section, we will compute the probabilities for a normal noisy game being (inverse-) ε -misinformed. For better readability, we split our analysis in 3 subsections. In Subsection 7.3.1, we recast some known results from game theory in a way that is more suitable for our analysis, whereas in Subsection 7.3.2, we develop some results that determine necessary and sufficient conditions for a misinformation game to be (inverse-) ε -misinformed. These results are then employed in Subsection 7.3.3 to compute the required probabilities. The respective results are summarized in Table 7.1 (for Subsection 7.3.1), Table 7.2 (for Subsection 7.3.2) and Tables 7.3, 7.4 and 7.5 (for Subsection 7.3.3).

7.3.1 Determining equilibrium strategies

For a 2×2 bimatrix game G , we denote by $\text{UGAIN}^G(x, i)$ the *utility gain* of strategy s_1 (compared to s_2) for player $x \in \{r, c\}$ when his/her opponent plays s_i , in game G . The reference to G will be omitted when obvious from the context. Note that $\text{UGAIN}(x, i)$ is determined by the elements of the payoff matrix of G (say $P = (P_r; P_c)$) as follows:

- For $x = r$, $\text{UGAIN}(r, i) = P_r[1, i] - P_r[2, i]$
- For $x = c$, $\text{UGAIN}(c, i) = P_c[i, 1] - P_c[i, 2]$

Intuitively, $\text{UGAIN}(x, i) > 0$ would mean that player x would play s_1 , if his/her opponent chose to play s_i , i.e., that s_1 is the best response (for x) to s_i . Similarly, $\text{UGAIN}(x, i) < 0$ would mean that player x would play s_2 , if his/her opponent chose to play s_i , i.e., that s_2 is the best response (for x) to s_i . Finally, when $\text{UGAIN}(x, i) = 0$, then player x is indifferent as to whether to play s_1 or s_2 , i.e., it has two pure best responses for his/her opponent's pure strategy s_i , indicating that the game is degenerate.

Example 7.1 (continued). We have $\text{UGAIN}(r, 1) = 3$, $\text{UGAIN}(r, 2) = -3$, $\text{UGAIN}(c, 1) = 2$, and $\text{UGAIN}(c, 2) = -2$. ■

Some well-known results from game theory for bimatrix games can be recast using the concept of $\text{UGAIN}(x, i)$. For example, the following proposition gives an equivalent formulation of the degeneracy criterion for 2×2 bimatrix games:

Proposition 20. *A 2×2 bimatrix game G is degenerate if and only if $\text{UGAIN}(x, i) = 0$ for some $x \in \{r, c\}, i \in \{1, 2\}$.*

Proof. Let's consider $G = \langle N, S, P \rangle$, for $P = (P_r; P_c)$. Suppose that G is degenerate. By definition, there is a pure strategy (say s_i , by player $x \in \{r, c\}$) that has two pure best responses. Suppose that $x = r, i = 1$. Then, since s_1, s_2 are equally preferred by c , it follows that $P_c[1, 1] = P_c[1, 2]$, i.e., $\text{UGAIN}(c, 1) = 0$. The other cases (i.e., when $x = c$ and/or $i = 2$) are analogous.

For the opposite, suppose that $\text{UGAIN}(r, 1) = 0$. Then $P_c[1, 1] = P_c[1, 2]$, so c has two pure best responses for the strategy s_1 of r , which means that G is degenerate. The proof is analogous for the other cases. □

When a non-degenerate 2×2 bimatrix game has a mixed Nash equilibrium, then its value is determined by $\text{UGAIN}(x, i)$:

Proposition 21. *Consider a non-degenerate 2×2 bimatrix game $G = \langle N, S, P \rangle$, for $P = (P_r; P_c)$. If $(p, 1 - p) \in \text{NE}_x(G)$ for some $0 < p < 1$, $x \in \{r, c\}$, then:*

$$p = \frac{\text{UGAIN}(\bar{x}, 2)}{\text{UGAIN}(\bar{x}, 2) - \text{UGAIN}(\bar{x}, 1)}$$

Proof. Suppose that $x = r$. From classical game theoretic results (e.g., see [Nisan et al., 2007a, Osborne and Rubinstein, 1994]), and our assumptions, we get that p will satisfy the following equation:

$$p \cdot P_c[1, 1] + (1 - p) \cdot P_c[2, 1] = p \cdot P_c[1, 2] + (1 - p) \cdot P_c[2, 2]$$

The result now follows trivially by solving this equation and applying the definition of $\text{UGAIN}(c, i)$.

Analogously, for the case where $x = c$, we get the following equation:

$$p \cdot P_r[1, 1] + (1 - p) \cdot P_r[1, 2] = p \cdot P_r[2, 1] + (1 - p) \cdot P_r[2, 2]$$

Solving it, as above, will give the required result. \square

Now consider a non-degenerate 2×2 bimatrix game G and some player $x \in \{r, c\}$. From classical results in game theory, we know that there are 4 possible cases for $NE_x(G)$, namely $NE_x(G) = \{(1, 0)\}$, $NE_x(G) = \{(0, 1)\}$, $NE_x(G) = \{(p, 1 - p)\}$ for some $0 < p < 1$ and $NE_x(G) = \{(1, 0), (0, 1), (p, 1 - p)\}$ for some $0 < p < 1$. If the game is degenerate, then there is one additional possibility, namely that $NE_x(G) = \{(p, 1 - p) \mid 0 \leq p \leq 1\} = \Sigma_x$.

For non-degenerate games, the value of $NE_x(G)$ can be determined using the following:

- $NE_x(G) = \{(1, 0)\}$ if and only if s_1 is dominant for x , or s_i is dominant for \bar{x} and s_1 is the best response for x on s_i .
- $NE_x(G) = \{(0, 1)\}$ if and only if s_2 is dominant for x , or s_i is dominant for \bar{x} and s_2 is the best response for x on s_i .
- $NE_x(G) = \{(p, 1 - p)\}$ for some $0 < p < 1$ if and only if no strategy is dominant for either player and no pure Nash equilibrium exists.
- $NE_x(G) = \{(1, 0), (0, 1), (p, 1 - p)\}$ for some $0 < p < 1$ if and only if no strategy is dominant for either player and two pure Nash equilibria exist.

The above conditions can also be expressed in terms of $\text{UGAIN}(x, i)$, as shown in Table 7.1. In the table, the various (mutually exclusive) cases are visualised for player r and for a non-degenerate game. The small figure in the rightmost column shows the depicted condition in terms of the relative order among the elements of P_r (blue lines) or P_c (yellow lines), which is determined by the sign (positive or negative) of $\text{UGAIN}(x, i)$. The first column provides a reference to the formulation of Proposition 22, where the above are formally stated and proved.

Before showing Proposition 22, for brevity, we introduce the following predicates to refer to the different cases with regards to the value of $NE_x(G)$:

- *Only-pure*: $OP_x^G(i)$, which is true if and only if the only equilibrium strategy for player x in game G is to play s_i , i.e.:
 $OP_x^G(1)$ if and only if $NE_x(G) = \{(1, 0)\}$
 $OP_x^G(2)$ if and only if $NE_x(G) = \{(0, 1)\}$
- *Only-mixed*: $OM_x^G(p)$, which is true if and only if the only equilibrium strategy for player x in game G is $(p, 1 - p)$ (where $0 < p < 1$), i.e.:
 $OM_x^G(p)$ if and only if $NE_x(G) = \{(p, 1 - p)\}$
- *Pure-and-mixed*: $PM_x^G(p)$, which is true if and only if player x has 3 equilibrium strategies in game G , two pure and one mixed, and the mixed one is $(p, 1 - p)$ (where $0 < p < 1$), i.e.:
 $PM_x^G(p)$ if and only if $NE_x(G) = \{(1, 0), (0, 1), (p, 1 - p)\}$
- *Ranged-only-mixed*: $ROM_x^G(\omega_1, \omega_2)$, which is true if and only if $OM_x^G(p)$ is true for some $\omega_1 < p < \omega_2$, i.e.:
 $ROM_x^G(\omega_1, \omega_2)$ if and only if $NE_x(G) = \{(p, 1 - p)\}$ for some p such that $\omega_1 < p < \omega_2$
- *Ranged-pure-and-mixed*: $RPM_x^G(\omega_1, \omega_2)$, which is true if and only if $PM_x^G(p)$ is true for some $\omega_1 < p < \omega_2$, i.e.:
 $RPM_x^G(\omega_1, \omega_2)$ if and only if $NE_x(G) = \{(1, 0), (0, 1), (p, 1 - p)\}$ for some p such that $\omega_1 < p < \omega_2$
- *Infinite-Nash*: IN_x^G , which is true if and only if player x has an infinite number of equilibrium strategies, namely the entire Σ_x (note that this is possible only for degenerate games), i.e.:
 IN_x^G if and only if $NE_x(G) = \Sigma_x$

When the game G is obvious from the context, we will omit the superscript G from the above. Now we can formally state Proposition 22, which formalizes the intuition of Table 7.1:

Proposition 22. *For any non-degenerate 2×2 bimatrix game the following hold:*

1. $OP_x(1)$ if and only if either one of the following is true:
 - (a) $(UGAIN(x, 1) > 0) \wedge (UGAIN(x, 2) > 0)$
 - (b) $(UGAIN(x, 1) > 0) \wedge (UGAIN(x, 2) < 0) \wedge (UGAIN(\bar{x}, 1) > 0) \wedge (UGAIN(\bar{x}, 2) > 0)$
 - (c) $(UGAIN(x, 1) < 0) \wedge (UGAIN(x, 2) > 0) \wedge (UGAIN(\bar{x}, 1) < 0) \wedge (UGAIN(\bar{x}, 2) < 0)$
2. $OP_x(2)$ if and only if either one of the following is true:
 - (a) $(UGAIN(x, 1) < 0) \wedge (UGAIN(x, 2) < 0)$
 - (b) $(UGAIN(x, 1) < 0) \wedge (UGAIN(x, 2) > 0) \wedge (UGAIN(\bar{x}, 1) > 0) \wedge (UGAIN(\bar{x}, 2) > 0)$

Case (from Proposition 22)	Value of $UGAIN(x, i)$, for:				Nash equilibrium strategies for $x = r$ ($NE_r(G)$)	Schematic depiction of the case
	$x = r$ $i = 1$	$x = r$ $i = 2$	$x = c$ $i = 1$	$x = c$ $i = 2$		
(1a)	>0	>0	$<>0$	$<>0$	$\{(1, 0)\}$	
(1b)	>0	<0	>0	>0	$\{(1, 0)\}$	
(1c)	<0	>0	<0	<0	$\{(1, 0)\}$	
(2a)	<0	<0	$<>0$	$<>0$	$\{(0, 1)\}$	
(2b)	<0	>0	>0	>0	$\{(0, 1)\}$	
(2c)	>0	<0	<0	<0	$\{(0, 1)\}$	
(3a)	>0	<0	<0	>0	$\{(p, 1 - p)\}$ $p = \frac{UGAIN(c, 2)}{UGAIN(c, 2) - UGAIN(c, 1)}$	
(3b)	<0	>0	>0	<0	$\{(p, 1 - p)\}$ $p = \frac{UGAIN(c, 2)}{UGAIN(c, 2) - UGAIN(c, 1)}$	
(4a)	>0	<0	>0	<0	$\{(1, 0), (0, 1), (p, 1 - p)\}$ $p = \frac{UGAIN(c, 2)}{UGAIN(c, 2) - UGAIN(c, 1)}$	
(4b)	<0	>0	<0	>0	$\{(1, 0), (0, 1), (p, 1 - p)\}$ $p = \frac{UGAIN(c, 2)}{UGAIN(c, 2) - UGAIN(c, 1)}$	

Table 7.1: Visualising the cases of Proposition 22, for $x = r$.

$$(c) \ (UGAIN(x, 1) > 0) \wedge (UGAIN(x, 2) < 0) \wedge (UGAIN(\bar{x}, 1) < 0) \wedge (UGAIN(\bar{x}, 2) < 0)$$

3. $OM_x(p)$ if and only if $p = \frac{UGAIN(\bar{x}, 2)}{UGAIN(\bar{x}, 2) - UGAIN(\bar{x}, 1)}$ and either one of the following is true:

$$(a) \ (UGAIN(x, 1) > 0) \wedge (UGAIN(x, 2) < 0) \wedge (UGAIN(\bar{x}, 1) < 0) \wedge (UGAIN(\bar{x}, 2) > 0)$$

$$(b) \ (UGAIN(x, 1) < 0) \wedge (UGAIN(x, 2) > 0) \wedge (UGAIN(\bar{x}, 1) > 0) \wedge (UGAIN(\bar{x}, 2) < 0)$$

4. $PM_x(p)$ if and only if $p = \frac{UGAIN(\bar{x}, 2)}{UGAIN(\bar{x}, 2) - UGAIN(\bar{x}, 1)}$ and either one of the following is true:

$$(a) \ (UGAIN(x, 1) > 0) \wedge (UGAIN(x, 2) < 0) \wedge (UGAIN(\bar{x}, 1) > 0) \wedge (UGAIN(\bar{x}, 2) < 0)$$

$$(b) \ (UGAIN(x, 1) < 0) \wedge (UGAIN(x, 2) > 0) \wedge (UGAIN(\bar{x}, 1) < 0) \wedge (UGAIN(\bar{x}, 2) > 0)$$

Proof. By Proposition 20, we conclude that $UGAIN(x, i) \neq 0$ for all $x \in \{r, c\}, i \in \{1, 2\}$. This means that the different (mutually exclusive) cases of the formulation of the proposition cover all possible cases for a non-degenerate game (see also Table 7.1). Thus, it suffices to show the “only if” part for each different case.

For (1a), note that player x will play $(1, 0)$ (i.e., s_1) regardless of the choice of \bar{x} , so $NE_x(G) = \{(1, 0)\}$ and $OP_x(1)$ is true.

For (1b), note that the only Nash equilibrium of G is $((1, 0), (1, 0))$, which proves the result. Next, (1c) is analogous to (1b).

The cases (2a), (2b), (2c) are analogous to (1a), (1b), (1c) respectively.

With regards to (3a), it can be easily shown that the game can have no pure Nash equilibrium. Thus, it must have a mixed one (by the result of Nash [Nash, 1951]). Moreover, it cannot have more than one mixed, as this would render it degenerate* (see [Nisan et al., 2007a], [Avis et al., 2010], [Osborne and Rubinstein, 1994]).

Thus, $NE_x(G) = \{(p, 1 - p)\}$, for some $0 < p < 1$. By Proposition 21, it follows that $p = \frac{UGAIN(\bar{x}, 2)}{UGAIN(\bar{x}, 2) - UGAIN(\bar{x}, 1)}$, which shows the result.

The case (3b) is analogous.

For (4a), we observe that the values of $UGAIN(x, i)$ imply that the game has exactly two pure Nash equilibria, namely: $((1, 0), (1, 0))$ and $((0, 1), (0, 1))$. By [Nisan et al., 2007a], [Osborne and Rubinstein, 1994], it must also have one (unique) mixed equilibrium.

Thus, $NE_x(G) = \{(1, 0), (0, 1), (p, 1 - p)\}$ for some $0 < p < 1$. Again, using Proposition 21, it follows that $p = \frac{UGAIN(\bar{x}, 2)}{UGAIN(\bar{x}, 2) - UGAIN(\bar{x}, 1)}$, which shows the result.

For (4b) the proof is analogous, except that here the pure Nash equilibria of G are: $((1, 0), (0, 1))$ and $((0, 1), (1, 0))$. \square

An analogous set of conditions determines whether the “ranged” versions of the above predicates are true:

Corollary 3. *Given a non-degenerate 2×2 bimatrix game G , the following hold:*

*Immediate consequence of Corollary 3.7 [Nisan et al., 2007a].

1. $ROM_x(\omega_1, \omega_2)$ if and only if $\omega_1 < \frac{UGAIN(\bar{x}, 2)}{UGAIN(\bar{x}, 2) - UGAIN(\bar{x}, 1)} < \omega_2$ and either one of the following is true:

$$(a) (UGAIN(x, 1) > 0) \wedge (UGAIN(x, 2) < 0) \wedge (UGAIN(\bar{x}, 1) < 0) \wedge (UGAIN(\bar{x}, 2) > 0)$$

$$(b) (UGAIN(x, 1) < 0) \wedge (UGAIN(x, 2) > 0) \wedge (UGAIN(\bar{x}, 1) > 0) \wedge (UGAIN(\bar{x}, 2) < 0)$$

2. $RPM_x(\omega_1, \omega_2)$ if and only if $\omega_1 < \frac{UGAIN(\bar{x}, 2)}{UGAIN(\bar{x}, 2) - UGAIN(\bar{x}, 1)} < \omega_2$ and either one of the following is true:

$$(a) (UGAIN(x, 1) > 0) \wedge (UGAIN(x, 2) < 0) \wedge (UGAIN(\bar{x}, 1) > 0) \wedge (UGAIN(\bar{x}, 2) < 0)$$

$$(b) (UGAIN(x, 1) < 0) \wedge (UGAIN(x, 2) > 0) \wedge (UGAIN(\bar{x}, 1) < 0) \wedge (UGAIN(\bar{x}, 2) > 0)$$

Example 7.1 (continued). We have that in Proposition 22 holds the case:

$$(UGAIN(x, 1) > 0) \wedge (UGAIN(x, 2) < 0) \wedge (UGAIN(\bar{x}, 1) > 0) \wedge (UGAIN(\bar{x}, 2) < 0)$$

for all $\{x, \bar{x}\} \in \{r, c\}$, thus this game has both pure and mixed Nash equilibria. Further, according to Corollary 3, if exist ω_1 and ω_2 then this game has Ranged-pure-and-mixed equilibria. \square

7.3.2 Misinformation games

In this subsection, we provide necessary and sufficient conditions for a misinformation game to be (inverse-) ε -misinformed. These are given in Propositions 23, 24, and use the notation previously introduced. Note that the propositions apply for all canonical misinformation games, not just noisy games. The results of this subsection are summarized in Table 7.2.

Proposition 23. Consider a canonical misinformation game $mG = \langle G^0, G^r, G^c \rangle$, where G^0 is a 2×2 bimatrix game and G^r, G^c are non-degenerate. Then, mG is ε -misinformed if and only if, for all $x \in \{r, c\}$, one of the following is true:

1. $OP_x^{G^0}(i)$ and $OP_x^{G^x}(i)$ for some $i \in \{1, 2\}$
2. $OM_x^{G^0}(p^0)$ for some $0 < p^0 < 1$ and $ROM_x^{G^x}(\omega_1, \omega_2)$, where $\omega_1 = \max\{0, p^0 - \varepsilon\}$, $\omega_2 = \min\{1, p^0 + \varepsilon\}$
3. $PM_x^{G^0}(p^0)$ for some $0 < p^0 < 1$ and $OP_x^{G^x}(1) \vee OP_x^{G^x}(2) \vee ROM_x^{G^x}(\omega_1, \omega_2) \vee RPM_x^{G^x}(\omega_1, \omega_2)$, where $\omega_1 = \max\{0, p^0 - \varepsilon\}$, $\omega_2 = \min\{1, p^0 + \varepsilon\}$
4. $IN_x^{G^0}$

Proof. By definition, mG is ε -misinformed if and only if for all $\sigma^* = (\sigma_r^*, \sigma_c^*) \in NME(mG)$ there exists $\sigma^0 = (\sigma_r^0, \sigma_c^0) \in NE(G^0)$ such that σ^*, σ^0 are ε -close. More formally:

$mG : \varepsilon$ misinformed

$$\begin{aligned}
&\Leftrightarrow \forall \sigma^* = (\sigma_r^*, \sigma_c^*) \in NME(mG) \exists \sigma^0 = (\sigma_r^0, \sigma_c^0) \in NE(G^0) : \sigma^* \in Cl^\varepsilon(\sigma^0) \\
&\Leftrightarrow \forall \sigma^* = (\sigma_r^*, \sigma_c^*) \in NME(mG), \sigma^* \in Cl^\varepsilon(NE(G^0)) \\
&\Leftrightarrow \forall \sigma^* = (\sigma_r^*, \sigma_c^*) \in NME(mG) \left(\sigma_r^* \in Cl^\varepsilon(NE_r(G^0)) \wedge \sigma_c^* \in Cl^\varepsilon(NE_c(G^0)) \right) \\
&\Leftrightarrow \forall \sigma_r^* \in NE_r(G^r), \sigma_r^* \in Cl^\varepsilon(NE_r(G^0)) \wedge \forall \sigma_c^* \in NE_c(G^c), \sigma_c^* \in Cl^\varepsilon(NE_c(G^0)) \\
&\Leftrightarrow \forall x \in \{r, c\} \forall \sigma_x^* \in NE_x(G^x), \sigma_x^* \in Cl^\varepsilon(NE_x(G^0))
\end{aligned}$$

Now let us fix some x and consider the different cases with regards to $NE_x(G^0)$:

- If $NE_x(G^0)$ contains a single pure strategy, i.e., $OP_x^{G^0}(i)$ is true for some $i \in \{1, 2\}$, then the expression $\forall \sigma_x^* \in NE_x(G^x), \sigma_x^* \in Cl^\varepsilon(NE_x(G^0))$ is true if and only if $NE_x(G^x)$ contains the same pure strategy, and no other, i.e., if and only if $OP_x^{G^x}(i)$ is true.
- If $NE_x(G^0)$ contains a single mixed strategy, i.e., $OM_x^{G^0}(p^0)$ is true for some $0 < p^0 < 1$, then the expression $\forall \sigma_x^* \in NE_x(G^x), \sigma_x^* \in Cl^\varepsilon(NE_x(G^0))$ is true if and only if $NE_x(G^x)$ contains a single mixed strategy that is ε -close to $(p^0, 1 - p^0)$, i.e., $ROM_x^{G^x}(\omega_1, \omega_2)$ is true, where $\omega_1 = \max\{0, p^0 - \varepsilon\}$, $\omega_2 = \min\{1, p^0 + \varepsilon\}$. Note that the max, min are necessary to cater for the case where $p^0 - \varepsilon$, $p^0 + \varepsilon$ are smaller than 0 or greater than 1, respectively.
- If $NE_x(G^0)$ contains two pure and one mixed strategies, i.e., $PM_x^{G^0}(p^0)$ is true for some $0 < p^0 < 1$, then the expression $\forall \sigma_x^* \in NE_x(G^x), \sigma_x^* \in Cl^\varepsilon(NE_x(G^0))$ is true if and only if $NE_x(G^x)$ contains either a pure or a mixed strategy that is ε -close to $(p^0, 1 - p^0)$. This is expressed by the expression in bullet #3 of the proposition.
- If $NE_x(G^0) = \Sigma_x$, i.e., $IN_x^{G^0}$ is true, then, no matter the contents of $NE_x(G^x)$, the expression $\forall \sigma_x^* \in NE_x(G^x), \sigma_x^* \in Cl^\varepsilon(NE_x(G^0))$ is true.

This, combined with the fact that these are the only cases with regards to the value of $NE_x(G^0)$, conclude the proof. \square

Proposition 24. Consider a canonical misinformation game $mG = \langle G^0, G^r, G^c \rangle$, where G^0 is a 2×2 bimatrix game and G^r, G^c are non-degenerate. Then, mG is inverse- ε -misinformed if and only if, for all $x \in \{r, c\}$, one of the following is true:

1. $OP_x^{G^0}(i)$ and $OP_x^{G^x}(i) \vee RPM_x^{G^x}(0, 1)$ for some $i \in \{1, 2\}$
2. $OM_x^{G^0}(p^0)$ for some $0 < p^0 < 1$ and $ROM_x^{G^x}(\omega_1, \omega_2) \vee RPM_x^{G^x}(\omega_1, \omega_2)$, $\omega_1 = \max\{0, p^0 - \varepsilon\}$, $\omega_2 = \min\{1, p^0 + \varepsilon\}$

3. $PM_x^{G^0}(p^0)$ for some $0 < p^0 < 1$ and $RPM_x^{G^x}(\omega_1, \omega_2)$, where $\omega_1 = \max\{0, p^0 - \varepsilon\}$, $\omega_2 = \min\{1, p^0 + \varepsilon\}$

4. $IN_x^{G^0}$ and $\varepsilon > 0.5$ and $RPM_x^{G^x}(\omega'_1, \omega'_2)$, where $\omega'_1 = \max\{0, 1 - \varepsilon\}$, $\omega'_2 = \min\{1, \varepsilon\}$

Proof. By definition, mG is inverse- ε -misinformed if and only if for all $\sigma^0 = (\sigma_r^0, \sigma_c^0) \in NE(G^0)$ there exists $\sigma^* = (\sigma_r^*, \sigma_c^*) \in NME(mG)$ such that σ^*, σ^0 are ε -close. More formally:

mG : inverse- ε -misinformed

$$\begin{aligned} &\Leftrightarrow \forall \sigma^0 = (\sigma_r^0, \sigma_c^0) \in NE(G^0) \exists \sigma^* = (\sigma_r^*, \sigma_c^*) \in NME(mG) : \sigma^* \in Cl^\varepsilon(\sigma^0) \\ &\Leftrightarrow \left(\forall \sigma_r^0 \in NE_r(G^0) \exists \sigma_r^* \in NE_r(G^x) : \sigma_r^* \in Cl^\varepsilon(NE_r(\sigma_r^0)) \right) \\ &\quad \wedge \left(\forall \sigma_c^0 \in NE_c(G^0) \exists \sigma_c^* \in NE_c(G^c) : \sigma_c^* \in Cl^\varepsilon(NE_c(\sigma_c^0)) \right) \\ &\Leftrightarrow \forall x \in \{r, c\} \forall \sigma_x^0 \in NE_x(G^0) \exists \sigma_x^* \in NE_x(G^x) : \sigma_x^* \in Cl^\varepsilon(NE_x(\sigma_x^0)) \end{aligned}$$

Now let us fix some x and consider the different cases with regards to $NE_x(G^0)$:

- If $NE_x(G^0)$ contains a single pure strategy, i.e., $OP_x^{G^0}(i)$ is true for some $i \in \{1, 2\}$, then the expression $\forall \sigma_x^0 \in NE_x(G^0) \exists \sigma_x^* \in NE_x(G^x) : \sigma_x^* \in Cl^\varepsilon(NE_x(\sigma_x^0))$ is true if and only if $NE_x(G^x)$ contains the same pure strategy, possibly in addition to others, i.e., (given that G^x is non-degenerate) if and only if $OP_x^{G^x}(i) \vee RPM_x^{G^x}(0, 1)$ is true.
- If $NE_x(G^0)$ contains a single mixed strategy, i.e., $OM_x^{G^0}(p^0)$ is true for some $0 < p^0 < 1$, then the expression $\forall \sigma_x^0 \in NE_x(G^0) \exists \sigma_x^* \in NE_x(G^x) : \sigma_x^* \in Cl^\varepsilon(NE_x(\sigma_x^0))$ is true if and only if $NE_x(G^x)$ contains a mixed strategy that is ε -close to $(p^0, 1 - p^0)$, possibly in addition to others, i.e., (given that G^x is non-degenerate) $ROM_x^{G^x}(\omega_1, \omega_2)$ is true, where $\omega_1 = \max\{0, p^0 - \varepsilon\}$, $\omega_2 = \min\{1, p^0 + \varepsilon\}$. Note that the max, min are necessary to cater for the case where $p^0 - \varepsilon, p^0 + \varepsilon$ are smaller than 0 or greater than 1, respectively.
- If $NE_x(G^0)$ contains two pure and one mixed strategies, i.e., $PM_x^{G^0}(p^0)$ is true for some $0 < p^0 < 1$, then the expression $\forall \sigma_x^0 \in NE_x(G^0) \exists \sigma_x^* \in NE_x(G^x) : \sigma_x^* \in Cl^\varepsilon(NE_x(\sigma_x^0))$ is true if and only if $NE_x(G^x)$ contains two pure and a mixed strategy that is ε -close to $(p^0, 1 - p^0)$, i.e., (given that G^x is non-degenerate) $RPM_x^{G^x}(\omega_1, \omega_2)$ is true, where $\omega_1 = \max\{0, p^0 - \varepsilon\}$, $\omega_2 = \min\{1, p^0 + \varepsilon\}$.
- If $NE_x(G^0) = \Sigma_x$, i.e., $IN_x^{G^0}$ is true, then, $\forall \sigma_x^0 \in NE_x(G^0) \exists \sigma_x^* \in NE_x(G^x) : \sigma_x^* \in Cl^\varepsilon(NE_x(\sigma_x^0))$ is true if and only if at least one of the strategies in $NE_x(G^x)$ is ε -close to each strategy in $NE_x(G^0)$. Given that $NE_x(G^x)$ is finite (because G^x is non-degenerate), this can only hold if $PM_x^{G^x}(p^x)$ for some p^x such that $(p, 1 - p) \in Cl^\varepsilon((p, 1 - p))$ for all $0 < p < 1$. From the latter, we conclude that $\varepsilon \geq 0.5$ and $\max\{0, p^0 - \varepsilon\} < p^x < \min\{1, p^0 + \varepsilon\}$, which leads to the requirement in bullet #4 of the proposition.

This, combined with the fact that these are the only cases with regards to the value of $NE_x(G^0)$, conclude the proof. \square

Condition on G^0	Condition on G^x	
	For ε -misinformed	For inverse- ε -misinformed
$OP_x^{G^0}(i)$	$OP_x^{G^x}(i)$	$OP_x^{G^x}(i) \vee$ $RPM_x^{G^x}(0, 1)$
$OM_x^{G^0}(p^0)$	$ROM_x^{G^x}(\omega_1, \omega_2)$	$ROM_x^{G^x}(\omega_1, \omega_2) \vee$ $RPM_x^{G^x}(\omega_1, \omega_2)$
$PM_x^{G^0}(p^0)$	$RPM_x^{G^x}(\omega_1, \omega_2) \vee$ $ROM_x^{G^x}(\omega_1, \omega_2) \vee$ $OP_x^{G^x}(1) \vee$ $OP_x^{G^x}(2)$	$RPM_x^{G^x}(\omega_1, \omega_2)$
$IN_x^{G^0}$	Always true	If $\varepsilon < 0.5$: always false If $\varepsilon \geq 0.5$: $RPM_x^{G^x}(\omega'_1, \omega'_2)$

In all the above:

$$0 < p^0 < 1,$$

$$\omega_1 = \max\{0, p^0 - \varepsilon\}, \omega_2 = \min\{1, p^0 + \varepsilon\},$$

$$\omega'_1 = \max\{0, 1 - \varepsilon\}, \omega'_2 = \min\{1, \varepsilon\}$$

Table 7.2: Scenarios for ε -misinformed and inverse- ε -misinformed

7.3.3 Probabilities

We will now exploit the results of the previous subsections, in order to compute the probabilities associated to various events, eventually leading up to the computation that a given normal noisy game mG is (inverse-) ε -misinformed. The results are summarized in Table 7.5, whereas intermediate results necessary to compute the above probabilities appear in Tables 7.3 and 7.4.

For a normal noisy game $mG \sim G^0 + \mathcal{N}(M, D)$, we define the family of random variables $U(y, x, i)$, such that, for any $x, y \in \{r, c\}, i \in \{1, 2\}$:

$$U(y, x, i) = \text{UGAIN}^{G^y}(x, i)$$

Applying formula (B.3) from Subsection B.1, we observe that $U(y, x, i) \sim \mathcal{N}(\mu_{U(y, x, i)}, d_{U(y, x, i)})$ for $\mu_{U(y, x, i)}, d_{U(y, x, i)}$ as shown in Table 7.3. The cdf and pdf of $U(y, x, i)$ (as resulting from formula (B.2) in Subsection B.1), as well as the probabilities for $U(y, x, i)$ taking certain values are also shown in the same table.

Results related to $U(y, x, i)$
$\mu_{U(y,r,i)} = (P_r^0[1, i] + M_r^y[1, i]) - (P_r^0[2, i] + M_r^y[2, i])$ $d_{U(y,r,i)} = \sqrt{(D_r^y[1, i])^2 + (D_r^y[2, i])^2}$ $\mu_{U(y,c,i)} = (P_c^0[i, 1] + M_c^y[i, 1]) - (P_c^0[i, 2] + M_c^y[i, 2])$ $d_{U(y,c,i)} = \sqrt{(D_c^y[i, 1])^2 + (D_c^y[i, 2])^2}$ $f_{U(y,x,i)}(u) = \frac{1}{d_{U(y,x,i)}} \phi\left(\frac{u - \mu_{U(y,x,i)}}{d_{U(y,x,i)}}\right)$ $F_{U(y,x,i)}(u) = \Phi\left(\frac{u - \mu_{U(y,x,i)}}{d_{U(y,x,i)}}\right)$ $\mathcal{P}[U(y, x, i) < 0] = F_{U(y,x,i)}(0)$ $\mathcal{P}[U(y, x, i) > 0] = 1 - F_{U(y,x,i)}(0)$ $\mathcal{P}[U(y, x, i) = 0] = 0$

Table 7.3: Formulas related to $U(y, x, i)$ for a given $mG \sim G^0 + \mathcal{N}(M, D)$.

Propositions 20 and 22 can now be employed to determine the probability that $NE_x(G^x)$ takes a certain value, based on the probabilities that $U(y, x, i)$ take certain values. More precisely, Lemma 11 is the counterpart of Proposition 20:

Lemma 11. *In any normal noisy game $mG = \langle G^0, G^r, G^c \rangle$, the probability that G^x is degenerate (for $x \in \{r, c\}$) is 0.*

Proof. The result is direct from Proposition 20 and the fact that $\mathcal{P}[U(x, y, i) = 0] = 0$ for any $x, y \in \{r, c\}$, $i \in \{1, 2\}$. \square

To formulate the counterpart of Proposition 22, the following lemma will prove helpful:

Lemma 12. *Consider two independent random variables $X \sim \mathcal{N}(\mu_X, d_X)$, $Y \sim \mathcal{N}(\mu_Y, d_Y)$, with pdfs f_X, f_Y respectively, and some $\Omega_1, \Omega_2 \in \mathbb{R} \cup \{-\infty\}$ such that $-\infty \leq \Omega_1 < \Omega_2 \leq 0$. Then:*

$$\mathcal{P}\left[\Omega_1 \leq \frac{X}{Y} \leq \Omega_2, X < 0, Y > 0\right] = \int_0^{+\infty} \left(\int_{\Omega_1 y}^{\Omega_2 y} f_X(x) dx \right) \frac{f_Y(y)}{y} dy$$

$$\mathcal{P}\left[\Omega_1 \leq \frac{X}{Y} \leq \Omega_2, X > 0, Y < 0\right] = \int_{-\infty}^0 \left(\int_{\Omega_1 y}^{\Omega_2 y} f_X(x) dx \right) \frac{f_Y(y)}{y} dy$$

Proof. For the first result, we observe that, since $\Omega_1 < \Omega_2 \leq 0$:

$$\Omega_1 \leq \frac{X}{Y} \leq \Omega_2 \bigwedge X < 0 \bigwedge Y > 0 \Leftrightarrow \Omega_1 \leq \frac{X}{Y} \leq \Omega_2 \bigwedge Y > 0$$

Thus, it suffices to compute the probability of the latter (simpler) event. Now, set $Z = \frac{X}{Y}$. Then $f_{Z|Y}(z|y) = f_X(z y)$, so:

$$f_{ZY}(z, y) = f_{Z|Y}(z|y) \cdot f_Y(y) = f_X(z y) \cdot f_Y(y)$$

Therefore:

$$\begin{aligned} \mathcal{P} \left[\Omega_1 \leq \frac{X}{Y} \leq \Omega_2, X < 0, Y > 0 \right] &= \mathcal{P} [\Omega_1 \leq Z \leq \Omega_2, Y > 0] \\ &= \int_0^{+\infty} \int_{\Omega_1}^{\Omega_2} f_{ZY}(z, y) \, dz \, dy \\ &= \int_0^{+\infty} \int_{\Omega_1}^{\Omega_2} f_X(z y) \cdot f_Y(y) \, dz \, dy \\ &= \int_0^{+\infty} \left(\int_{\Omega_1}^{\Omega_2} f_X(z y) \, dz \right) f_Y(y) \, dy \\ &= \int_0^{+\infty} \left(\int_{\Omega_1 y}^{\Omega_2 y} \frac{1}{y} f_X(x) \, dx \right) f_Y(y) \, dy \\ &= \int_0^{+\infty} \left(\int_{\Omega_1 y}^{\Omega_2 y} f_X(x) \, dx \right) \frac{f_Y(y)}{y} \, dy \end{aligned}$$

The proof of the second result is completely analogous. \square

The next proposition determines the probability that $NE_x(G^x)$ will have each of its possible values (see also Table 7.4):

Proposition 25. *Consider a normal noisy game $mG \sim G^0 + \mathcal{N}(M, D)$, and some $x \in \{r, c\}$. Then, the probabilities $\mathcal{P} [OP_x^{G^x}(1)]$, $\mathcal{P} [OP_x^{G^x}(2)]$, $\mathcal{P} [ROM_x^{G^x}(\omega_1, \omega_2)]$, $\mathcal{P} [RPM_x^{G^x}(\omega_1, \omega_2)]$ and $\mathcal{P} [IN_x^{G^x}]$ are as shown in Table 7.4.*

Proof. The results on $OP_x^{G^x}(i)$ ($i \in \{1, 2\}$) are direct consequences of Proposition 22, the fact that $U(y, x, i)$ are normal random variables as described in Table 7.3, and the independence/mutual exclusiveness of the involved random variables (which allow us to use the restricted disjunction/conjunction formulas from formula (B.4), Subsection B.1).

For the case of $ROM_x^{G^x}(\omega_1, \omega_2)$, applying Corollary 3, we get that $ROM_x^{G^x}(\omega_1, \omega_2)$ is true if

Results related to $NE_x(G^x)$ (Proposition 25)

$$\begin{aligned}
\mathcal{P} \left[\text{OP}_x^{G^x}(1) \right] &= (1 - F_{U(x,x,1)}(0)) \cdot (1 - F_{U(x,x,2)}(0)) \\
&\quad + (1 - F_{U(x,x,1)}(0)) \cdot F_{U(x,x,2)}(0) \cdot (1 - F_{U(x,\bar{x},1)}(0)) \cdot (1 - F_{U(x,\bar{x},2)}(0)) \\
&\quad + F_{U(x,x,1)}(0) \cdot (1 - F_{U(x,x,2)}(0)) \cdot F_{U(x,\bar{x},1)}(0) \cdot F_{U(x,\bar{x},2)}(0) \\
\mathcal{P} \left[\text{OP}_x^{G^x}(2) \right] &= F_{U(x,x,1)}(0) \cdot F_{U(x,x,2)}(0) \\
&\quad + F_{U(x,x,1)}(0) \cdot (1 - F_{U(x,x,2)}(0)) \cdot (1 - F_{U(x,\bar{x},1)}(0)) \cdot (1 - F_{U(x,\bar{x},2)}(0)) \\
&\quad + (1 - F_{U(x,x,1)}(0)) \cdot F_{U(x,x,2)}(0) \cdot F_{U(x,\bar{x},1)}(0) \cdot F_{U(x,\bar{x},2)}(0) \\
\mathcal{P} \left[\text{ROM}_x^{G^x}(\omega_1, \omega_2) \right] &= (1 - F_{U(x,x,1)}(0)) \cdot F_{U(x,x,2)}(0) \cdot \\
&\quad \int_0^{+\infty} \left(\int_{\frac{\omega_1-1}{\omega_1} u_2}^{\frac{\omega_2-1}{\omega_2} u_2} f_{U(x,\bar{x},1)}(u_1) \, du_1 \right) \frac{f_{U(x,\bar{x},2)}(u_2)}{u_2} \, du_2 \\
&\quad + F_{U(x,x,1)}(0) \cdot (1 - F_{U(x,x,2)}(0)) \cdot \\
&\quad \int_{-\infty}^0 \left(\int_{\frac{\omega_1-1}{\omega_1} u_2}^{\frac{\omega_2-1}{\omega_2} u_2} f_{U(x,\bar{x},1)}(u_1) \, du_1 \right) \frac{f_{U(x,\bar{x},2)}(u_2)}{u_2} \, du_2 \\
\mathcal{P} \left[\text{RPM}_x^{G^x}(\omega_1, \omega_2) \right] &= (1 - F_{U(x,x,1)}(0)) \cdot F_{U(x,x,2)}(0) \cdot \\
&\quad \int_{-\infty}^0 \left(\int_{\frac{\omega_1-1}{\omega_1} u_2}^{\frac{\omega_2-1}{\omega_2} u_2} f_{U(x,\bar{x},1)}(u_1) \, du_1 \right) \frac{f_{U(x,\bar{x},2)}(u_2)}{u_2} \, du_2 + \\
&\quad F_{U(x,x,1)}(0) \cdot (1 - F_{U(x,x,2)}(0)) \cdot \\
&\quad \int_0^{+\infty} \left(\int_{\frac{\omega_1-1}{\omega_1} u_2}^{\frac{\omega_2-1}{\omega_2} u_2} f_{U(x,\bar{x},1)}(u_1) \, du_1 \right) \frac{f_{U(x,\bar{x},2)}(u_2)}{u_2} \, du_2 \\
\mathcal{P} \left[\text{IN}_x^{G^x} \right] &= 0
\end{aligned}$$

Table 7.4: Various probabilities pertaining to a given $mG \sim G^0 + \mathcal{N}(M, D)$ (see also Proposition 25).

and only if:

$$\begin{aligned}
&(\text{UGAIN}^{G^x}(x, 1) > 0) \bigwedge (\text{UGAIN}^{G^x}(x, 2) < 0) \bigwedge (\text{UGAIN}^{G^x}(\bar{x}, 1) < 0) \bigwedge \\
&(\text{UGAIN}^{G^x}(\bar{x}, 2) > 0) \bigwedge \left(\omega_1 < \frac{\text{UGAIN}^{G^x}(\bar{x}, 2)}{\text{UGAIN}^{G^x}(\bar{x}, 2) - \text{UGAIN}^{G^x}(\bar{x}, 1)} < \omega_2 \right) \\
&\vee \\
&(\text{UGAIN}^{G^x}(x, 1) < 0) \bigwedge (\text{UGAIN}^{G^x}(x, 2) > 0) \bigwedge (\text{UGAIN}^{G^x}(\bar{x}, 1) > 0) \bigwedge \\
&(\text{UGAIN}^{G^x}(\bar{x}, 2) < 0) \bigwedge \left(\omega_1 < \frac{\text{UGAIN}^{G^x}(\bar{x}, 2)}{\text{UGAIN}^{G^x}(\bar{x}, 2) - \text{UGAIN}^{G^x}(\bar{x}, 1)} < \omega_2 \right)
\end{aligned}$$

Obviously, the above disjunction contains mutually exclusive events, so the probability $\mathcal{P}[\text{ROM}_x^{G^x}(\omega_1, \omega_2)]$ is the sum of the probability of each disjunct (by the restricted disjunctive formula – see formula (B.4), Subsection B.1). So, let us compute the probability of the first disjunct.

We observe that the events $\text{UGAIN}^{G^x}(x, 1)$, $\text{UGAIN}^{G^x}(x, 2)$ are independent to each other and also independent to the other conjuncts. Moreover:

$$\omega_1 < \frac{\text{UGAIN}^{G^x}(\bar{x}, 2)}{\text{UGAIN}^{G^x}(\bar{x}, 2) - \text{UGAIN}^{G^x}(\bar{x}, 1)} < \omega_2 \Leftrightarrow \frac{\omega_1 - 1}{\omega_1} < \frac{\text{UGAIN}^{G^x}(\bar{x}, 1)}{\text{UGAIN}^{G^x}(\bar{x}, 2)} < \frac{\omega_2 - 1}{\omega_2}$$

Thus, we can apply Lemma 12 for the last three conjuncts (for $\Omega_1 = \frac{\omega_1 - 1}{\omega_1}$, $\Omega_2 = \frac{\omega_2 - 1}{\omega_2}$), getting that the probability of the first conjunction is equal to:

$$(1 - F_{U(x, x, 1)}(0)) \cdot F_{U(x, x, 2)}(0) \cdot \int_0^{+\infty} \left(\int_{\Omega_1 u_2}^{\Omega_2 u_2} f_{U(x, \bar{x}, 1)}(u_1) du_1 \right) \frac{f_{U(x, \bar{x}, 2)}(u_2)}{u_2} du_2$$

Working analogously for the second disjunct, and summing the resulting probability with the one above, we get the result.

For $\text{RPM}_x^{G^x}(\omega_1, \omega_2)$, we work analogously, applying the second bullet of Corollary 3 as above. For $\text{IN}_x^{G^x}$, we observe that if $\text{IN}_x^{G^x}$ is true, then G^x is degenerate, which has probability 0. \square

Proposition 25 (and the respective Table 7.4), combined with Proposition 23 (and the respective Table 7.2) easily leads to the following theorems (summarized in Table 7.5):

Theorem 7. *Consider a normal noisy game $mG \sim G^0 + \mathcal{N}(M, D)$. Then:*

$$\mathcal{P}[mG : \varepsilon\text{-misinformed}] = \mathcal{P}_r^{\text{mis}} \cdot \mathcal{P}_c^{\text{mis}}$$

where, for $x \in \{r, c\}$, $\mathcal{P}_x^{\text{mis}}$ is determined by the second column of Table 7.5.

Proof. The proof is direct from Proposition 23 (and the respective Table 7.2), combined with the fact that the different cases in the disjunction are mutually exclusive, so we can use the restricted disjunction formula of (B.4) in Subsection B.1. \square

Theorem 8. *Consider a normal noisy game $mG \sim G^0 + \mathcal{N}(M, D)$. Then:*

$$\mathcal{P}[mG : \text{inverse-}\varepsilon\text{-misinformed}] = \mathcal{P}_r^{\text{inv}} \cdot \mathcal{P}_c^{\text{inv}}$$

where, for $x \in \{r, c\}$, $\mathcal{P}_x^{\text{inv}}$ is determined by the third column of Table 7.5.

Proof. The proof is direct from Proposition 23 (and the respective Table 7.2), combined with the fact that the different cases in the disjunction are mutually exclusive, so we can use the restricted disjunction formula of (B.4) in Subsection B.1. \square

Condition on G^0 (value of $NE_x(G^0)$)	Probability $\mathcal{P}_x^{mis} (x \in \{r, c\})$ for ε -misinformed (Theorem 7)	Probability $\mathcal{P}_x^{inv} (x \in \{r, c\})$ for inverse- ε -misinformed (Theorem 8)
$OP_x^{G^0}(i)$	$\mathcal{P} \left[OP_x^{G^x}(i) \right]$	$\mathcal{P} \left[OP_x^{G^x}(i) \right] +$ $\mathcal{P} \left[RPM_x^{G^x}(0, 1) \right]$
$OM_x^{G^0}(p^0)$	$\mathcal{P} \left[ROM_x^{G^x}(\omega_1, \omega_2) \right]$	$\mathcal{P} \left[ROM_x^{G^x}(\omega_1, \omega_2) \right] +$ $\mathcal{P} \left[RPM_x^{G^x}(\omega_1, \omega_2) \right]$
$PM_x^{G^0}(p^0)$	$\mathcal{P} \left[RPM_x^{G^x}(\omega_1, \omega_2) \right] +$ $\mathcal{P} \left[ROM_x^{G^x}(\omega_1, \omega_2) \right] +$ $\mathcal{P} \left[OP_x^{G^x}(1) \right] +$ $\mathcal{P} \left[OP_x^{G^x}(2) \right]$	$\mathcal{P} \left[RPM_x^{G^x}(\omega_1, \omega_2) \right]$
$IN_x^{G^0}$	1	If $\varepsilon \leq 0.5$: 0 If $\varepsilon > 0.5$: $\mathcal{P} \left[RPM_x^{G^x}(\omega'_1, \omega'_2) \right]$

In all the above:

$$i \in \{1, 2\}, 0 < p^0 < 1,$$

$$\omega_1 = \max\{0, p^0 - \varepsilon\}, \omega_2 = \min\{1, p^0 + \varepsilon\},$$

$$\omega'_1 = \max\{0, 1 - \varepsilon\}, \omega'_2 = \min\{1, \varepsilon\}$$

Table 7.5: Probabilities for ε -misinformed and inverse- ε -misinformed ($\mathcal{P}_r^{mis} \cdot \mathcal{P}_c^{mis}$ and $\mathcal{P}_r^{inv} \cdot \mathcal{P}_c^{inv}$ respectively – see also Theorems 7, 8)

7.4 Results for Noisy games

The results of Section 7.3 provide the formulas to compute the probability of a given normal noisy game to be (inverse-) ε -misinformed (i.e., behaviourally consistent). In this section, we explore the properties of these formulas, to understand better their behaviour.

To do so, we first observe that the probability of a normal noisy game $mG \sim G^0 + \mathcal{N}(M, D)$ being behaviourally consistent is essentially a function of:

- The tolerance ε .
- The payoff matrix of the actual game of mG . This affects the probabilities in two ways: first, because it determines the equilibria of G^0 , and, thus, the case to consider in Table 7.5; second, because it affects $\mu_{U(y,x,i)}$ (see Table 7.3).
- The noise pattern, determined by the matrices M, D .

In the following subsections, we study the effect of each of these parameters on the probability of mG being (inverse-) ε -misinformed, as well as the relationship between the two metrics of

behavioural consistency.

7.4.1 Effect of modifying tolerance (ε)

With regards to tolerance (ε), we expect that larger values of tolerance would translate to higher probability of behavioural consistency. Although this is true, we also observe that there are several cases where increasing tolerance does not affect the probability of behavioural consistency. The following proposition clarifies the situation:

Condition on G^0 (value of $NE(G^0)$)	Monotonicity properties	
	For ε -misinformed	For inverse- ε -misinformed
$OP_r^{G^0}(i) \wedge OP_c^{G^0}(j)$ for some $i, j \in \{1, 2\}$	Constant for all $\varepsilon \geq 0$	
$OM_r^{G^0}(p^0) \wedge OM_c^{G^0}(q^0)$ for some $0 < p^0 < 1, 0 < q^0 < 1$	Strictly increasing for $0 \leq \varepsilon \leq \max\{p^0, q^0, 1 - p^0, 1 - q^0\}$, constant otherwise	
$PM_r^{G^0}(p^0) \wedge PM_c^{G^0}(q^0)$ for some $0 < p^0 < 1, 0 < q^0 < 1$	Strictly increasing for $0 \leq \varepsilon \leq \max\{p^0, q^0, 1 - p^0, 1 - q^0\}$, constant otherwise	
$IN_r^{G^0} \vee IN_c^{G^0}$	Constant for all $\varepsilon \geq 0$	Strictly increasing for $0.5 \leq \varepsilon \leq 1$, constant otherwise

Table 7.6: Effect of tolerance on behavioural consistency (monotonicity)

Proposition 26. Consider some $mG \sim G^0 + \mathcal{N}(M, D)$ and non-negative $\varepsilon_1, \varepsilon_2$, such that $\varepsilon_1 < \varepsilon_2$. Then:

1. If $NE(G^0)$ contains a single pure strategy, then:
 - $\mathcal{P}[mG : \varepsilon_1\text{-misinformed}] = \mathcal{P}[mG : \varepsilon_2\text{-misinformed}]$
 - $\mathcal{P}[mG : \text{inverse-}\varepsilon_1\text{-misinformed}] = \mathcal{P}[mG : \text{inverse-}\varepsilon_2\text{-misinformed}]$
2. If $NE(G^0)$ is finite and $((p^0, 1 - p^0), (q^0, 1 - q^0)) \in NE(G^0)$ for some $0 < p^0 < 1$, $0 < q^0 < 1$, then:
 - (a) If $\max\{p^0, q^0, 1 - p^0, 1 - q^0\} \leq \varepsilon_1$, then:
 - $\mathcal{P}[mG : \varepsilon_1\text{-misinformed}] = \mathcal{P}[mG : \varepsilon_2\text{-misinformed}]$
 - $\mathcal{P}[mG : \text{inverse-}\varepsilon_1\text{-misinformed}] = \mathcal{P}[mG : \text{inverse-}\varepsilon_2\text{-misinformed}]$

(b) If $\max\{p^0, q^0, 1 - p^0, 1 - q^0\} \leq \varepsilon_1$, then:

- $\mathcal{P}[mG : \varepsilon_1\text{-misinformed}] < \mathcal{P}[mG : \varepsilon_2\text{-misinformed}]$
- $\mathcal{P}[mG : \text{inverse-}\varepsilon_1\text{-misinformed}] < \mathcal{P}[mG : \text{inverse-}\varepsilon_2\text{-misinformed}]$

3. If $NE(G^0)$ is infinite, then:

(a) If $\varepsilon_1 \geq 1$ or $\varepsilon_2 \leq 0.5$, then:

- $\mathcal{P}[mG : \varepsilon_1\text{-misinformed}] = \mathcal{P}[mG : \varepsilon_2\text{-misinformed}]$
- $\mathcal{P}[mG : \text{inverse-}\varepsilon_1\text{-misinformed}] = \mathcal{P}[mG : \text{inverse-}\varepsilon_2\text{-misinformed}]$

(b) If $\varepsilon_1 < 1$ and $\varepsilon_2 > 0.5$, then:

- $\mathcal{P}[mG : \varepsilon_1\text{-misinformed}] = \mathcal{P}[mG : \varepsilon_2\text{-misinformed}]$
- $\mathcal{P}[mG : \text{inverse-}\varepsilon_1\text{-misinformed}] < \mathcal{P}[mG : \text{inverse-}\varepsilon_2\text{-misinformed}]$

Proof. We first observe that, for any $x \in \{r, c\}$ and any a, b, c such that: $0 \leq a \leq b \leq c \leq 1$, we have that:

$$\mathcal{P}[\text{ROM}_x^{G^x}(a, c)] = \mathcal{P}[\text{ROM}_x^{G^x}(a, b)] + \mathcal{P}[\text{ROM}_x^{G^x}(b, c)] \quad (\text{ROM1})$$

$$\mathcal{P}[\text{ROM}_x^{G^x}(a, c)] = 0 \Leftrightarrow a = c \quad (\text{ROM2})$$

$$\mathcal{P}[\text{RPM}_x^{G^x}(a, c)] = \mathcal{P}[\text{RPM}_x^{G^x}(a, b)] + \mathcal{P}[\text{RPM}_x^{G^x}(b, c)] \quad (\text{RPM1})$$

$$\mathcal{P}[\text{RPM}_x^{G^x}(a, c)] = 0 \Leftrightarrow a = c \quad (\text{RPM2})$$

From Theorem 7, and for $i = 1, 2$:

$$\mathcal{P}[mG : \varepsilon_i\text{-misinformed}] = P_{r,i} \cdot P_{c,i},$$

where $P_{r,i}, P_{c,i}$ are determined by the second column of Table 7.5 for the respective ε_i .

Similarly, from Theorem 8, and for $i = 1, 2$:

$$\mathcal{P}[mG : \text{inverse-}\varepsilon_i\text{-misinformed}] = P'_{r,i} \cdot P'_{c,i},$$

where $P'_{r,i}, P'_{c,i}$ are determined by the third column of Table 7.5 for the respective ε_i .

Now, let us focus on the first bullet of the proposition. By Tables 7.3, 7.4, 7.5, it is easy to conclude that, for any $x \in \{r, c\}$, $i \in \{1, 2\}$, the computation of $P_{x,i}, P'_{x,i}$ is not affected by the value of ε_i , and, thus: $P_{x,1} = P_{x,2}$, $P'_{x,1} = P'_{x,2}$ for $x \in \{r, c\}$, which shows the result.

Now, let us focus on the second bullet, and let us consider $P_{r,i}, P'_{r,i}$ first. Set:

$$\begin{aligned} \omega_{1,1} &= \max\{0, p^0 - \varepsilon_1\}, & \omega_{2,1} &= \min\{1, p^0 + \varepsilon_1\}, \\ \omega_{1,2} &= \max\{0, p^0 - \varepsilon_2\}, & \omega_{2,2} &= \min\{1, p^0 + \varepsilon_2\} \end{aligned}$$

Since $0 \leq \varepsilon_1 < \varepsilon_2$, we get that: $0 \leq \omega_{1,2} \leq \omega_{1,1} \leq \omega_{2,1} \leq \omega_{2,2} \leq 1$. Moreover, since $\varepsilon_1 < \varepsilon_2$, it follows that:

$$\omega_{1,1} = \omega_{1,2} \Leftrightarrow \omega_{1,1} = \omega_{1,2} = 0 \Leftrightarrow \begin{pmatrix} p^0 \leq \varepsilon_1 \\ \text{and} \\ p^0 \leq \varepsilon_2 \end{pmatrix} \Leftrightarrow p^0 \leq \varepsilon_1$$

Analogously:

$$\omega_{2,1} = \omega_{2,2} \Leftrightarrow \omega_{2,1} = \omega_{2,2} = 1 \Leftrightarrow \begin{pmatrix} 1 - p^0 \leq \varepsilon_1 \\ \text{and} \\ 1 - p^0 \leq \varepsilon_2 \end{pmatrix} \Leftrightarrow 1 - p^0 \leq \varepsilon_1$$

Using the order among $\omega_{i,j}$, and by applying (ROM1) twice, we get that:

$$\mathcal{P} \left[\text{ROM}_r^{Gr}(\omega_{1,2}, \omega_{2,2}) \right] = \mathcal{P} \left[\text{ROM}_r^{Gr}(\omega_{1,2}, \omega_{1,1}) \right] + \mathcal{P} \left[\text{ROM}_r^{Gr}(\omega_{1,1}, \omega_{2,1}) \right] + \mathcal{P} \left[\text{ROM}_r^{Gr}(\omega_{2,1}, \omega_{2,2}) \right]$$

Now given the fact that probabilities are non-negative, and (ROM2), we have:

$$\begin{aligned} \mathcal{P} \left[\text{ROM}_r^{Gr}(\omega_{1,1}, \omega_{2,1}) \right] &\leq \mathcal{P} \left[\text{ROM}_r^{Gr}(\omega_{1,2}, \omega_{2,2}) \right] \text{ and:} \\ \mathcal{P} \left[\text{ROM}_r^{Gr}(\omega_{1,1}, \omega_{2,1}) \right] &= \mathcal{P} \left[\text{ROM}_r^{Gr}(\omega_{1,2}, \omega_{2,2}) \right] \Leftrightarrow \\ &\omega_{1,1} = \omega_{1,2} \text{ and } \omega_{2,1} = \omega_{2,2} \end{aligned}$$

Using analogous reasoning we get:

$$\begin{aligned} \mathcal{P} \left[\text{RPM}_r^{Gr}(\omega_{1,1}, \omega_{2,1}) \right] &\leq \mathcal{P} \left[\text{RPM}_r^{Gr}(\omega_{1,2}, \omega_{2,2}) \right] \text{ and:} \\ \mathcal{P} \left[\text{RPM}_r^{Gr}(\omega_{1,1}, \omega_{2,1}) \right] &= \mathcal{P} \left[\text{RPM}_r^{Gr}(\omega_{1,2}, \omega_{2,2}) \right] \Leftrightarrow \\ &\omega_{1,1} = \omega_{1,2} \text{ and } \omega_{2,1} = \omega_{2,2} \end{aligned}$$

Using the above, and Tables 7.3, 7.4, 7.5, we can easily conclude that $P_{r,1} \leq P_{r,2}$ and $P'_{r,1} \leq P'_{r,2}$. Moreover:

$$P_{r,1} = P_{r,2} \Leftrightarrow \begin{pmatrix} \omega_{1,1} = \omega_{1,2} \\ \text{and} \\ \omega_{2,1} = \omega_{2,2} \end{pmatrix} \Leftrightarrow \begin{pmatrix} p^0 \leq \varepsilon_1 \\ \text{and} \\ 1 - p^0 \leq \varepsilon_1 \end{pmatrix}$$

Analogously:

$$P'_{r,1} = P'_{r,2} \Leftrightarrow \begin{pmatrix} p^0 \leq \varepsilon_1 \\ \text{and} \\ 1 - p^0 \leq \varepsilon_1 \end{pmatrix}$$

Reasoning analogously for the case of $P_{c,i}, P'_{c,i}$, we get:

$$\text{For } P_{c,1} \leq P_{c,2} : P_{c,1} = P_{c,2} \Leftrightarrow \begin{pmatrix} q^0 \leq \varepsilon_1 \\ \text{and} \\ 1 - q^0 \leq \varepsilon_1 \end{pmatrix}$$

$$\text{For } P'_{c,1} \leq P'_{c,2} : P'_{c,1} = P'_{c,2} \Leftrightarrow \begin{pmatrix} q^0 \leq \varepsilon_1 \\ \text{and} \\ 1 - q^0 \leq \varepsilon_1 \end{pmatrix}$$

By the hypothesis of the second bullet with regards to $NE(G^0)$, Tables 7.3, 7.4, 7.5, and the above relations, the cases (2a), (2b) of the Theorem follow easily.

Now let us focus on the third bullet. First, we observe that, by Table 7.5, the result is obvious for the case of ε -misinformed, so let us focus on the case of inverse- ε -misinformed. If $\varepsilon_2 \leq 0.5$, then $\varepsilon_1 \leq 0.5$, so the result is again obvious by Table 7.5. So let us focus on the scenario where $\varepsilon_2 > 0.5$.

To show the result for this case, we use an approach similar to the one employed for the second bullet. In particular, we consider $P'_{r,i}$ first. Set:

$$\begin{aligned} \omega'_{1,1} &= \max\{0, 1 - \varepsilon_1\}, & \omega'_{2,1} &= \min\{1, \varepsilon_1\}, \\ \omega'_{1,2} &= \max\{0, 1 - \varepsilon_2\}, & \omega'_{2,2} &= \min\{1, \varepsilon_2\} \end{aligned}$$

Using an analogous procedure (as in the second bullet), and the fact that $0 \leq \varepsilon_1 < \varepsilon_2$, we conclude that:

$$\begin{aligned} 0 &\leq \omega'_{1,2} \leq \omega'_{1,1} \leq \omega'_{2,1} \leq \omega'_{2,2} \leq 1 \\ \omega'_{1,1} &= \omega'_{1,2} \Leftrightarrow \varepsilon_1 \leq 1 \\ \omega'_{2,1} &= \omega'_{2,2} \Leftrightarrow \varepsilon_1 \leq 1 \end{aligned}$$

Also, using (RPM1), (RPM2), and the fact that probabilities are non-negative, we get, as in the second bullet:

$$\begin{aligned} \mathcal{P} \left[\text{RPM}_r^{Gr}(\omega'_{1,1}, \omega'_{2,1}) \right] &\leq \mathcal{P} \left[\text{RPM}_r^{Gr}(\omega'_{1,2}, \omega'_{2,2}) \right] \\ \mathcal{P} \left[\text{RPM}_r^{Gr}(\omega'_{1,1}, \omega'_{2,1}) \right] &= \mathcal{P} \left[\text{RPM}_r^{Gr}(\omega'_{1,2}, \omega'_{2,2}) \right] \\ &\Leftrightarrow \omega'_{1,1} = \omega'_{1,2} \text{ and } \omega'_{2,1} = \omega'_{2,2} \end{aligned}$$

Therefore, given that $\varepsilon_1 > \varepsilon_2 > 0.5$:

$$P'_{r,1} \leq P'_{r,2} \text{ and } P'_{r,1} = P'_{r,2} \Leftrightarrow \varepsilon_1 \leq 1$$

Working analogously for $P'_{c,i}$, we get:

$$P'_{c,1} \leq P'_{c,2} \text{ and } P'_{c,1} = P'_{c,2} \Leftrightarrow \varepsilon_1 \leq 1$$

By the hypothesis of the second bullet with regards to $NE(G^0)$, Tables 7.3, 7.4, 7.5, and the above relations, the remaining subcases of (3a), (3b) of the Theorem follow easily. \square

Condition on G^0 (value of $NE_x(G^0)$)	Minimal value for probability P_x ($x \in \{r, c\}$) for ε -misinformed	Minimal value for probability P_x ($x \in \{r, c\}$) for inverse- ε -misinformed
$OP_x^{G^0}(i)$ for some $i \in \{1, 2\}$	$\mathcal{P} \left[OP_x^{G^x}(i) \right]$	$\mathcal{P} \left[OP_x^{G^x}(i) \right] +$ $\mathcal{P} \left[RPM_x^{G^x}(0, 1) \right]$
$OM_x^{G^0}(p^0)$ for some $0 < p^0 < 1$	0	0
$PM_x^{G^0}(p^0)$ for some $0 < p^0 < 1$	$\mathcal{P} \left[OP_x^{G^x}(1) \right] +$ $\mathcal{P} \left[OP_x^{G^x}(2) \right]$	0
$IN_x^{G^0}$	1	0

(a) Minimal values

Condition on G^0 (value of $NE_x(G^0)$)	Maximal value for probability P_x ($x \in \{r, c\}$) for ε -misinformed	Maximal value for probability P_x ($x \in \{r, c\}$) for inverse- ε -misinformed
$OP_x^{G^0}(i)$ for some $i \in \{1, 2\}$	$\mathcal{P} \left[OP_x^{G^x}(i) \right]$	$\mathcal{P} \left[OP_x^{G^x}(i) \right] +$ $\mathcal{P} \left[RPM_x^{G^x}(0, 1) \right]$
$OM_x^{G^0}(p^0)$ for some $0 < p^0 < 1$	$\mathcal{P} \left[ROM_x^{G^x}(0, 1) \right]$	$\mathcal{P} \left[ROM_x^{G^x}(0, 1) \right] +$ $\mathcal{P} \left[RPM_x^{G^x}(0, 1) \right]$
$PM_x^{G^0}(p^0)$ for some $0 < p^0 < 1$	1	$\mathcal{P} \left[RPM_x^{G^x}(0, 1) \right]$
$IN_x^{G^0}$	1	$\mathcal{P} \left[RPM_x^{G^x}(0, 1) \right]$

(b) Maximal values

Table 7.7: Minimal and maximal values for the probabilities of mG being (inverse-) ε -misinformed (resulting by multiplying \mathcal{P}_r^{mis} with \mathcal{P}_c^{mis} and \mathcal{P}_r^{inv} with \mathcal{P}_c^{inv} respectively)

Proposition 26 has several interesting consequences. First, we note that the probability for a given mG to be (inverse-) ε -misinformed is non-decreasing with respect to ε . When there is a pure Nash equilibrium, the choice of ε is irrelevant to the value of these probabilities. When there is a mixed Nash equilibrium (case 2 of the proposition), there is a limit above which ε does not affect the value of the related probability; this limit depends on the actual mixed equilibrium, but it is always equal to, or larger than 0.5, and smaller than 1. Finally, in the case where there is an infinite number of equilibria, ε affects the probabilities only for certain values (between 0.5 and 1, and only for the inverse- ε -misinformed case), as detailed in case 3 of Proposition 26. These are summarised in Table 7.6.

Our results (and Table 7.6) indicate that the minimal value for the probability of mG being (inverse-) ε -misinformed is given for $\varepsilon = 0$. Its maximal value is taken for an appropriate ε (depending on the case); in all cases $\varepsilon = 1$ would also give that maximal value. These maximal/minimal values can be easily deduced by Table 7.5 for the above choices of ε , and are given in Table 7.7 for convenience. Note that the actual result for the minimal/maximal values results by multiplying \mathcal{P}_r^{mis} with \mathcal{P}_r^{mis} , and \mathcal{P}_r^{inv} with \mathcal{P}_r^{inv} for ε -misinformed and inverse- ε -misinformed respectively.

Another important result (albeit relatively obvious) is that the probability of mG being (inverse-) ε -misinformed, viewed as a function of ε , is continuous. This is a direct consequence of the results in Tables 7.3, 7.4, 7.5. An important consequence of this fact, by well-known results of calculus, is that, for any given target value for the probabilities of (inverse-) ε -misinformed, there exists some ε whose application would result to that value for the respective probability.

7.4.2 Effect of changing the game (G^0) and the mean (M)

Consider a misinformation game $mG \sim G^0 + \mathcal{N}(M, D)$, and let us informally ponder on the effect of bias in the noise of a game. A biased noise is noise whose mean M is non-zero, i.e., $M \neq [\mathbf{0}]$. Let us consider only player r , for simplicity. In such a scenario, we know that $G^r \sim G^0 + \mathcal{N}(M^r, D^r)$. Observe that this is the same as writing $G^r \sim (G^0 + M^r) + \mathcal{N}([\mathbf{0}], D^r)$. Using this simple reasoning, the computation of the probabilities of behavioural consistency for mG for biased noise can be reduced to computations related to some \overline{mG} with unbiased noise ($M = [\mathbf{0}]$), whose actual game will be the sum of G^0 and M^x .

However, there are two caveats here. First, since M^r may be different than M^c , our original misinformation game is essentially reduced to two different misinformation games (say $\overline{mG}_r, \overline{mG}_c$), i.e., one per player. Second, in the case where the equilibria of G^0 are different than the equilibria of $G^0 + M^x$, care should be taken to consult the proper line in Table 7.5 while computing the probability of mG being (inverse-) ε -misinformed. In particular, the line to consider should be the one related to the equilibria of G^0 , not $G^0 + M^x$. This means that the probability of mG being (inverse-) ε -misinformed may not be the same as the respective probability for $\overline{mG}_r, \overline{mG}_c$.

To prove the above ideas formally, we start with the following proposition:

Proposition 27. *Consider two noisy games $mG \sim G^0 + \mathcal{N}(M, D)$, $\overline{mG} \sim \overline{G^0} + \mathcal{N}(\overline{M}, D)$. Suppose that there exists $a \in \mathbb{R}$, $x \in \{r, c\}$ such that $G^0 + M^x = \overline{G^0} + \overline{M}^x + [a]$. Then:*

- For any $i \in \{1, 2\}$,

$$\mathcal{P} \left[OP_x^{G^x}(i) \right] = \mathcal{P} \left[OP_x^{\overline{G}^x}(i) \right]$$

- For any $0 \leq \omega_1 \leq \omega_2 \leq 1$,

$$\mathcal{P} \left[ROM_x^{G^x}(\omega_1, \omega_2) \right] = \mathcal{P} \left[ROM_x^{\overline{G}^x}(\omega_1, \omega_2) \right]$$

- For any $0 \leq \omega_1 \leq \omega_2 \leq 1$,

$$\mathcal{P} \left[RPM_x^{G^x}(\omega_1, \omega_2) \right] = \mathcal{P} \left[RPM_x^{\overline{G}^x}(\omega_1, \omega_2) \right]$$

Proof. From Table 7.3, we observe that, for the given x , and for any $i \in \{1, 2\}$:

$$\begin{aligned} \mu_{U(x,r,i)} &= (P_r^0[1, i] + M_r^x[1, i]) - (P_r^0[2, i] + M_r^x[2, i]) \\ &= (\overline{P}_r^0[1, i] + \overline{M}_r^x[1, i] + a) - (\overline{P}_r^0[2, i] + \overline{M}_r^x[2, i] + a) = \overline{\mu}_{U(x,r,i)}. \end{aligned}$$

Analogously, we can show that $\mu_{U(x,c,i)} = \overline{\mu}_{U(x,c,i)}$ for any $i \in \{1, 2\}$. Also, it is clear that $d_{U(x,y,i)} = \overline{d}_{U(x,y,i)}$ for any $y \in \{r, c\}$, $i \in \{1, 2\}$. Combining these two facts, the results are obvious. \square

Proposition 27 implies that, given a noisy game $mG \sim G^0 + \mathcal{N}(M, D)$ and a player $x \in \{r, c\}$, we can generate some other noisy game (say \overline{mG}), whose probabilities related to the various outcomes (equilibria) of the game \overline{G}^x of \overline{mG} are identical to the respective ones for G^x (in mG). As a matter of fact, there is an infinite number of noisy games that satisfy this property: for any given $\overline{G^0}$ we can find an infinite number of \overline{M} that do this, and for any given \overline{M} we can find an infinite number of $\overline{G^0}$ that do this. This observation motivates us to consider some interesting special cases, formalised as corollaries below.

The first interesting case is when $\overline{M} = [\mathbf{0}]$. Given a noisy game mG , the following corollary shows that the probabilities related to the various outcomes (equilibria) of the game G^x in mG can be predicted by looking at a properly defined noisy game \overline{mG} where the noise is unbiased (i.e., $\overline{M} = [\mathbf{0}]$). Formally:

Corollary 4. *Consider a noisy game $mG \sim G^0 + \mathcal{N}(M, D)$, and some $x \in \{r, c\}$. Set $\overline{G^0} = G^0 + M^x$, and $\overline{mG} \sim \overline{G^0} + \mathcal{N}([\mathbf{0}], D)$. Then:*

- For any $i \in \{1, 2\}$,

$$\mathcal{P} \left[OP_x^{G^x}(i) \right] = \mathcal{P} \left[OP_x^{\overline{G}^x}(i) \right]$$

- For any $0 \leq \omega_1 \leq \omega_2 \leq 1$,

$$\mathcal{P} \left[\text{ROM}_x^{G^x}(\omega_1, \omega_2) \right] = \mathcal{P} \left[\text{ROM}_x^{\overline{G^x}}(\omega_1, \omega_2) \right]$$

- For any $0 \leq \omega_1 \leq \omega_2 \leq 1$,

$$\mathcal{P} \left[\text{RPM}_x^{G^x}(\omega_1, \omega_2) \right] = \mathcal{P} \left[\text{RPM}_x^{\overline{G^x}}(\omega_1, \omega_2) \right]$$

Combining Corollary 4 with Theorems 7, 8, it is easy to compute the probability that mG is (inverse-) ε -misinformed, using the respective probabilities for \overline{mG} . This is one of the main results of this subsection, as it allows us to restrict our study to noisy games with unbiased noise only.

An interesting observation is that Corollary 4 applies for some $x \in \{r, c\}$. Thus, we need to define two different \overline{mG} (one for each player $x \in \{r, c\}$) in order to compute the probability that mG is (inverse-) ε -misinformed. The following corollary holds for both $x \in \{r, c\}$ (and thus foregoes this need), but applies only when $M^r = M^c$, i.e., when the noise received by the two players has the same bias:

Corollary 5. *Consider a noisy game $mG \sim G^0 + \mathcal{N}(M, D)$, where $M = (M^*; M^*)$. Set $\overline{G^0} = G^0 + M^*$, and $\overline{mG} \sim \overline{G^0} + \mathcal{N}([0], D)$. Then:*

- For any $i \in \{1, 2\}$ and $x \in \{r, c\}$,

$$\mathcal{P} \left[\text{OP}_x^{G^x}(i) \right] = \mathcal{P} \left[\text{OP}_x^{\overline{G^x}}(i) \right]$$

- For any $0 \leq \omega_1 \leq \omega_2 \leq 1$ and $x \in \{r, c\}$,

$$\mathcal{P} \left[\text{ROM}_x^{G^x}(\omega_1, \omega_2) \right] = \mathcal{P} \left[\text{ROM}_x^{\overline{G^x}}(\omega_1, \omega_2) \right]$$

- For any $0 \leq \omega_1 \leq \omega_2 \leq 1$ and $x \in \{r, c\}$,

$$\mathcal{P} \left[\text{RPM}_x^{G^x}(\omega_1, \omega_2) \right] = \mathcal{P} \left[\text{RPM}_x^{\overline{G^x}}(\omega_1, \omega_2) \right]$$

Proposition 27 and Corollary 4 provide the probability of the different events to occur (e.g., the probability that G^x has a certain equilibrium), but do not directly provide the probability for mG being (inverse-) ε -misinformed. Indeed, since G^0 and $\overline{G^0}$ may have different equilibria, the computation of the probabilities for mG and \overline{mG} being (inverse-) ε -misinformed may use different rows in Table 7.5. This is unnecessary only when the two games have the same equilibria:

Corollary 6. Consider a noisy game $mG \sim G^0 + \mathcal{N}(M, D)$. Set:

$$\overline{G^0} = G^0 + M^r, \quad \widehat{G^0} = G^0 + M^c, \quad \overline{mG} \sim \overline{G^0} + \mathcal{N}([0], D), \quad \widehat{mG} \sim \widehat{G^0} + \mathcal{N}([0], D)$$

If $NE(G^0) = NE(\overline{G^0}) = NE(\widehat{G^0})$ then:

- $\mathcal{P}[mG : \varepsilon\text{-misinformed}] = \overline{\mathcal{P}_r^{mis}} \cdot \widehat{\mathcal{P}_c^{mis}}$
- $\mathcal{P}[mG : \text{inverse-}\varepsilon\text{-misinformed}] = \overline{\mathcal{P}_r^{inv}} \cdot \widehat{\mathcal{P}_c^{inv}}$

where $\overline{\mathcal{P}_r^{mis}}, \overline{\mathcal{P}_r^{inv}}, \widehat{\mathcal{P}_c^{mis}}, \widehat{\mathcal{P}_c^{inv}}$ are the probabilities of Table 7.5 for $\overline{mG}, \widehat{mG}$ respectively.

Note that, in Corollary 6, the computation of the probability for mG to be (inverse-) ε -misinformed, occurs via the combination of quantities from two different noisy games $(\overline{mG}, \widehat{mG})$. As with Corollary 4, this can be avoided when the noise received by the two players has the same bias, in which case we get a direct computation of the related probability:

Corollary 7. Consider the noisy game $mG \sim G^0 + \mathcal{N}(M, D)$, where $M = (M^*; M^*)$. Set $\overline{G^0} = G^0 + M^*$ and $\overline{mG} \sim \overline{G^0} + \mathcal{N}([0], D)$. If $NE(G^0) = NE(\overline{G^0})$ then:

- $\mathcal{P}[mG : \varepsilon\text{-misinformed}] = \mathcal{P}[\overline{mG} : \varepsilon\text{-misinformed}]$
- $\mathcal{P}[mG : \text{inverse-}\varepsilon\text{-misinformed}] = \mathcal{P}[\overline{mG} : \text{inverse-}\varepsilon\text{-misinformed}]$

Corollary 7 is the most specific result, as it gives us a method of computing the probabilities of a noisy game being (inverse-) ε -misinformed using the respective probabilities of another noisy game, under specific assumptions.

The last proposition of this subsection follows easily from Proposition 27, and shows an elegant, and expected, property of noisy games. In particular, changing the payoff matrix of a game by adding any fixed constant number to all payoffs, does not modify the probability of the respective noisy game to be (inverse-) ε -misinformed (for a fixed noise pattern). This is expected, because the addition of a fixed number in the payoffs does not change the structure of the game, and, thus, the two games are considered “equivalent” in standard game theory. The proposition below includes a more complex version of this statement, showing that the same is true for the noise pattern: adding a fixed amount of bias across the board does not modify the respective probabilities. Formally:

Proposition 28. Consider a noisy game $mG \sim G^0 + \mathcal{N}(M, D)$, and constant numbers $a_G, a_r, a_c \in \mathbb{R}$. Set $\overline{G^0} = G^0 + [a_G]$ and $\overline{M} = (\overline{M}^r; \overline{M}^c)$, where $\overline{M}^x = M^x + [a_x]$ for $x \in \{r, c\}$. Moreover, set $\overline{mG} \sim \overline{G^0} + \mathcal{N}(\overline{M}, D)$. Then:

- $\mathcal{P}[mG : \varepsilon\text{-misinformed}] = \mathcal{P}[\overline{mG} : \varepsilon\text{-misinformed}]$
- $\mathcal{P}[mG : \text{inverse-}\varepsilon\text{-misinformed}] = \mathcal{P}[\overline{mG} : \text{inverse-}\varepsilon\text{-misinformed}]$

Proof. Take any $x \in \{r, c\}$. Set $b_x = -a_G - a_x$. We observe that $G^0 + M^x = \overline{G^0} + \overline{M^x} + [b_x]$. Thus, by Proposition 27, we get, for $x \in \{r, c\}$:

- For any $i \in \{1, 2\}$,

$$\mathcal{P} \left[\text{OP}_x^{G^x}(i) \right] = \mathcal{P} \left[\text{OP}_x^{\overline{G^x}}(i) \right]$$

- For any $0 \leq \omega_1 \leq \omega_2 \leq 1$,

$$\mathcal{P} \left[\text{ROM}_x^{G^x}(\omega_1, \omega_2) \right] = \mathcal{P} \left[\text{ROM}_x^{\overline{G^x}}(\omega_1, \omega_2) \right]$$

- For any $0 \leq \omega_1 \leq \omega_2 \leq 1$,

$$\mathcal{P} \left[\text{RPM}_x^{G^x}(\omega_1, \omega_2) \right] = \mathcal{P} \left[\text{RPM}_x^{\overline{G^x}}(\omega_1, \omega_2) \right]$$

In addition, game theoretic results tell us that $NE(G^0) = NE(\overline{G^0})$. Combining the above with Theorems 7, 8 and Table 7.5, the result follows directly. \square

7.4.3 Effect of modifying noise intensity (D)

Given a misinformation game $mG \sim G^0 + \mathcal{N}(M, D)$ the distortion in the behavior of the agents regarding the limit cases in the deviation d of the noise. Initially, we prove the effect of d in two independent random variables that follow the normal distribution. With that in hand, we can provide the results regarding the probabilities of behavioural consistency for mG for $d \rightarrow \{0, +\infty\}$.

Lemma 13. *Take a random variable X such that $X \sim \mathcal{N}(\mu, d)$. Then, the following hold:*

$$\lim_{d \rightarrow 0} F(x) = \begin{cases} 0 & , \text{ when } x < \mu \\ 0.5 & , \text{ when } x = \mu \\ 1 & , \text{ when } x > \mu \end{cases}$$

$$\lim_{d \rightarrow +\infty} F(x) = 0.5$$

$$\lim_{d \rightarrow 0} f(x) = \begin{cases} 0 & , \text{ when } x \neq \mu \\ +\infty & , \text{ when } x = \mu \end{cases}$$

, and

$$\lim_{d \rightarrow +\infty} f(x) = 0$$

Proof. For the first and second equation, the result follows from equation B.2 from Subsection B.1, by observing the value of the quantity $\lim_{d \rightarrow 0} \frac{x - \mu}{d}$, for the various cases related to μ , given that $d > 0$.

For the third and fourth equation, we observe, using equation B.2 from Subsection ??, that, for $x = \mu$, we get that $f(x) = f(\mu) = \frac{1}{d\sqrt{2\pi}}$, so the result is direct for $x = \mu$. For $x \neq \mu$, the following hold:

$$f(x) = \frac{1}{d\sqrt{2\pi}} e^{-\left(\frac{x-\mu}{d\sqrt{2}}\right)^2} = \frac{1}{\sqrt{2\pi}} e^{-\left(\log d + \left(\frac{x-\mu}{d\sqrt{2}}\right)^2\right)} = \frac{1}{\sqrt{2\pi}} e^{-\left(\log d \left(1 + \frac{(x-\mu)^2}{2} \frac{1}{d^2 \log d}\right)\right)}$$

We also observe that, using L' Hospital's rule, the following can be shown (given that $d > 0$):

$$\lim_{d \rightarrow 0^+} \frac{1}{d^2 \log d} = \lim_{d \rightarrow 0^+} \frac{d^{-2}}{\log d} = \lim_{d \rightarrow 0^+} \frac{-2d^{-3}}{d^{-1}} = \lim_{d \rightarrow 0^+} -2d^{-2} = -\infty$$

Using the latter, we conclude that:

$$\lim_{d \rightarrow 0^+} f(x) = \lim_{d \rightarrow 0^+} \frac{1}{\sqrt{2\pi}} e^{-\left(\log d \left(1 + \frac{(x-\mu)^2}{2} \frac{1}{d^2 \log d}\right)\right)} = 0$$

when $x \neq \mu$, which concludes the proof. \square

Lemma 14. Take a random variable X such that $X \sim \mathcal{N}(\mu, d)$. Consider also some $\alpha, \beta \in \mathbb{R} \cup \{-\infty, +\infty\}$ such that $\alpha \leq \beta$. Then, the following hold:

$$\lim_{d \rightarrow 0} \left(\int_{\alpha}^{\beta} f(x) dx \right) = \begin{cases} 1 & , \text{ when } \alpha < \mu < \beta \\ 0.5 & , \text{ when } \alpha = \mu < \beta \\ & \text{ or } \alpha < \mu = \beta \\ 0 & , \text{ otherwise} \end{cases}$$

, and

$$\lim_{d \rightarrow +\infty} \left(\int_{\alpha}^{\beta} f(x) dx \right) = 0$$

Proof. We observe that $\int_{\alpha}^{\beta} f(x) dx = F(\beta) - F(\alpha)$. The results now follow from Lemma 13. \square

Lemma 15. Take two independent random variables X, Y such that $X \sim \mathcal{N}(\mu_X, d_X)$, $Y \sim \mathcal{N}(\mu_Y, d_Y)$, where $\mu_X \neq 0$ or $\mu_Y \neq 0$. Set $Z = \frac{X}{Y}$. Then:

1. When $\mu_Y \neq 0$:

$$\lim_{d_X \rightarrow 0} \lim_{d_Y \rightarrow 0} F_Z(z) = \begin{cases} 0 & , \text{ when } z < \frac{\mu_X}{\mu_Y} \\ 0.5 & , \text{ when } z = \frac{\mu_X}{\mu_Y} \\ 1 & , \text{ when } z > \frac{\mu_X}{\mu_Y} \end{cases}$$

2. When $\mu_Y = 0, \mu_X \neq 0$:

$$\lim_{d_X \rightarrow 0} \lim_{d_Y \rightarrow 0} F_Z(z) = 0.5$$

Proof. For any $z \in \mathbb{R}$ the following holds:

$$F_Z(z) = \mathcal{P} \left[\frac{X}{Y} \leq z \right] = \mathcal{P} \left[\frac{X - zY}{Y} \leq 0 \right]$$

Now set $W = X - zY$. By the above, we get that:

$$F_Z(z) = \mathcal{P} [W \leq 0 \wedge Y > 0] + \mathcal{P} [W \geq 0 \wedge Y < 0]$$

We also note that $W \sim \mathcal{N}(\mu_X - z\mu_Y, d_X + z^2 d_Y)$. Therefore: $\lim_{d_X \rightarrow 0} \lim_{d_Y \rightarrow 0} F_W(w) = \lim_{d_W \rightarrow 0} F_W(w)$, for which Lemma 13 applies.

Now we will consider various cases.

Case 1: $\mu_Y > 0$.

We observe that:

$$0 \leq \mathcal{P} [W \geq 0 \wedge Y < 0] \leq \mathcal{P} [Y < 0]$$

and

$$\lim_{d_X \rightarrow 0} \lim_{d_Y \rightarrow 0} \mathcal{P} [Y < 0] = 0$$

Therefore,

$$\lim_{d_X \rightarrow 0} \lim_{d_Y \rightarrow 0} \mathcal{P} [W \geq 0 \wedge Y < 0] = 0$$

Moreover,

$$\mathcal{P} [W \leq 0 \wedge Y > 0] = \mathcal{P} [W \leq 0] + \mathcal{P} [Y > 0] - \mathcal{P} [W \leq 0 \vee Y > 0]$$

and

$$1 \geq \mathcal{P} [W \leq 0 \vee Y > 0] \geq \mathcal{P} [Y > 0]$$

and

$$\lim_{d_X \rightarrow 0} \lim_{d_Y \rightarrow 0} \mathcal{P} [Y > 0] = 1$$

Therefore,

$$\lim_{d_X \rightarrow 0} \lim_{d_Y \rightarrow 0} \mathcal{P} [W \leq 0 \vee Y > 0] = 1$$

Hence,

$$\lim_{d_X \rightarrow 0} \lim_{d_Y \rightarrow 0} \mathcal{P} [W \leq 0 \bigwedge Y > 0] = \lim_{d_X \rightarrow 0} \lim_{d_Y \rightarrow 0} \mathcal{P} [W \leq 0]$$

Combining the above results we get that

$$\lim_{d_X \rightarrow 0} \lim_{d_Y \rightarrow 0} F_Z(z) = \lim_{d_X \rightarrow 0} \lim_{d_Y \rightarrow 0} \mathcal{P} [W \leq 0]$$

Consequently,

$$\lim_{d_X \rightarrow 0} \lim_{d_Y \rightarrow 0} F_Z(z) = \lim_{d_X \rightarrow 0} \lim_{d_Y \rightarrow 0} F_W(0)$$

Therefore, by Lemma 13:

$$\lim_{d_X \rightarrow 0} \lim_{d_Y \rightarrow 0} F_Z(z) = \begin{cases} 0 & , \text{ when } 0 < \mu_X - z\mu_Y \\ 0.5 & , \text{ when } 0 = \mu_X - z\mu_Y \\ 1 & , \text{ when } 0 > \mu_X - z\mu_Y \end{cases}$$

Considering also the fact that $\mu_Y > 0$, the result follows easily for this case.

Case 2: $\mu_Y < 0$.

Set $Y' = -Y$. Then, $Y' \sim \mathcal{N}(-\mu_Y, d_Y)$, i.e., $\mu_{Y'} = -\mu_Y > 0$. Moreover, set $Z' = \frac{X}{Y'}$. For Z' , case 1 applies, thus:

$$\lim_{d_X \rightarrow 0} \lim_{d_Y \rightarrow 0} F_{Z'}(z) = \begin{cases} 0 & , \text{ when } z < \frac{\mu_X}{\mu_{Y'}} \\ 0.5 & , \text{ when } z = \frac{\mu_X}{\mu_{Y'}} \\ 1 & , \text{ when } z > \frac{\mu_X}{\mu_{Y'}} \end{cases}$$

Equivalently,

$$\lim_{d_X \rightarrow 0} \lim_{d_Y \rightarrow 0} F_{Z'}(z) = \begin{cases} 0 & , \text{ when } z < -\frac{\mu_X}{\mu_Y} \\ 0.5 & , \text{ when } z = -\frac{\mu_X}{\mu_Y} \\ 1 & , \text{ when } z > -\frac{\mu_X}{\mu_Y} \end{cases}$$

Furthermore:

$$\begin{aligned} F_Z(z) &= \mathcal{P} \left[\frac{X}{Y} \leq z \right] = \mathcal{P} \left[\frac{X}{-Y'} \leq z \right] = \mathcal{P} \left[\frac{X}{Y'} \geq -z \right] \\ &= 1 - \mathcal{P} \left[\frac{X}{Y'} \leq -z \right] = 1 - \mathcal{P} [Z' \leq -z] = 1 - F_{Z'}(-z) \end{aligned}$$

Combining the above equations, the result for this case follows easily. This concludes the proof for the scenario where $\mu_Y \neq 0$.

Case 3: $\mu_Y = 0, \mu_X > 0$.

We observe that, in this case, $\mu_W = \mu_X > 0$, and thus:

$$\lim_{d_X \rightarrow 0} \lim_{d_Y \rightarrow 0} \mathcal{P}[W \leq 0] = 0, \quad \lim_{d_X \rightarrow 0} \lim_{d_Y \rightarrow 0} \mathcal{P}[W \geq 0] = 1$$

Using analogous arguments with case 1, we get:

$$0 \leq \mathcal{P}[W \leq 0 \bigwedge Y > 0] \leq \mathcal{P}[W \leq 0]$$

and

$$\lim_{d_X \rightarrow 0} \lim_{d_Y \rightarrow 0} \mathcal{P}[W \leq 0] = 0$$

Therefore,

$$\lim_{d_X \rightarrow 0} \lim_{d_Y \rightarrow 0} \mathcal{P}[W \leq 0 \bigwedge Y > 0] = 0$$

Moreover,

$$\mathcal{P}[W \geq 0 \bigwedge Y < 0] = \mathcal{P}[W \geq 0] + \mathcal{P}[Y < 0] - \mathcal{P}[W \geq 0 \bigvee Y < 0]$$

and

$$1 \geq \mathcal{P}[W \geq 0 \bigvee Y < 0] \geq \mathcal{P}[W \geq 0]$$

and

$$\lim_{d_X \rightarrow 0} \lim_{d_Y \rightarrow 0} \mathcal{P}[W \geq 0] = 1$$

Therefore,

$$\lim_{d_X \rightarrow 0} \lim_{d_Y \rightarrow 0} \mathcal{P}[W \geq 0 \bigvee Y < 0] = 1$$

Hence,

$$\lim_{d_X \rightarrow 0} \lim_{d_Y \rightarrow 0} \mathcal{P}[W \geq 0 \bigwedge Y < 0] = \lim_{d_X \rightarrow 0} \lim_{d_Y \rightarrow 0} \mathcal{P}[Y < 0]$$

Combining the above results with Lemma 13 and the fact that $\mu_Y = 0$ we get that, for all $z \in \mathbb{R}$:

$$\lim_{d_X \rightarrow 0} \lim_{d_Y \rightarrow 0} F_Z(z) = \lim_{d_X \rightarrow 0} \lim_{d_Y \rightarrow 0} \mathcal{P}[Y < 0] = 0.5$$

This proves the result for this case.

Case 4: $\mu_Y = 0$, $\mu_X < 0$.

Set $X' = -X$. Then, $X' \sim \mathcal{N}(-\mu_X, d_X)$, i.e., $\mu_{X'} = -\mu_X > 0$. Moreover, set $Z' = \frac{X'}{Y}$. For Z' , case 3 applies, thus, for all $z \in \mathbb{R}$:

$$\lim_{d_X \rightarrow 0} \lim_{d_Y \rightarrow 0} F_{Z'}(z) = 0.5$$

Furthermore, for all $z \in \mathbb{R}$:

$$\begin{aligned} F_Z(z) &= \mathcal{P} \left[\frac{X}{Y} \leq z \right] = \mathcal{P} \left[\frac{-X'}{Y} \leq z \right] = \mathcal{P} \left[\frac{X'}{Y} \geq -z \right] = \\ &= 1 - \mathcal{P} \left[\frac{X'}{Y} \leq -z \right] = 1 - \mathcal{P} [Z' \leq -z] = 1 - F_{Z'}(-z) \end{aligned}$$

Consequently,

$$\lim_{d_X \rightarrow 0} \lim_{d_Y \rightarrow 0} F_Z(z) = 1 - \lim_{d_X \rightarrow 0} \lim_{d_Y \rightarrow 0} F_{Z'}(-z) = 0.5$$

This concludes the proof. \square

Lemma 16. Take two independent random variables X, Y such that $X \sim \mathcal{N}(\mu_X, d_X)$, $Y \sim \mathcal{N}(\mu_Y, d_Y)$, where $\mu_X \neq 0$ or $\mu_Y \neq 0$. Consider also some $\Omega_1, \Omega_2 \in \mathbb{R} \cup \{-\infty, +\infty\}$ such that $\Omega_1 < \Omega_2$. Then:

1. When $\mu_Y \neq 0$:

$$\lim_{d_X \rightarrow 0} \lim_{d_Y \rightarrow 0} \mathcal{P} \left[\Omega_1 \leq \frac{X}{Y} \leq \Omega_2 \right] = \begin{cases} 1 & , \text{ when } \Omega_1 < \frac{\mu_X}{\mu_Y} < \Omega_2 \\ 0.5 & , \text{ when } \frac{\mu_X}{\mu_Y} \in \{\Omega_1, \Omega_2\} \\ 0 & , \text{ otherwise} \end{cases}$$

2. When $\mu_Y = 0, \mu_X \neq 0$:

$$\lim_{d_X \rightarrow 0} \lim_{d_Y \rightarrow 0} \mathcal{P} \left[\Omega_1 \leq \frac{X}{Y} \leq \Omega_2 \right] = \begin{cases} 1 & , \text{ when } \Omega_1 = -\infty, \Omega_2 = +\infty \\ 0.5 & , \text{ when } \Omega_1 = -\infty, \Omega_2 \neq +\infty \\ & , \text{ or } \Omega_1 \neq -\infty, \Omega_2 = +\infty \\ 0 & , \text{ otherwise} \end{cases}$$

Proof. Set $Z = \frac{X}{Y}$. We observe that:

$$\mathcal{P} \left[\Omega_1 \leq \frac{X}{Y} \leq \Omega_2 \right] = F_Z(\Omega_2) - F_Z(\Omega_1)$$

The result now follows by considering the various cases and applying Lemma 15 as appropriate. \square

Lemma 17. Take two independent random variables X, Y such that $X \sim \mathcal{N}(\mu_X, d_X)$, $Y \sim \mathcal{N}(\mu_Y, d_Y)$, where $\mu_X \neq 0$ and $\mu_Y \neq 0$. Consider also some $\Omega_1, \Omega_2 \in \mathbb{R} \cup \{-\infty\}$ such that $-\infty \leq \Omega_1 < \Omega_2 \leq 0$. Set:

$$I(X, Y) = \int_0^{+\infty} \left(\int_{\Omega_1 y}^{\Omega_2 y} f_X(x) dx \right) \frac{f_Y(y)}{y} dy$$

Then, the following hold:

$$\lim_{d_X \rightarrow 0} \lim_{d_Y \rightarrow 0} I(X, Y) = \begin{cases} 1 & , \text{ when } \mu_X < 0 \text{ and } \mu_Y > 0 \text{ and } \Omega_1 < \frac{\mu_X}{\mu_Y} < \Omega_2 \\ 0.5 & , \text{ when } \mu_X < 0 \text{ and } \mu_Y > 0 \text{ and } \frac{\mu_X}{\mu_Y} \in \{\Omega_1, \Omega_2\} \\ 0 & , \text{ otherwise} \end{cases}$$

, and

$$\lim_{d_X \rightarrow +\infty} \lim_{d_Y \rightarrow +\infty} I(X, Y) = 0.$$

Proof. By Lemma 12, it follows that:

$$I = \mathcal{P} \left[\Omega_1 \leq \frac{X}{Y} \leq \Omega_2, X < 0, Y > 0 \right]$$

When $\mu_X > 0$:

$$0 \leq \mathcal{P} \left[\Omega_1 \leq \frac{X}{Y} \leq \Omega_2, X < 0, Y > 0 \right] \leq \mathcal{P} [X < 0]$$

Then,

$$0 \leq \lim_{d_X \rightarrow 0} \lim_{d_Y \rightarrow 0} \mathcal{P} \left[\Omega_1 \leq \frac{X}{Y} \leq \Omega_2, X < 0, Y > 0 \right] \leq \lim_{d_X \rightarrow 0} \lim_{d_Y \rightarrow 0} \mathcal{P} [X < 0] = 0$$

Analogously, when $\mu_Y < 0$:

$$0 \leq \mathcal{P} \left[\Omega_1 \leq \frac{X}{Y} \leq \Omega_2, X < 0, Y > 0 \right] \leq \mathcal{P} [Y > 0]$$

Then,

$$0 \leq \lim_{d_X \rightarrow 0} \lim_{d_Y \rightarrow 0} \mathcal{P} \left[\Omega_1 \leq \frac{X}{Y} \leq \Omega_2, X < 0, Y > 0 \right] \leq \lim_{d_X \rightarrow 0} \lim_{d_Y \rightarrow 0} \mathcal{P} [Y > 0] = 0$$

When $\mu_X < 0, \mu_Y > 0$, we observe that:

$$\lim_{d_X \rightarrow 0} \lim_{d_Y \rightarrow 0} \mathcal{P} [X < 0, Y > 0] = \lim_{d_X \rightarrow 0} \lim_{d_Y \rightarrow 0} \mathcal{P} [X < 0] \mathcal{P} [Y > 0] = 1$$

and

$$\lim_{d_X \rightarrow 0} \lim_{d_Y \rightarrow 0} \mathcal{P} \left[\Omega_1 \leq \frac{X}{Y} \leq \Omega_2 \bigvee (X < 0, Y > 0) \right] \geq \lim_{d_X \rightarrow 0} \lim_{d_Y \rightarrow 0} \mathcal{P} [X < 0, Y > 0] = 1$$

Therefore:

$$\begin{aligned}
\lim_{d_X \rightarrow 0} \lim_{d_Y \rightarrow 0} I(X, Y) &= \lim_{d_X \rightarrow 0} \lim_{d_Y \rightarrow 0} \mathcal{P} \left[\Omega_1 \leq \frac{X}{Y} \leq \Omega_2, X < 0, Y > 0 \right] \\
&= \lim_{d_X \rightarrow 0} \lim_{d_Y \rightarrow 0} \mathcal{P} \left[\Omega_1 \leq \frac{X}{Y} \leq \Omega_2 \right] + \mathcal{P}[X < 0, Y > 0] \\
&\quad - \mathcal{P} \left[\Omega_1 \leq \frac{X}{Y} \leq \Omega_2 \vee (X < 0, Y > 0) \right] \\
&= \lim_{d_X \rightarrow 0} \lim_{d_Y \rightarrow 0} \mathcal{P} \left[\Omega_1 \leq \frac{X}{Y} \leq \Omega_2 \right]
\end{aligned}$$

Combining the above results with Lemma 16, and by considering the various cases on μ_X , μ_Y and $\frac{\mu_X}{\mu_Y}$, the result follows. \square

Although adding a fixed constant number to the game's payoffs does not modify the respective probabilities (Proposition 28), this is not the case when changing the “scale” of a game (by multiplying all its payoffs by a constant number, say $\lambda > 0$). In particular, changing the scale of a game will affect its “resilience” to noise, without changing the game's properties and behaviour, because it increases the “amount of noise” necessary to change the sign of the various $\text{UGAIN}(y, x)_i$. As a matter of fact, multiplying the payoffs by a sufficiently large number would minimize the effect of the noise, as its effects on the payoffs would be, comparatively smaller (analogously, using a sufficiently small positive number would maximize the effect of the noise).

In Proposition 29 (and especially in Corollary 8), we quantify this effect, by showing that we need to multiply the noise intensity (standard deviation) by λ^2 in order for the noise to have the same effect on a game scaled by λ . Formally:

Proposition 29. *Consider a normal noisy game $mG \sim G^0 + \mathcal{N}(M, D)$ and some $\lambda > 0$. Set: $\overline{G^0} = \lambda G^0$, $\overline{M} = \lambda M$ and $\overline{D} = \lambda^2 D$, and consider the normal noisy game $\overline{mG} = \overline{G^0} + \mathcal{N}(\overline{M}, \overline{D})$. Then:*

- $\mathcal{P}[mG : \varepsilon\text{-misinformed}] = \mathcal{P}[\overline{mG} : \varepsilon\text{-misinformed}]$
- $\mathcal{P}[mG : \text{inverse-}\varepsilon\text{-misinformed}] = \mathcal{P}[\overline{mG} : \text{inverse-}\varepsilon\text{-misinformed}]$

An obvious and interesting corollary of Propositions 28 and 29 is the following:

Corollary 8. *Consider a normal noisy game $mG \sim G^0 + \mathcal{N}([0], D)$ and some $\lambda > 0$, $k \in \mathbb{R}$. Set: $\overline{G^0} = \lambda G^0 + k$ and $\overline{D} = \lambda^2 D$, and consider the normal noisy game $\overline{mG} = \overline{G^0} + \mathcal{N}([0], \overline{D})$. Then:*

- $\mathcal{P}[mG : \varepsilon\text{-misinformed}] = \mathcal{P}[\overline{mG} : \varepsilon\text{-misinformed}]$
- $\mathcal{P}[mG : \text{inverse-}\varepsilon\text{-misinformed}] = \mathcal{P}[\overline{mG} : \text{inverse-}\varepsilon\text{-misinformed}]$

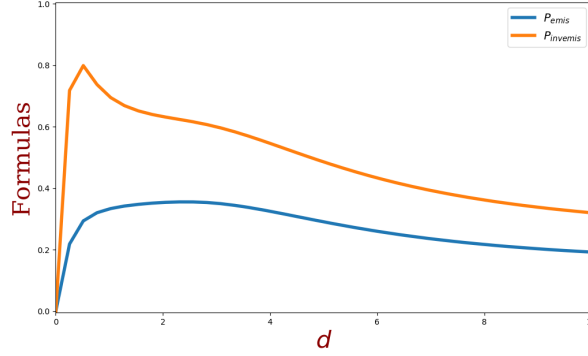


Figure 7.2: Example showing that $\mathcal{P}[mG : \varepsilon\text{-misinformed}]$ (blue), $\mathcal{P}[mG : \text{inverse-}\varepsilon\text{-misinformed}]$ (orange) (vertical axis) have non-monotonic dependence to noise (horizontal axis).

From the previous theoretical results, the effect of noise in the outcome of an abstract 2×2 bimatrix game has the following characteristics: for small values of the noise intensity (standard deviation), players almost surely have the same behaviour as in the actual game, whereas for large noise intensity, the behaviour of players cannot be predicted as their games will be almost random. Also, observe that the formulas giving the probabilities for (inverse)- ε -misinformed are continuous with respect to the standard deviation. Given the above, one would expect that, by increasing the standard deviation, we would monotonically transit from the first extreme to the second. However, this does not always hold, as the following counter-example shows.

Example 7.2. Consider as actual game the classical Prisoner's Dilemma (see Figure 7.3a), which has a pure Nash equilibrium with strategy profile $((0, 1), (0, 1))$. We produce a noisy game, in which the noise only affects the upper left elements of the actual payoff matrix, where we add noise according to a random variable following the normal distribution $\mathcal{N}(0, d^2)$. From Theorems 7, 8, we can compute the probabilities for this game to be (inverse)- ε -misinformed. The result is shown in Figure 7.2, where we plot $\mathcal{P}[mG : \varepsilon\text{-misinformed}]$ (blue line) and $\mathcal{P}[mG : \text{inverse-}\varepsilon\text{-misinformed}]$ (orange line), for $d \in (0, 10)$. As is obvious by this figure, these functions are not monotonic with respect to d .

7.5 Discussion and Experiments

In this section we report on experiments that validate our basic results, and we investigate the effect of noise on the players' decisions, for the four 2×2 bimatrix games shown in Figure ???. The games were chosen to capture the following cases: i) dominant equilibrium (Prisoner's

Dilemma), ii) unique mixed Nash equilibrium (Matching Pennies), iii) multiple Nash equilibria (Battle of the Sexes), and iv) dominant equilibrium that coincides with the optimal outcome (Win-Win).

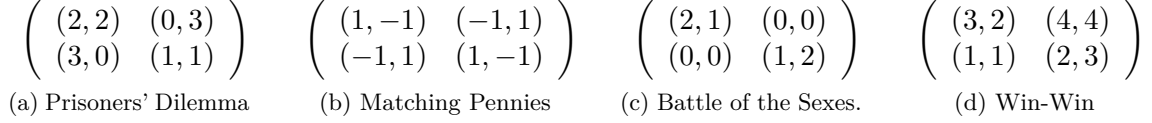


Figure 7.3: Test cases.

7.5.1 Theoretical and Experimental Computation of the Probability that a Game is (Inverse-) ε -misinformed

We consider that the actual game undergoes an additive noise that follows the normal distribution $\mathcal{N}([0], [d^2])$ where $d \in \{0.001, 0.5, 1, \dots, 10\}$. We compare the theoretical values of probabilities that we get from Theorems 7, 8, with the respective values calculated through Monte Carlo simulations. The Monte Carlo simulations were conducted as follows: we generate a game G^0 , which can be one of the four games shown in Figure 7.3. Then, for each of the above values for d , we create the respective noisy game $mG = G^0 + \mathcal{N}([0], [d^2])$. To be more precise, we generate a misinformation game, where the misinformation stems from the incorporation of additive noise stemming from one random experiment that follows the above distribution ($\mathcal{N}([0], [d^2])$). We derive the natural misinformed equilibrium and check about ε -closeness. We perform 3,000 repetitions of the above process and calculate:

- a) the percentage of games that are ε -misinformed (i.e., all nmes of mG are ε -close to one Nash equilibrium of G^0 , according to the first bullet of Definition 54),
- b) the percentage of games that are inverse ε -misinformed (i.e., all Nash equilibria of G^0 are ε -close to one nme of mG , according to the second bullet of Definition 54).

We repeat the simulations for two different values of ε ($\varepsilon \in \{10^{-2}, 10^{-3}\}$). The results are shown in Figures 7.4-7.6. As the Prisoner's Dilemma and the Win-Win games both have a unique pure Nash equilibrium, their behavioural consistency is similar. Hence, Figure 7.4 shows both cases. In all subplots, we have plots of two colours. The blue ones depict the computations for $\varepsilon = 10^{-2}$, whereas the red ones depict the computations for $\varepsilon = 10^{-3}$.

In part (a) of the figures, the horizontal axis depicts the different values for the standard deviation d of the noise, and the vertical axis depicts the probability of a game being ε -misinformed according to Theorem 7 (solid line) or the probability of a game being ε -misinformed according to the Monte Carlo simulations (dotted lines). The same hold for part (b) of the figures, but for the inverse- ε -misinformed case (Theorem 8). In both subfigures, a

high value of the probability calculated in the vertical axis implies a small effect of noise on players' decisions. As expected, the theoretical results are very close to the experimental ones.

These figures give rise to various remarks concerning the influence of noise to players' strategic choices. Although some general patterns emerge, the effect of noise in the behavioural consistency of the game greatly depends on the type and number of Nash equilibria that it has, so we split our analysis in 3 different cases.

Case 1: Unique Pure Nash equilibrium

The case of a unique pure Nash equilibrium appears in the Prisoner's Dilemma and Win-Win games, whose behaviour is depicted in Figure 7.4.

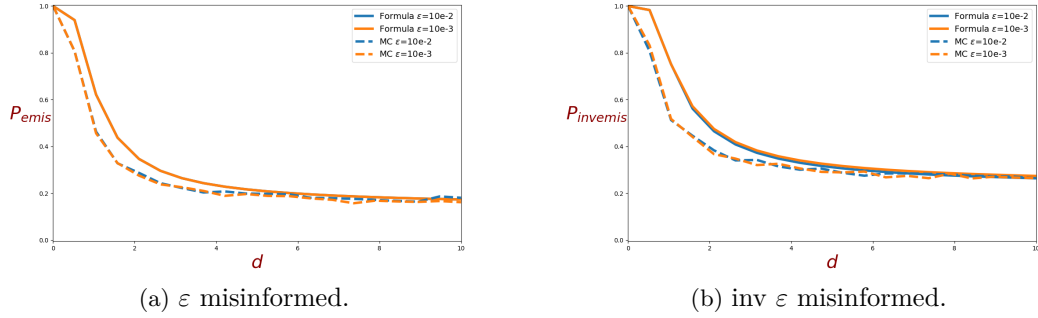


Figure 7.4: Monte Carlo (MC) simulation and probabilistic formulas for Prisoner's Dilemma and Win-Win games. Vertical axis: (7.4a) $\mathcal{P}[mG : \varepsilon\text{-misinformed}]$, (7.4b) $\mathcal{P}[mG : \text{inverse-}\varepsilon\text{-misinformed}]$. Horizontal axis: noise intensity d .

For Prisoner's Dilemma, we observe that, for small values of the standard deviation ($d \ll 1$), the nme of mG will usually be the same as the NE of the original game (G^0). Thus, both probabilities $\mathcal{P}_{emis}[mG; \varepsilon]$ and $\mathcal{P}_{invemis}[mG; \varepsilon]$ will have values close to 1. As d increases, noise will produce misinformation games G^r, G^c with different Nash equilibria than that of the actual game (different means non-close, by definition, in this case) with an increasing probability, thereby reducing the probability for behavioural consistency.

As d increases further, each of the different possible sets of equilibria will appear with almost equal probability in G^r, G^c , leading to a convergence in the plots of Figure 7.4. In particular, $\mathcal{P}_{emis}[mG; \varepsilon]$ converges to approximately 14%, whereas $\mathcal{P}_{invemis}[mG; \varepsilon]$ converges to approximately 25%. This can be theoretically predicted by observing Table 7.1. For a large enough noise, the original orderings among the elements of the payoff matrix become increasingly irrelevant, and the actual orderings in each of G^r, G^c become totally random. As a result, the equilibrium strategy for r in G^r will be a pure one with a probability of

6/8 (3/8 for each strategy), a mixed one with 1/8 probability, and pure-and-mixed with 1/8 probability. The same is true of course for c in G^c . Combining these observations with Table 7.5 and Theorems 7, 8, we get the above numbers for the convergence of $\mathcal{P}_{emis}[mG; \varepsilon]$, $\mathcal{P}_{invemis}[mG; \varepsilon]$.

Similar remarks hold for the Win-Win game that has one pure Nash equilibrium strategy profile (namely, $((1, 0), (0, 1))$).

Case 2: Unique Mixed Nash equilibrium

The case of a unique mixed Nash equilibrium appears in the Matching Pennies game, which has one Nash equilibrium strategy profile $((1/2, 1/2), (1/2, 1/2))$, and whose behaviour is depicted in Figure 7.5.

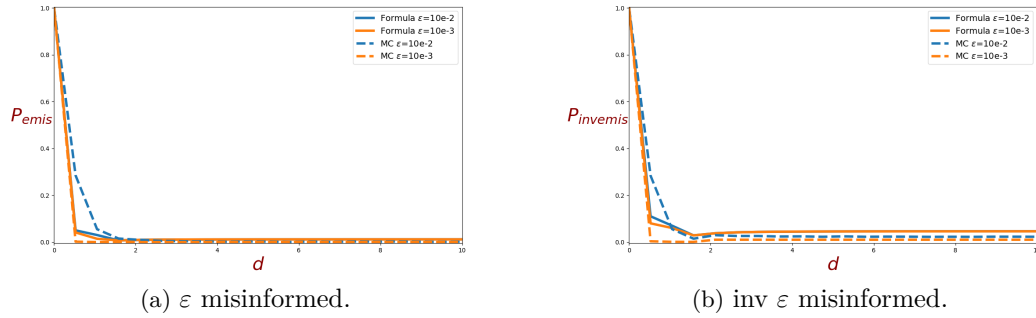


Figure 7.5: Monte Carlo (MC) simulation and probabilistic formulas for Matching Pennies. Vertical axis: (7.5a) $\mathcal{P}[mG : \varepsilon\text{-misinformed}]$, (7.5b) $\mathcal{P}[mG : \text{inverse-}\varepsilon\text{-misinformed}]$. Horizontal axis: noise intensity d .

As in case 1, we observe that for small values of the standard deviation ($d \ll 1$), mG will have the same nme as the NE in G^0 . Thus, both probabilities $\mathcal{P}_{emis}[mG; \varepsilon]$ and $\mathcal{P}_{invemis}[mG; \varepsilon]$ will have values close to 1. As d increases, noise will produce games G^r, G^c with different Nash equilibria than that of the actual game G^0 , and the respective probabilities fall sharply (much faster compared to the Prisoner's Dilemma case), converging to a value close to 0 for large values of the standard deviation. This is explained by the fact that, although a mixed nme is achieved in some of the produced games, this is often not close to the actual mixed one, leading to games that are (usually) not (inverse-) ε -misinformed. For example, for $\varepsilon = 10^{-2}$, the function $\mathcal{P}_{emis}[mG; \varepsilon]$ converges at around 0.03%.

Case 3: Multiple Nash equilibria

The case of two pure and one mixed Nash equilibrium appears in the Battle of the Sexes game, whose behaviour is depicted in Figure 7.6. The Nash equilibrium strategy profiles of Battle of the Sexes are: $\{((1, 0), (1, 0)), ((0, 1), (0, 1)), ((2/3, 1/3), (1/3, 2/3))\}$.

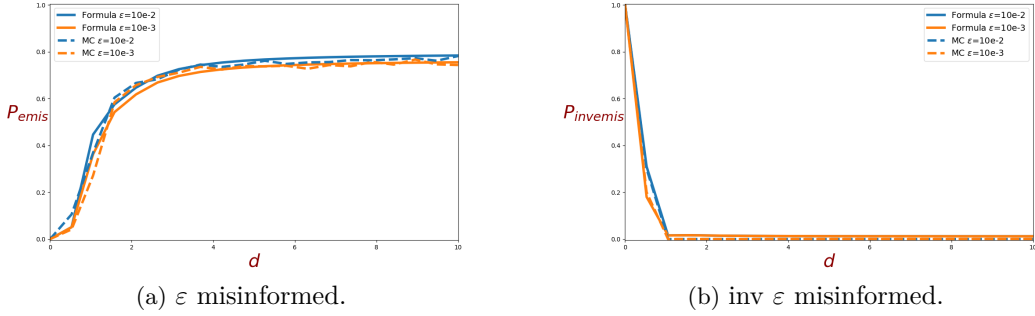


Figure 7.6: Monte Carlo (MC) simulation and probabilistic formulas for Battle of the Sexes. Vertical axis: $\mathcal{P}[mG : \varepsilon\text{-misinformed}]$, $\mathcal{P}[mG : \text{inverse-}\varepsilon\text{-misinformed}]$. Horizontal axis: noise intensity d .

Unlike other games, we observe that the Battle of the Sexes has zero probability of being ε -misinformed for small values of d . This is explained by the fact that, for small values of d , G^r, G^c will be very similar to G^0 , each giving 3 equilibrium strategies (for the respective player). Thus, there are 9 nmes, one for each combination of equilibrium strategies (see Definition ??), so some of them will not be ε -close to one of the three equilibria of G^0 . By Definition 54 this means that the respective game is not ε -misinformed, so $\mathcal{P}_{emis}[mG; \varepsilon]$ will be close to 0.

As d increases, and the games G^r, G^c become less and less predictable, the probability of being ε -misinformed becomes larger, reaching a plateau at around 72%. The explanation here is analogous to the one given for the other two cases: in order for a misinformation game to *not* be ε -misinformed, it should either have one pure equilibrium (but not one of the two that are in the equilibria of G^0), or it should have one mixed equilibrium (but not ε -close to the one of G^0). Based on the analysis of the Prisoner's Dilemma game, the probability of the former is around 28%; based on the analysis of the Matching Pennies game, the probability of the latter is close to 0; combining these observations, we conclude that a plateau at around 72% is reasonable.

For the inverse- ε -misinformed case (part (b) of Figure 7.6), small values of d result to high values for $\mathcal{P}_{invemis}[mG; \varepsilon]$, as expected. As d increases, the probability decreases at a rate even faster than the one observed for Matching Pennies, eventually converging at a value close to 0. This is explained by the fact that, in order for the game to be inverse- ε -misinformed,

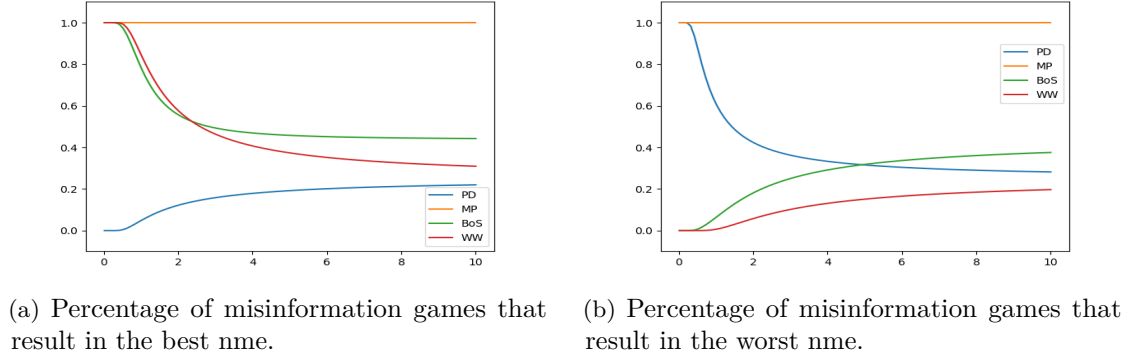


Figure 7.7: Strategy profiles in terms of efficiency.

it should have, among other things, also a mixed equilibrium that is close to the respective mixed of G^0 . As we established in Case 2 above, this has a very low probability for large values of d .

7.5.2 Optimal strategy profiles in terms of efficiency

In this subsection, we report on experiments that investigate whether the misinformation game mG that results from a given actual game G^0 has natural misinformed equilibria that are best or worst in terms of efficiency (social welfare). We then evaluate the effect of noise on each of the four games under consideration.

We performed Monte Carlo simulations as in the previous section and calculated:

- the percentage P_{best} of misinformation games that have a natural misinformed equilibrium that maximizes social welfare (best nme),
- the percentage P_{worst} of misinformation games that have a natural misinformed equilibrium that minimizes social welfare (worst nme).

We repeat the simulations for all values of d in $\{0.02, 0.04, \dots, 10\}$ and for $\varepsilon = 10^{-2}$. The results are shown in Figures 7.7a and 7.7b.

In Matching Pennies, as it is a constant-sum game, all strategy profiles provide the same level of social welfare, so the respective line is flat, regardless of the value of d (see Figures 7.7a and 7.7b). In other words, the noise has no effect with respect to the optimal outcome.

In Prisoners' Dilemma, the best strategy profile is $((1, 0), (1, 0))$ and the worst one is $((0, 1), (0, 1))$ which coincides with the pure NE of the actual game G^0 . We observe that, for small values of d , only a few repetitions provide the best nme (Figure 7.7a), while most of them provide the worst nme (Figure 7.7b); this is in line with the results given in the previous

subsection. As d increases, the percentage of games resulting in the best strategy increases too, implying that noise has a positive effect on Prisoners' Dilemma.

In the Battle of the Sexes, the best strategy profiles are $((1, 0), (1, 0))$ and $((0, 1), (0, 1))$ (these are also the pure Nash equilibria of the actual game), and the worst strategy profiles are $((1, 0), (0, 1))$ and $((0, 1), (1, 0))$. We observe that, for small values of d , most of the misinformation games result in one of the best strategy profiles (Figure 7.7a). As d increases, this percentage decreases, implying that noise has a negative effect on the Battle of the Sexes: players are not forced to choose better strategies.

In the Win-Win game, the best strategy profile is $((1, 0), (0, 1))$ and the worst one is $((0, 1), (1, 0))$. The same observations as in the Battle of the Sexes hold for the Win-Win game.

To summarize, as the percentage P_{best} increases (or P_{worst} decreases) with respect to d , noise is beneficial. This is the case for Prisoners' Dilemma. On the contrary, noise deteriorates the efficiency of the system if the percentage P_{best} decreases (or P_{worst} increases) with respect to d as in Win-Win and Battle of the Sexes games. Finally, the efficiency of the system is independent of the noise in the Matching Pennies game.

Given the above, as expected, noise deteriorates the social welfare in games where the original Nash equilibrium is already "good" for the social welfare (Battle of the Sexes, Win-Win), as it induces a more "random" behaviour. On the contrary, it improves the situation in games where the original equilibrium is "bad" (e.g., Prisoner's Dilemma). In constant sum games (e.g., Matching Pennies), noise has no effect with regards to the social welfare.

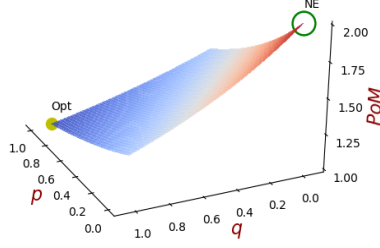
7.5.3 PoM vs PoA

In this subsection, we compare the price of anarchy PoA with the price of misinformation PoM for the four games of interest. Both metrics measure social welfare, with or without misinformation respectively, and take values that are higher than or equal to 1.

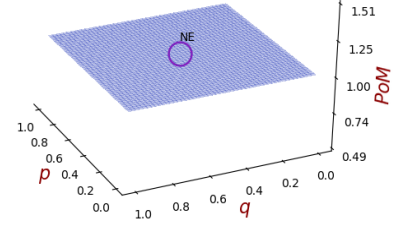
Given a bimatrix game G with payoff matrix $P = (P_r; P_c)$ we use Definition 27 to compute PoM for all values of pairs (p, q) , where $p, q \in [0, 1]$. The values of p, q are non other than the values in the joint strategy profile $\sigma = (\vec{p}, \vec{q}) = ((p, 1 - p), (q, 1 - q))$. In formula 4.2, the quantities in the fraction are given by the formula $SW(\sigma) = \mathbf{p}^T(P_r + P_c)\mathbf{q}$. The respective graphs are shown in Figures 7.8a-7.8d.

We can make the following observations on social welfare planes of Figures 7.8a-7.8d that present the range of values of PoM :

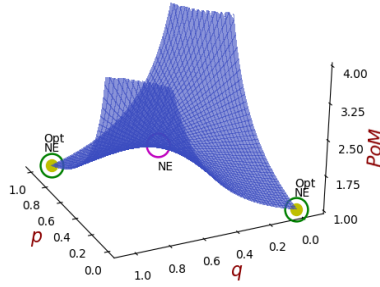
1. In Prisoner's Dilemma we note that the social welfare plane is monotonic (see Figure 7.8a). The minimum value is in the bottom left corner ("bluest") and the maximum value is in the upper right corner ("redest"). We know that the PoA in this game is 2, which is equal to the minimum social welfare, so any distortion in the payoff matrices of the game does not deteriorate the efficiency of the game, and $PoM \leq PoA$, for every level



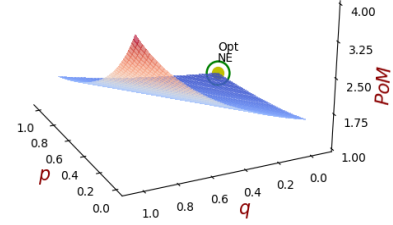
(a) PoM plane for the Prisoner's Dilemma.



(b) PoM plane for the Matching Pennies.



(c) PoM plane for the Battle of the Sexes.



(d) PoM plane for the Win-Win.

Figure 7.8: PoM plane for games in Table 7.3.

of noise.

2. In Matching Pennies we observe that the social welfare plane is constant (Figure 7.8b). That is, PoM remains constant as any combination of the values of the payoff matrix results in the same social welfare value. Thus, noise may affect the strategic behaviour of players, but keeps the social welfare constant. Note that, in zero-sum games such as Matching Pennies, the value of PoM and PoA cannot be calculated (the denominator of the respective formulas takes the value of zero). To mitigate this inconvenience we add proper values to each element of the payoff matrices and produce a constant-sum game, without affecting the strategic behaviour of players.
3. In Battle of the Sexes we observe that the two pure Nash equilibria of the game are the optimal strategic behaviours (Figure 7.8c). Thus, PoA depends on the mixed Nash equilibrium, and noise could improve or degrade the efficiency of the system.

4. In Win-Win, the unique Nash equilibrium coincides with the optimal one, thus $PoA = 1$ (Figure 7.8d). Therefore, any misinformation cannot improve the outcome of this game, and $PoM \geq PoA$.

Further numerical experiments provided in Section B.5.

Chapter 8

Mechanism Design

8.1 Introduction

In Chapter 4 proved that given a sufficient misinformation, anything is possible in terms of social welfare, Propositions 4-5 and Corollary 1. Therefore, in this chapter, we investigate methods to lead players' behaviour to a socially improved outcome. Coordination mechanisms were introduced for this purpose in [Christodoulou et al., 2009, Christodoulou et al., 2014], in which the authors propose a theoretical framework where modifications of the game lead to a reduced fraction, compared to the PoA , of the worst Nash equilibrium in the modified game to the social optimum of the original game. This has been applied to many classes of games such as load balancing and congestion games [Nisan et al., 2007b].

Clearly, misinformation could lead players to strategic choices that are different from the ones they would make in the absence of misinformation. This includes choices that are actually beneficial (from the perspective of social welfare) for the players. Inspired by this observation, we combine misinformation games and coordination mechanisms in order to deal with the following main question:

Is it possible for the designer of a game to misinform players regarding the game parameters, in order to provide incentives for a better (or even optimal) behaviour in terms of social welfare?

We positively answer this question, and provide a novel way for applying coordination mechanisms using the concept of misinformation, thereby establishing a connection between the classical coordination mechanisms and misinformation games. Towards this direction, we introduce a modification of PoM metric to measure the impact of misinformation in games compared to the socially optimum situation.

As in classical coordination mechanisms, where the designer modifies the game in order to minimize the ratio between the worst Nash equilibrium of the modified game and the social optimum of the original game, we propose a similar approach where the designer misinforms players. Next, we compare the worst natural misinformed equilibrium (i.e., the worst result of misinformation) with the social optimum of the actual game. The resulting ratio is the

Price of Misinformation, Definition 4.2. Observe that the Price of Misinformation could be less than the Price of Anarchy, resulting in an improved behaviour (from the perspective of social welfare) of the players, compared to the scenario without misinformation.

A key difference between classical coordination mechanisms and coordination mechanisms with misinformation is that in the first case the designer can determine and influence the actual interaction as a whole, whereas in the latter the designer chooses and imposes the subjective views of the players, i.e., the misinformed views. Thus, in the first case the designer modifies the actual game specification, whereas in mechanisms with misinformation the designer changes players' (subjective) information, but has no power over the actual game specification. In this chapter, we consider this setting, i.e., scenarios where the designer cannot impose a different game specification, but can misinform players about the actual set up.

We consider the problem under assumptions about the number of misinformed views that the designer can spawn. In particular, we study scenarios where the designer has bounded capabilities with regard to the number of different misinformed views that can be spawned (see Section 8.3).

We apply the above ideas for single-commodity non-atomic congestion games with n parallel links and affine cost functions. We first adapt the concept of misinformation to the class of non-atomic congestion games, (see Section 4.4). Further, we design a polynomial-time algorithm for computing a pure Nash equilibrium in a network, based on Information Theory techniques (see Section 8.2), and more specifically on the *waterfilling solution*, [Cover and Thomas, 2006, Fasoulakis et al., 2019]*. Moreover, we describe a mechanism for designing misinformation games with an optimal Price of Misinformation (and thus better social outcomes) under various assumptions, (Section 8.3).

8.2 The waterfilling algorithm

In this Section, we provide an algorithm that computes a *pure Nash equilibrium* in a single-commodity non-atomic congestion game with n parallel links, and affine latency functions, inspired by the well-known waterfilling theorem of Information theory. To the best of our knowledge, there is not a similar algorithm in the bibliography.

One of the fundamental problems in wireless communications is the problem of the allocation of a budget of power in a constant number of different quality (different noise levels) and independent wireless communication channels in order to maximize the sum of the transmission rate. The optimal solution of this problem is given by the well-known *waterfilling theorem*. Namely, the algorithm fills with water (power) the channels in a way that minimizes the maximum level of water, where the level of water is the maximum value of power plus noise in the channels that are used. At the end of the algorithm, the noise plus the water in the channels that are used is the same (see Figure 8.1 for two different waterfilling allocations

*Further, see Example B.1 in Appendix B.3.

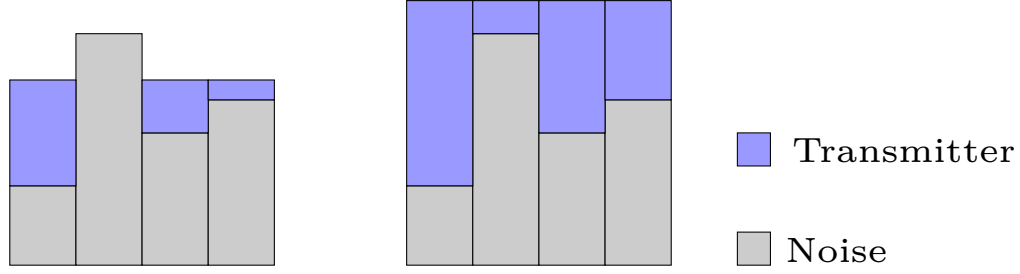


Figure 8.1: Two possible cases of waterfilling with different budget of power.

depending on the transmitter's power).

Interestingly, this idea can be used to find a Nash equilibrium flow allocation in a single-commodity non-atomic congestion game with n parallel links, and affine latency functions. Below, we give a polynomial-time algorithm based on this idea.

Algorithm 3 Waterfilling approach algorithm for computing a *pure Nash equilibrium* in single-commodity non-atomic congestion games with n parallel links and affine latency functions.

Input: n parallel links with affine latency functions $a_k x_k + b_k$, with $a_k, b_k > 0$, for any $k \in \{1, 2, \dots, n\}$.

Output: A pure Nash equilibrium allocation.

Sort links in an increasing order based on b_k .

for $1 \leq i \leq n$ **do**

 Solve the Linear program (Algorithm 4), for $t = b_{i+1}$ and $j = i$. If it returns a feasible solution x^* , then STOP and return x^* .

end

By the definition of Nash equilibrium, we know that there is at least one *Nash equilibrium* in which all links that are used have the same latency, $v = a_i x_i + b_i$, and any link k that is not used has a latency no less than v , or in other words $a_k \cdot 0 + b_k \geq v$. However, we do not know *a priori* the value of v , but we do know that the possible values are between the intervals of $[b_1, b_2], (b_2, b_3], \dots, (b_{n-1}, b_n], (b_n, +\infty)$, since b_i s are sorted in an increasing order. Note that if the optimal threshold is in interval $[b_{k-1}, b_k]$, then in the Nash equilibrium allocation we will have exactly $k - 1$ links. We exhaustively search the optimal values in an increasing order for any possible interval, see Algorithm 3. For any interval we solve a linear program to check if there is an allocation with the Nash equilibrium properties, if there is such an allocation we return it. Since, we know the existence of such an equilibrium, our algorithm always returns a feasible solution at the end. It is easy to see that the total computational time of the algorithm is polynomial.

Algorithm 4 Linear program**Input:** A positive threshold t and an index j .**Output:** x .

$$\begin{aligned}
& \text{minimize } v \\
& \text{s.t.} \\
& x_i = \frac{v - b_i}{a_i}, \text{ for } i \leq j. \\
& \sum_{i=1}^n x_i = r. \\
& x_i \geq 0, v \geq 0 \text{ and } v \leq t.
\end{aligned}$$

8.3 Coordination mechanisms with misinformation

Now let us focus on the case of single-commodity misinformation non-atomic congestion games, where the actual game has n parallel links and affine latency functions. Further, we restate the main question of the paper, that is how we can use misinformation in order to improve the performance of single-commodity non-atomic congestion games with n parallel links in terms of Social Cost.

Specifically, note that if we properly change the coefficients of the latency functions of the misinformed games then the flow according to the worst natural misinformed equilibrium will change. To that direction we choose to increase the coefficients of the latency function from a_k, b_k to \hat{a}_k^j, \hat{b}_k^j (one for each different subjective view Γ^j respectively). Further, we assume that the designer has the constraint that he can provide a limited number of misinformed views.

We will show that it is always possible to find a unique natural misinformed equilibrium that coincides with the optimal allocation, in terms of social welfare, by appropriately changing the coefficients. The constructed misinformation game $m\Gamma = \langle \Gamma^0, \Gamma^1, \dots, \Gamma^N \rangle$, with θ splitting, that will be derived in the next Subsection, will have the following properties:

- i) $\Gamma^0 = \Gamma$ (the case of n parallel links),
- ii) $\Gamma^0 = \langle G, l, s, t, 1 \rangle$ and $\Gamma^j = \langle G, l^j, s, t, 1 \rangle$, with $l_k(x_k) = a_k x_k + b_k$ and $l_{j,k}(x_k) = \hat{a}_k^j x_k + \hat{b}_k^j$, respectively.
- iii) $\theta = \langle \theta^1, \dots, \theta^N \rangle$.

Next, we provide a methodology so as to construct a misinformation game in such a way that the unique natural misinformed equilibrium in the misinformation game is an optimal allocation of the actual game. Towards that direction we give a simple algorithm that takes as input a single-commodity non-atomic congestion game Γ^0 , the optimal allocation in Γ^0 ,

and an arbitrary break down (partition) of the links that are used in the optimal flow. Let $x^* = (x_1^*, x_2^*, \dots, x_n^*)$ be the optimal solution, which we can easily find in polynomial time as a minimization of a convex function, and an abstract partition (k_1, \dots, k_m) of the allocation x^* over the parallel links that are used; m is the number of misinformed views that the designer provides to the players. E.g., if $n = 3$ with $x_i^* > 0$ for any i , then $k_1 = \{1, 2\}$, $k_3 = \{3\}$ is a possible partition. With Algorithm 5 we construct a misinformation game $m\Gamma$, where players perform optimally in terms of Social Cost.

Algorithm 5 Coordination mechanism algorithm for an abstract partition (k_1, \dots, k_m) of n

Input : An actual game $\Gamma^0 = \langle G, l^j, s, t, 1 \rangle$
 An optimal allocation $x^* = (x_1^*, \dots, x_n^*)$
 A partition over the links, (k_1, \dots, k_m) .
Output: A misinformation game $m\Gamma$.
 A splitting θ .

```

while  $1 \leq i \leq m$  do
  New allocation  $y^*$ :
    
$$y_j^* = \begin{cases} \frac{x_j^*}{\sum_{t \in \{k_i\}} x_t^*}, & j \in k_i \\ 0, & \text{elsewhere} \end{cases}$$

  Apply Algorithm 6 for  $y^*$  to
  construct the latency functions of  $\Gamma^i$ .
   $\theta^i = \sum_{j \in \{k_i\}} x_j^*$ .
end
 $m\Gamma \leftarrow \langle \Gamma^0, \Gamma^1, \dots, \Gamma^m \rangle$ .
 $\theta \leftarrow \langle \theta^1, \dots, \theta^m \rangle$ .
```

We run Algorithm 5 and we get the misinformation game. Note that, it produces one misinformation view Γ^i at a time. Then, calling Algorithm 6 inside Algorithm 5, we take the coefficients for the latency functions for Γ^i . Then, Γ^i is entailed with ease, to produce $m\Gamma$.

At the beginning of Algorithm 6, we initialize v by setting it equal to the maximum of the costs of the latency functions over the links that are used in the allocation y^* . Then, we increase the b_i for the links that are not used in the allocation y^* in order to make them no less than the cost v . For any link i that is used in the allocation y^* we can increase the b_i in such a way that the cost of this link with the allocation y_i^* is equal to the cost v . This procedure can be done in polynomial-time by solving a system of linear inequalities. For any y^* it is easy to see that Algorithm 3 gives a unique pure Nash equilibrium in the modified game for the players that have this view. Taking the natural misinformed equilibrium we construct the allocation x^* , which is the optimal allocation of the actual game, hence $PoM = 1$.

In Figure 8.2 we provide schematically the pipeline regarding coordination mechanism

Algorithm 6 Coordination mechanism algorithm

Input : Coefficients a_i, b_i of latency functions
for any link i . The allocation y^* .

Output: New coefficients of latency functions.

Put $v = \max_{i: y_i^* > 0} \{a_i y_i^* + b_i\}$.

Find \hat{b}_i , for any i ,

s.t.

$\hat{b}_i \geq v$, for any i such that $y_i^* = 0$,

$a_i y_i^* = v - \hat{b}_i$, for any i such that $y_i^* > 0$,

$\hat{b}_i \geq b_i$, for any i .

return \hat{b} .

with misinformation.

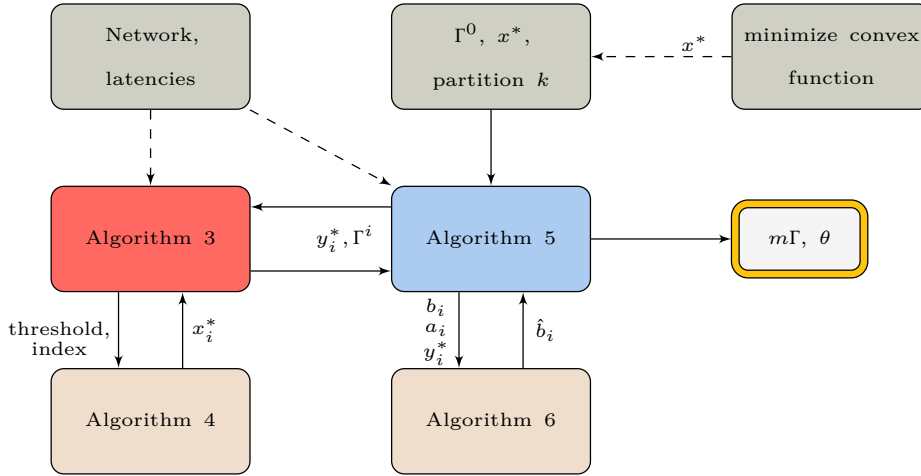


Figure 8.2: Coordination mechanism with misinformation.

Further, similar mechanism can be used in the case where the designer can influence only a fragment of the players. Hence, he/she can construct a mechanism where the θ^1 portion gets misinformed in order to improve Social Cost, whereas the rest use the actual game, i.e., the resulting misinformation game would be $m\Gamma = \langle \Gamma^0, \Gamma^1, \Gamma^2 \rangle$, where $\Gamma^2 = \Gamma^0$. In parallel, the splitting θ becomes $\langle \theta^1, \theta^2 \rangle$.

The first step is to reconsider the optimal allocation for the θ^1 fragment of the flow, taking into account the fact that there is a fixed part of the players θ^2 will route according to Γ^2 , which has the same latency functions as Γ^0 . There we reconsider the coefficients of the latency functions as they experience the additional cost of the Nash equilibrium flow of the

unmisinformed fragment θ^2 . Afterwards, we implement our mechanism and get the desired $m\Gamma$.

Observe that as we can affect only θ^1 portion of the players, the rest should route (possibly sub-optimally) according to the actual specifications. Thus, is evident that in case where designer can influence a limited part of the flow it is possible that $PoM > 1$.

Chapter 9

Conclusions

After presenting our work, its applicability and the relevant results we now revisit the main outcomes from a more high-level standpoint providing a research path and the contributions regarding each area of the dissertation. Afterwards, we remark prominent future research directions.

9.1 Research path

In this dissertation, we studied the effect of misinformation in a multi-agent system, considering that the participants are intelligent and rational. Namely, each agent possibly has the wrong information regarding the actual interaction. Initially, starting from classical games in game theory we presented the basic setting of our model, in Chapter 4, and we provided several classes of games where it can be applied, but we restricted in “single shot” interactions. Then, in Chapter 5, we expanded our model so as to include interactions where agents take several decisions sequentially. The sequential interaction raises several issues in respect of the reconsideration and introspection that agents had to do. Hence, in Chapter 5, we further established an epistemic approach, see Figure 9.1. Afterwards, we presented a learning algorithm in order to compute the equilibria concepts that have emerged from our model, in Chapter 6. Then, we studied the vulnerability of agents’ decisions in the case where their information is susceptible to noise, in Chapter 7. Finally, in Chapter 8, we presented and analyzed a setting where a designer can take advantage of the misinformation and improved the efficiency of the multi-agent system.

9.2 Synopsis of Contributions

9.2.1 Misinformation Games

Chapter 4 is the bedrock of this dissertation. It is motivated by the idea that misinformation is a fact of life in most multi-player interactions, and thus having the formal machinery to analyse misinformation can help understand many real-world phenomena. Towards this aim,

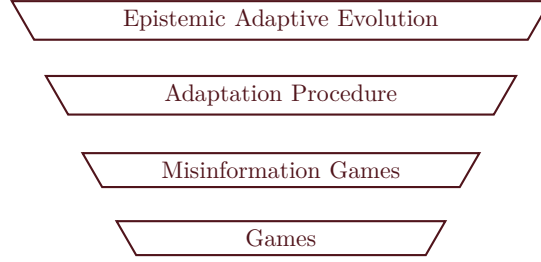


Figure 9.1: Top view of the structure of misinformation games model.

we introduced a novel game-theoretic framework, called *misinformation games*. Further, we provided a metric that measures the deterioration/improvement of a system, in terms of social welfare, due to misinformation in the views of the players, we called this metric *Price of Misinformation (PoM)*. Afterwards, we applied misinformation games in the areas of load-balancing games and congestion games, thus establishing the applicability of our model.

9.2.2 Adaptation Procedure

Next, we enriched misinformation games with a methodology for studying game-playing scenarios where misinformed players revised their game-related information as they interact with their environment. We considered the case where the revision is based on the received payoffs, which are publicly announced in each game iteration. To formalize the process, we defined the *Adaptation Procedure*, which describes the changes in the decisions of the players as they obtain new information for the environment. This leads to a new equilibrium concept, called the *stable misinformed equilibrium*, which is the strategy profile(s) that the players choose when the Adaptation Procedure has stabilized. Unsurprisingly, players do not need to fully learn the actual specifications of the interaction in order for the Adaptation Procedure to end.

9.2.3 Epistemic Adaptive Evolution

Afterwards, we expand our methodology so as to provide the players with the formal machinery to revise their knowledge and beliefs with regards to: i) the information that their opponents have, and ii) the way they make their decisions. Initially, each player has his/her own subjective view that is provisioned by his/her knowledge and beliefs. We expand misinformation games in order to include the beliefs of each participant, we call such games *epistemic misinformation games*. Moreover, since each participant comes along with an “irrational” decision must reconsider his/her mindset regarding the other participants, we call this process *Epistemic Adaptive Evolution*. With this at hand, each player can learn the way that other players interact and exploit this. This lead to a new equilibrium concept, called the *stable epistemic*

misinformed equilibrium.

We also provided different ways to visualize the process and showed several properties related to the aforementioned procedures and the respective equilibria, pertaining to its existence, termination, length etc.

9.2.4 Learning Dynamics

In Chapter 6 we presented a novel learning algorithm that obtains last-iterate convergence to the equilibrium point for a zero-sum game, called Forward Looking Best-Response Multiplicative Weights Update (FLBR-MWU) method. A structural characteristic of our methodology is that given a state x , the FLBR-MWU method first generates an intermediate state taking a best-response step. However, instead of continuing from that step, it goes back to the original state x in order to generate a new state x' . As both theoretical and experimental results show FLBR-MWU dynamics attain significant acceleration regarding the state-of-the-art dynamics such as OMWU and OMD, see Section 6.3.

With this at hand, we have an online learning technique to compute natural and stable misinformed equilibria in the case where both subjective and actual specifications belong to the class of zero-sum games.

9.2.5 Noisy Games

Afterwards, in Chapter 7 we introduce a subclass of misinformation games called *noisy games*, where players receive the information of the game with a distortion affecting the elements of the payoff matrices, due to additive noise that follows a normal distribution. We study the influence of the parameters of the noise in the strategic behavior of the players, and we derive probabilistic formulas that capture this influence. Namely, we analyze the cases where a player suffering from noise would not alter his/her choices if the noise disappears and vice versa, we call this analysis *behavioural consistency*, see Subsection 7.2.3. As a result, we define the concepts of $(inv)\varepsilon$ -misinformed players, and we establish probabilistic formulas that quantify the probability of a strategy profile to be (inv) -*e-mis* due to noise, Section 7.3. Our analysis restricted in the case of two players' bimatrix games with two strategies per player, due to highly intricate formulas and complex mathematical objects that are hard to be represented.

Afterwards, we perform numerical experiments using four benchmark bimatrix games, see Figure 7.3. Initially, we compare the probabilistic formulas with Monte Carlo simulations. Then, we derive general remarks as to the efficiency of the system regarding the additive noise, in terms of social welfare. We quantify these results using the *Price of Misinformation* metric, in order to depict how benevolent/malevolent could be the noise regarding game's performance.

Undeniably the two players' bimatrix games with two strategies per player is a very

restricting setting. Unsurprisingly, however, even in this simple setting our analysis highlighted the richness, intricacy and interdependence of the probabilistic events, mathematical objects and techniques that are involved, which further resulted in complex formulations.

9.2.6 Mechanism Design

Finally, we explored the use of *misinformation* as a novel and powerful method for Coordination mechanisms, in Chapter 8. We applied this idea in single-commodity non-atomic congestion games with parallel links and affine latency functions. Specifically, our goal was to steer players' behaviour towards the socially optimum allocation, by misinforming them regarding the latency functions of the network.

Towards that direction, we provide two polynomial-time algorithms. The first finds a Nash equilibrium flow allocation in a single-commodity non-atomic congestion game with n parallel links and affine cost functions. The second takes as input an abstract partition over the links that are used in the optimal allocation and creates a misinformation game whose subjective games follow the required specification. Consequently, its natural misinformed equilibrium is the optimum allocation in the actual game.

9.3 Directions for Future Work and Research

We argue that the concept of misinformation games has the potential to explain various phenomena, and raises several interesting problems to be studied from different perspectives.

Misinformation Games. An interesting future work would be in the directions of: i) proving tighter bounds regarding *PoM*, and ii) transfusing misinformation game framework in several areas of game theory.

Adaptation Procedure Initiated by the ending notes of Chapter 5, we acknowledge that there are many directions for improvement such as: i) dropping the assumption that players fully update their information (e.g., they may not be able to observe other players' payoffs); ii) considering different players' attitudes towards the new information and/or towards the realization that their knowledge is inaccurate; iii) considering the case where players experiment with sub-optimal actions; iv) considering cases where the players attempt to exploit other players' inaccurate knowledge, (e.g., by playing non-equilibrium, higher-payoff strategies), by observing their actions and making inferences as to what they know and what they don't; v) considering the scenario where the actual game specifications also evolve over time.

Epistemic Adaptive Evolution. As this part concludes the misinformation games framework there are a lot of different directions where we can expand our methodology, such as: i)

study the ways a player exploits the strategic thinking of his/her opponent, ii) implement our framework using Logical Calculus, and iii) evaluate it in real-life situations. Further, the follow-ups of the Adaptation Procedure can also be considered as future work for the Epistemic Adaptive Evolution.

Learning Dynamics Here, future directions are three-folded. First, the enhancement of FLBR-MWU method in order to compute all the equilibria concepts that arise in this dissertation, without the zero-sum constraint. Second, from an algorithmic perspective, as our results suggest, the implementation of a best-response step is a flexible add-on that can be easily attached to a wide variety of GAN training methods and provides noticeable gains in performance and stability. Third, we consider as a very interesting future step to implement FLBR-MWU method for more general classes of games and objective functions. Additionally, it could be interesting to examine modifications of FLBR-MWU method, such as adaptive schemes for ξ and η throughout the iterations.

Noisy Games. There are several potential research directions for this part of the study. Specifically, we could derive analogous probabilistic formulas for other classes of noise distributions. Further, it would be very intriguing to expand our formulas in cases of bimatrix games with abstract, yet bounded, number of pure strategies, or in cases of $N > 2$ number of players.

Moreover, an intriguing future step would be to provide tools to quantify the *sensitivity* of a game to random noise, i.e., determine “how much noise” the game can withstand so that the behaviour of the players remains close (under the above sense) to the expected ones, with a certain probability. A related research question is how sensitivity is affected by inconsequential changes in the game specification (e.g., change of scale). This could be used as a tool for game designers to improve their designs and make them more robust to unexpected circumstances. Thus, enhancing the theoretical results presented in Chapter 8.

Mechanism Design An immediate future direction is to use the methodology of misinformation games in order to design mechanisms for serial-parallel networks and general latency functions. It is also interesting to prove bounds regarding *PoM*.

Recently, authors in [Omidshafiei et al., 2019] introduced a general descriptive multi-agent evaluation method, called α -Rank. They focus on Empirical game theory and specifically on the concept of empirical games, or meta-games, and the convergence of their dynamics to the Nash equilibria [Tuyls et al., 2018]. Interestingly, they provide a new paradigm. In particular, they allow the dynamics to roll out and enable strong (i.e., non-transient) agents to emerge and weak (i.e., transient) agents to vanish naturally through their long-term interactions. On the contrary, the classical *Nash approach* tries to identify static points in the simplex that

capture simultaneous best response behaviours of the agents. In other words, the strategies chosen by the new approach are those favored by evolutionary selection, as opposed to the Nash strategies, which are simultaneous best-responses.

The new approach provides substantial gains regarding many pathologies of the Nash equilibrium concept, such as the tractability issue. Further is a natural candidate to resolve the incompatibility of the classical Nash approach with the dynamic behaviours of agents in interacting systems.

In the case of the misinformation games framework, we use the classical Nash equilibrium concept as an ingredient of our model, whereas the resulting equilibrium concepts (*nme*, *sme*, *ene*, and *stene*) can be considered as *weak* Nash equilibrium concepts. The paradigm we offer with this framework is that: i) interactions should be analysed under the perspective of the subjective views of the participants, and ii) the revision of the views is taking place simultaneously in the knowledge and the beliefs of the participants.

Bringing together these two paradigms, we may come up with a unified framework where the interaction in a multi-agent system unfolded in an evolutionary manner and the solutions are justified through epistemic revision techniques. This agglomeration would provide a strong evolutionary revision methodology. Thus, it would be very interesting to plug in the α -Rank method into our framework.

To sum up, we deem that this dissertation could be expatiated and applied to several and diverse areas of Discipline, such as Economics (i.e. in stock markets where dealers try to manipulate their opponents), (Cyber-)security (i.e. defenders and attackers with limited resources), Telecommunications (e.g. in cases where information passes through communication channels), and Artificial Intelligence (i.e. intelligent agents learn a procedure or derive conclusions as to the information of the others).

While the road ahead is still long, we have come far enough to say with some confidence that there is a lot of promise in the paradigm presented in this thesis. As we surround ourselves with more and more technology, data, complex social rules, inconsistent news, cognitive inadequacy, conflicting theories, non-provable assumptions, etc. maybe it is time to start thinking about interactions under the scope that participants are far from being considered as correctly informed.

Taking a distant stance against this dissertation we should wonder what's the meaning of *misinformation games* paradigm. We can consider the core of this as a *low knowledge process*, meaning that the participants do not and shall not ever be in place to know (and to know that they know, etc.) entirely the correct specifications ("truth"). Contrary to *high knowledge process*, such as probabilistic approaches, where the state of mind of the participants

is something like “the truth is out there”, or in other words the “truth is consist of elements on my grasp (or near grasp)” that need to be tuned properly. But how plausible is the latter paradigm in real-life interactions? What is that prominent “truth”? Is it everlasting? Are we, even slightly, certain of what we are looking for? In case scholars can provide affirmative answers to that questions, and many more of the same spirit, then *low knowledge process* is just an exercise without any interest and *high knowledge process* is inapt per mare per terra. Otherwise, the *low knowledge process* is the man of the hour.

Bibliography

- [Acar and Meir, 2020] Acar, E. and Meir, R. (2020). Distance-based equilibria in normal-form games. In *34th AAAI Conference on Artificial Intelligence, AAAI 2020, 32th Innovative Applications of Artificial Intelligence Conference, IAAI 2020, 10th AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2020*, pages 1750–1757. AAAI Press.
- [Acemoglu et al., 2018] Acemoglu, D., Makhdoumi, A., Malekian, A., and Ozdaglar, A. E. (2018). Informational Braess’ paradox: The effect of information on traffic congestion. *Operations Research*, 66(4):893–917.
- [Adler, 2013] Adler, I. (2013). The equivalence of linear programs and zero-sum games. *International Journal of Game Theory*, 42:165–177.
- [Aghassi and Bertsimas, 2006] Aghassi, M. and Bertsimas, D. (2006). Robust game theory. *Math. Program.*, 107(1-2):231–273.
- [Antos and Pfeffer, 2010] Antos, D. and Pfeffer, A. (2010). Representing bayesian games without a common prior. In *9th International Conference on Autonomous Agents and Multiagent Systems (AAMAS’10)*, pages 1457–1458. IFAAMAS.
- [Arora et al., 2012] Arora, S., Hazan, E., and Kale, S. (2012). The multiplicative weights update method: a meta-algorithm and applications. *Theory Comput.*, 8(1):121–164.
- [Arrow and Green, 1973] Arrow, K. J. and Green, J. R. (1973). Notes on expectations equilibria in bayesian settings institute for mathematical studies in the social sciences. Technical report. Working Paper No. 33, Stanford University.
- [Aumann, 1987] Aumann, R. J. (1987). Correlated equilibrium as an expression of bayesian rationality. *Econometrica*, 55(1):1–18.
- [Aumann and Brandenburger, 1995] Aumann, R. J. and Brandenburger, A. (1995). Epistemic conditions for nash equilibrium. *Econometrica*, 63:1161–1180.
- [Avis et al., 2010] Avis, D., Rosenberg, G., Savani, R., and von Stengel, B. (2010). Enumeration of nash equilibria for two-player games. *Economic Theory*, 42:9–37.
- [Azizian et al., 2020] Azizian, W., Scieur, D., Mitliagkas, I., Lacoste-Julien, S., and Gidel, G. (2020). Accelerating smooth games by manipulating spectral shapes. In *AISTATS*.

- [Bach and Tsakas, 2014] Bach, C. W. and Tsakas, E. (2014). Pairwise epistemic conditions for nash equilibrium. *Games Econ. Behav.*, 85:48–59.
- [Bailey and Piliouras, 2018] Bailey, J. P. and Piliouras, G. (2018). Multiplicative weights update in zero-sum games. In *Conference on Economics and Computation (EC’18)*, pages 321–338.
- [Bailey and Piliouras, 2019] Bailey, J. P. and Piliouras, G. (2019). Fast and furious learning in zero-sum games: Vanishing regret with non-vanishing step sizes. In *Advances in Neural Information Processing Systems 32: Annual Conference on Neural Information Processing Systems 2019, NeurIPS 2019*, pages 12977–12987.
- [Bakker et al., 2021] Bakker, C., Bhattacharya, A., Chatterjee, S., and Vrabie, D. L. (2021). Metagames and hypergames for deception-robust control. *ACM Trans. Cyber-Phys. Syst.*, 5(3).
- [Balcan and Braverman, 2017] Balcan, M. and Braverman, M. (2017). Nash equilibria in perturbation-stable games. *Theory of Computing*, 13(1):1–31.
- [Balcan et al., 2009] Balcan, M.-F., Blum, A., and Mansour, Y. (2009). The price of uncertainty. In *10th ACM Conference on Electronic Commerce (EC)*, pages 285–294. ACM.
- [Banks et al., 2020] Banks, D., Gallego, V., Naveiro, R., and Ríos Insua, D. (2020). Adversarial risk analysis: An overview. *WIREs Computational Statistics*.
- [Bárány et al., 2007] Bárány, I., Vempala, S. S., and Vetta, A. (2007). Nash equilibria in random games. *Random Structures and Algorithms*, 31:391–405.
- [Beck, 2017] Beck, A. (2017). *First-Order Methods in Optimization*. Society for Industrial and Applied Mathematics.
- [Belardinelli and Rendsvig, 2020] Belardinelli, G. and Rendsvig, R. K. (2020). Awareness logic: A kripke-based rendition of the heifetz-meier-schipper model. *ArXiv*, abs/2012.12982.
- [Bennett, 1980] Bennett, P. G. (1980). Hypergames: Developing a model of conflict. *Futures*, 12(6):489–507.
- [Bilò et al., 2010] Bilò, V., Fanelli, A., Flammini, M., and Moscardelli, L. (2010). When ignorance helps: Graphical multicast cost sharing games. *Theoretical Computer Science*, 411(3):660–671.
- [Boyd and Vandenberghe, 2004] Boyd, S. and Vandenberghe, L. (2004). *Convex optimization*. Cambridge university press.

- [Brown, 2020] Brown, P. N. (2020). Partial altruism is worse than complete selfishness in nonatomic congestion games.
- [Brown et al., 2017] Brown, P. N., Borowski, H. P., and Marden, J. R. (2017). Are multiagent systems resilient to communication failures?
- [Bubeck, 2015] Bubeck, S. (2015). Convex optimization: Algorithms and complexity. *Found. Trends Mach. Learn.*, 8(3-4):231–357.
- [Burkov and Chaib-draa, 2009] Burkov, A. and Chaib-draa, B. (2009). Effective learning in the presence of adaptive counterparts. *J. Algorithms*, 64(4):127–138.
- [Caragiannis et al., 2006] Caragiannis, I., Kaklamanis, C., and Kanellopoulos, P. (2006). Taxes for linear atomic congestion games. In *Algorithms - ESA 2006, 14th Annual European Symposium, Zurich, Switzerland, Proceedings*, volume 4168, pages 184–195. Springer.
- [Caragiannis et al., 2010a] Caragiannis, I., Kaklamanis, C., and Kanellopoulos, P. (2010a). Taxes for linear atomic congestion games. *ACM Transactions on Algorithms*, 7(1):13:1–13:31.
- [Caragiannis et al., 2010b] Caragiannis, I., Kaklamanis, C., Kanellopoulos, P., Kyropoulou, M., and Papaioannou, E. (2010b). The impact of altruism on the efficiency of atomic congestion games. In *Trustworthy Global Computing - 5th International Symposium, TGC 2010, Munich, Germany, Revised Selected Papers*, volume 6084, pages 172–188. Springer.
- [Cesa-Bianchi and Lugosi, 2006] Cesa-Bianchi, N. and Lugosi, G. (2006). *Prediction, Learning, and Games*. Cambridge University Press.
- [Chaib-Draa, 2001] Chaib-Draa, B. (2001). Hypergame analysis in multiagent environments. In *Proceedings of the AAAI Spring Symposium*, pages 147–149.
- [Chen et al., 2014] Chen, P., de Keijzer, B., Kempe, D., and Schäfer, G. (2014). Altruism and its impact on the price of anarchy. *ACM Trans. Economics and Comput.*, 2(4):17:1–17:45.
- [Chen et al., 2009] Chen, X., Deng, X., and Teng, S. (2009). Settling the complexity of computing two-player nash equilibria. *J. ACM*, 56(3):14:1–14:57.
- [Cheung and Piliouras, 2020] Cheung, Y. K. and Piliouras, G. (2020). Chaos, extremism and optimism: Volume analysis of learning in games. In *Advances in Neural Information Processing Systems 33: Annual Conference on Neural Information Processing Systems 2020, (NeurIPS’20)*.
- [Chiang et al., 2012] Chiang, C., Yang, T., Lee, C., Mahdavi, M., Lu, C., Jin, R., and Zhu, S. (2012). Online optimization with gradual variations. In *25th Annual Conference on Learning Theory COLT’12*, pages 6.1–6.20.

- [Cho et al., 2019] Cho, J.-H., Zhu, M., and Singh, M. (2019). *Modeling and Analysis of Deception Games Based on Hypergame Theory*, pages 49–74. Springer International Publishing.
- [Christodoulou et al., 2009] Christodoulou, G., Koutsoupias, E., and Nanavati, A. (2009). Coordination mechanisms. *Theoretical Computer Science*, 410(36):3327–3336.
- [Christodoulou et al., 2014] Christodoulou, G., Mehlhorn, K., and Pyrga, E. (2014). Improving the price of anarchy for selfish routing via coordination mechanisms. *Algorithmica*, 69(3):619–640.
- [Cole et al., 2003] Cole, R., Dodis, Y., and Roughgarden, T. (2003). Pricing network edges for heterogeneous selfish users. In *35th Annual ACM Symposium on Theory of Computing*, pages 521–530.
- [Cole et al., 2018] Cole, R., Lianas, T., and Nikolova, E. (2018). When does diversity of agent preferences improve outcomes in selfish routing? In *27th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 173–179.
- [Copic and Galeotti, 2006] Copic, J. and Galeotti, A. (2006). Awareness as an equilibrium notion: Normal-form games.
- [Correa et al., 2008] Correa, J. R., Schulz, A. S., and Moses, N. E. S. (2008). A geometric approach to the price of anarchy in nonatomic congestion games. *Games and Economic Behavior*, 64(2):457–469.
- [Cover and Thomas, 2006] Cover, T. M. and Thomas, J. A. (2006). *Elements of Information Theory*. John Wiley and Sons.
- [Dai et al., 2018] Dai, B., Shaw, A., Li, L., Xiao, L., He, N., Liu, Z., Chen, J., and Song, L. (2018). SBEED: convergent reinforcement learning with nonlinear function approximation. In *35th International Conference on Machine Learning, ICML’18*, pages 1133–1142.
- [Daskalakis et al., 2007] Daskalakis, C., Dimakis, A. G., and Mossel, E. (2007). Connectivity and equilibrium in random games. *CoRR*, abs/math/0703902.
- [Daskalakis et al., 2009] Daskalakis, C., Goldberg, P. W., and Papadimitriou, C. H. (2009). The complexity of computing a nash equilibrium. *SIAM J. Comput.*, 39(1):195–259.
- [Daskalakis et al., 2018] Daskalakis, C., Ilyas, A., Syrgkanis, V., and Zeng, H. (2018). Training GANs with optimism. In *Proceedings of the International Conference on Learning Representations (ICLR’18)*.
- [Daskalakis and Panageas, 2019] Daskalakis, C. and Panageas, I. (2019). Last-iterate convergence: Zero-sum games and constrained min-max optimization. In *ITCS’19*.

- [Dekel et al., 1999] Dekel, E., Fudenberg, D., and Levine, D. K. (1999). Payoff information and self-confirming equilibrium. *Journal of Economic Theory*, 89(2):165–185.
- [Dresher, 1970] Dresher, M. (1970). Probability of a pure equilibrium point in n -person games. *Journal of Combinatorial Theory, Series A*, 8:134–145.
- [Easley and Kiefer, 1988] Easley, D. and Kiefer, N. M. (1988). Controlling a Stochastic Process with Unknown Parameters. *Econometrica*, 56(5):1045–1064.
- [Esponda and Pouzo, 2016a] Esponda, I. and Pouzo, D. (2016a). Berk-nash equilibrium: A framework for modeling agents with misspecified models. *Econometrica*, 84(3):1093–1130.
- [Esponda and Pouzo, 2016b] Esponda, I. and Pouzo, D. (2016b). Equilibrium in misspecified markov decision processes. *ERN: Other Microeconomics: Decision-Making under Risk & Uncertainty (Topic)*.
- [Esponda and Pouzo, 2019] Esponda, I. and Pouzo, D. (2019). Asymptotic behavior of bayesian learners with misspecified models. *arXiv: Theoretical Economics*.
- [Etessami and Yannakakis, 2010] Etessami, K. and Yannakakis, M. (2010). On the complexity of Nash equilibria and other fixed points. *SIAM J. Comput.*, 39(6):2531–2597.
- [Facchinei and Pang, 2003] Facchinei, F. and Pang, J. S. (2003). Finite-dimensional variational inequalities and complementarity problems.
- [Fagin and Halpern, 1987] Fagin, R. and Halpern, J. Y. (1987). Belief, awareness, and limited reasoning. *Artif. Intell.*, 34:39–76.
- [Fasoulakis et al., 2021] Fasoulakis, M., Markakis, E., Pantazis, Y., and Varsos, C. (2021). Forward looking best-response multiplicative weights update methods for bilinear zero-sum games. *CoRR*, abs/2106.03579.
- [Fasoulakis et al., 2019] Fasoulakis, M., Traganitis, A., and Ephremides, A. (2019). Jamming in multiple independent Gaussian channels as a game. In *8th International EAI Conference on Game Theory for Networks (GameNets’19)*, volume 277, pages 3–8.
- [Feinberg, 2020] Feinberg, Y. (2020). Games with unawareness. *The B.E. Journal of Theoretical Economics*.
- [Fleischer et al., 2004] Fleischer, L., Jain, K., and Mahdian, M. (2004). Tolls for heterogeneous selfish users in multicommodity networks and generalized congestion games. In *45th Symposium on Foundations of Computer Science (FOCS’04)*, pages 277–285.
- [Fotakis and Spirakis, 2008] Fotakis, D. and Spirakis, P. G. (2008). Cost-balancing tolls for atomic network congestion games. *Internet Mathematics*, 5(4):343–363.

- [Freund and Schapire, 1999] Freund, Y. and Schapire, R. E. (1999). Adaptive game playing using multiplicative weights. *Games and Economic Behavior*, 29:79–103.
- [Fudenberg and Kamada, 2015] Fudenberg, D. and Kamada, Y. (2015). Rationalizable partition-confirmed equilibrium. *Theoretical Economics*, 10:775–806.
- [Fudenberg and Levine, 1993] Fudenberg, D. and Levine, D. K. (1993). Self-confirming equilibrium. *Econometrica*, 61(3):523–545.
- [Fudenberg et al., 2016] Fudenberg, D., Romanyuk, G., and Strack, P. (2016). Active learning with misspecified beliefs. *ERN: Search*.
- [Galer, 2007] Galer, O. (2007). *Discrete Dynamical Systems*. Springer-Verlag.
- [Gharesifard and Cortés, 2011] Gharesifard, B. and Cortés, J. (2011). Stealthy strategies for deception in hypergames with asymmetric information. In *2011 50th IEEE Conference on Decision and Control and European Control Conference*, pages 5762–5767.
- [Gharesifard and Cortés, 2012] Gharesifard, B. and Cortés, J. (2012). Evolution of players’ misperceptions in hypergames under perfect observations. *IEEE Trans. Autom. Control.*, 57(7):1627–1640.
- [Gidel et al., 2019] Gidel, G., Berard, H., Vignoud, G., Vincent, P., and Lacoste-Julien, S. (2019). A variational inequality perspective on generative adversarial networks. In *7th International Conference on Learning Representations, ICLR 2019*.
- [Goodfellow et al., 2014] Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y. (2014). Generative Adversarial Nets. In *Proceedings of Annual Conference on Neural Information Processing Systems (NIPS ’14)*, pages 2672–2680.
- [Halpern and Rêgo, 2014] Halpern, J. Y. and Rêgo, L. C. (2014). Extensive games with possibly unaware players. *Math. Soc. Sci.*, 70:42–58.
- [Harsanyi, 1967] Harsanyi, J. C. (1967). Games with incomplete information played by “Bayesian” players, I–III Part I. The basic model. *Management Science*, 14(3):159–182.
- [Heidhues et al., 2018] Heidhues, P., KE“szegi, B., and Strack, P. (2018). Convergence in misspecified learning models with endogenous actions. *Econometrics: Econometric Model Construction*.
- [Heidhues et al., 2017] Heidhues, P., oszegi, B. K., and Strack, P. (2017). Unrealistic expectations and misguided learning. *CSN: Economics (Topic)*.

- [Heifetz et al., 2006] Heifetz, A., Meier, M., and Schipper, B. C. (2006). Interactive unawareness. *J. Econ. Theory*, 130:78–94.
- [Heifetz et al., 2008] Heifetz, A., Meier, M., and Schipper, B. C. (2008). A canonical model for interactive unawareness. *Games and Economic Behavior*, 62(1):304–324.
- [Heifetz et al., 2013] Heifetz, A., Meier, M., and Schipper, B. C. (2013). Unawareness, beliefs, and speculative trade. *Games Econ. Behav.*, 77:100–121.
- [Heinsalu, 2011] Heinsalu, S. (2011). Universal type spaces with unawareness. *Microeconomics: Decision-Making under Risk & Uncertainty eJournal*.
- [Hoi et al., 2018] Hoi, S. C. H., Sahoo, D., Lu, J., and Zhao, P. (2018). Online learning: A comprehensive survey. *CoRR*, abs/1802.02871.
- [Jordan, 1991] Jordan, J. (1991). Bayesian learning in normal form games. *Games and Economic Behavior*, 3(1):60–81.
- [Kleer and Schäfer, 2016] Kleer, P. and Schäfer, G. (2016). The impact of worst-case deviations in non-atomic network routing games. In *9th International Symposium Algorithmic Game Theory (SAGT)*, pages 129–140.
- [Kleer and Schäfer, 2017a] Kleer, P. and Schäfer, G. (2017a). Path deviations outperform approximate stability in heterogeneous congestion games. In *Algorithmic Game Theory - 10th International Symposium, SAGT 2017, L'Aquila, Italy, Proceedings*, volume 10504, pages 212–224.
- [Kleer and Schäfer, 2017b] Kleer, P. and Schäfer, G. (2017b). Tight inefficiency bounds for perception-parameterized affine congestion games. In *Algorithms and Complexity - 10th International Conference, CIAC 2017, Athens, Greece, Proceedings*, volume 10236, pages 381–392.
- [Korpelevich, 1976] Korpelevich, G. M. (1976). The extragradient method for finding saddle points and other problems.
- [Koutsoupias and Papadimitriou, 1999] Koutsoupias, E. and Papadimitriou, C. (1999). Worst-case equilibria. In *16th annual conference on Theoretical aspects of Computer Science (STACS)*, pages 404–413.
- [Kovach et al., 2015] Kovach, N. S., Gibson, A. S., and Lamont, G. B. (2015). Hypergame theory: A model for conflict, misperception, and deception. *Game Theory*, 2015.
- [Lavi and Swamy, 2007] Lavi, R. and Swamy, C. (2007). Truthful mechanism design for multi-dimensional scheduling via cycle monotonicity. In *8th ACM Conference on Electronic Commerce (EC'07)*, pages 252–261.

- [Lei et al., 2021] Lei, Q., Nagarajan, S. G., Panageas, I., and Wang, X. (2021). Last iterate convergence in no-regret learning: constrained min-max optimization for convex-concave landscapes. In *24th International Conference on Artificial Intelligence and Statistics, AISTATS'21*, pages 1441–1449.
- [Leme and Tardos, 2010] Leme, R. P. and Tardos, É. (2010). Pure and Bayes-Nash price of anarchy for generalized second price auction. In *51th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 735–744.
- [Lerer et al., 2019] Lerer, A., Hu, H., Foerster, J., and Brown, N. (2019). Improving policies via search in cooperative partially observable games.
- [Liang and Stokes, 2019] Liang, T. and Stokes, J. (2019). Interaction matters: A note on non-asymptotic local convergence of generative adversarial networks. In *22nd International Conference on Artificial Intelligence and Statistics, AISTATS'19*, pages 907–915.
- [Littlestone and Warmuth, 1994] Littlestone, N. and Warmuth, M. K. (1994). The weighted majority algorithm. *Inf. Comput.*, 108(2):212–261.
- [Luce and Raiffa, 1957] Luce, R. D. and Raiffa, H. (1957). *Games and Decisions: Introduction and Critical Survey*. John Wiley and Sons, New York.
- [Meier and Schipper, 2014] Meier, M. and Schipper, B. C. (2014). Bayesian games with unawareness and unawareness perfection. *Economic Theory*, 56:219–249.
- [Meir and Parkes, 2015a] Meir, R. and Parkes, D. C. (2015a). Congestion games with distance-based strict uncertainty. In *29th AAAI Conference on Artificial Intelligence*, pages 986–992.
- [Meir and Parkes, 2015b] Meir, R. and Parkes, D. C. (2015b). Playing the wrong game: Smoothness bounds for congestion games with behavioral biases. *SIGMETRICS Performance Evaluation Review*, 43(3):67–70.
- [Meir and Parkes, 2018] Meir, R. and Parkes, D. C. (2018). Playing the wrong game: Bounding externalities in diverse populations of agents. In *17th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*, pages 86–94.
- [Mertikopoulos et al., 2019] Mertikopoulos, P., Lecouat, B., Zenati, H., Foo, C., Chandrasekhar, V., and Piliouras, G. (2019). Optimistic mirror descent in saddle-point problems: Going the extra (gradient) mile. In *7th International Conference on Learning Representations, ICLR 2019*.
- [Mertikopoulos et al., 2018] Mertikopoulos, P., Papadimitriou, C. H., and Piliouras, G. (2018). Cycles in adversarial regularized learning. In *29th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'18)*, pages 2703–2717.

- [Mokhtari et al., 2020] Mokhtari, A., Ozdaglar, A. E., and Pattathil, S. (2020). A unified analysis of extra-gradient and optimistic gradient methods for saddle point problems: Proximal point approach. In *Proceedings of The 23rd International Conference on Artificial Intelligence and Statistics, AISTATS'20*, pages 1497–1507.
- [Monderer and Tennenholtz, 2003] Monderer, D. and Tennenholtz, M. (2003). k-implementation. In *Proceedings 4th ACM Conference on Electronic Commerce (EC'03)*, pages 19–28.
- [Nash, 1951] Nash, J. F. (1951). Non-cooperative games. *The Annals of Mathematics*, 54(2):286–295.
- [Nemirovski, 2004] Nemirovski, A. (2004). Prox-method with rate of convergence $o(1/t)$ for variational inequalities with lipschitz continuous monotone operators and smooth convex-concave saddle point problems. *SIAM J. Optim.*, 15:229–251.
- [Nikolova and Moses, 2015] Nikolova, E. and Moses, N. E. S. (2015). The burden of risk aversion in mean-risk selfish routing. In *16th ACM Conference on Economics and Computation (EC)*, pages 489–506.
- [Nisan et al., 2007a] Nisan, N., Roughgarden, T., Tardos, E., and Vazirani, V. V. (2007a). *Algorithmic Game Theory*. Cambridge University Press, USA.
- [Nisan et al., 2007b] Nisan, N., Roughgarden, T., Tardos, E., and Vazirani, V. V. (2007b). *Algorithmic Game Theory*. Cambridge Univ. Press.
- [Norde, 1999] Norde, H. (1999). Bimatrix games have quasi-strict equilibria. *Math. Program.*, 85(1):35–49.
- [Nyarko, 1991] Nyarko, Y. (1991). Learning in mis-specified models and the possibility of cycles. *Journal of Economic Theory*, 55:416–427.
- [Omidshafiei et al., 2019] Omidshafiei, S., Papadimitriou, C. H., Piliouras, G., Tuyls, K., Rowland, M., Lespiau, J.-B., Czarnecki, W. M., Lanctot, M., Pérolat, J., and Munos, R. (2019). α -rank: Multi-agent evaluation by evolution. *arXiv: Multiagent Systems*.
- [Osborne and Rubinstein, 1994] Osborne, M. J. and Rubinstein, A. (1994). A course in game theory.
- [Ozbay, 2007] Ozbay, E. Y. (2007). Unawareness and strategic announcements in games with uncertainty. In *11th Conference on Theoretical Aspects of Rationality and Knowledge*, pages 231–238.
- [Peterson, 2009] Peterson, M. (2009). *An Introduction to Decision Theory*. Cambridge University Press.

- [Pigou, 1920] Pigou, A. C. (1920). *The Economics of Welfare*. Macmillan.
- [Quarteroni et al., 2006] Quarteroni, A., Sacco, R., and Saleri, F. (2006). *Numerical Mathematics (Texts in Applied Mathematics)*. Springer-Verlag.
- [Quattropani and Scarsini, 2020] Quattropani, M. and Scarsini, M. (2020). Efficiency of equilibria in random binary games. *CoRR*, abs/2007.08518.
- [Rakhlin and Sridharan, 2013] Rakhlin, S. and Sridharan, K. (2013). Optimization, learning, and games with predictable sequences. In *Proceedings of the Advances in Neural Information Processing Systems*, pages 3066–3074.
- [Ray et al., 2008] Ray, D., King-Casas, B., Montague, P. R., and Dayan, P. (2008). Bayesian model of behaviour in economic games. In *Advances in Neural Information Processing Systems 21, NIPS08*, pages 1345–1352.
- [Rinott and Scarsini, 2000] Rinott, Y. and Scarsini, M. (2000). On the number of pure strategy nash equilibria in random games. *Games Econ. Behav.*, 33(2):274–293.
- [Robinson, 1951] Robinson, J. (1951). An iterative method of solving a game. *Annals of Mathematics*, 54(2):296–301.
- [Romanyuk et al., 2017] Romanyuk, G., Strack, P., and Fudenberg, D. (2017). Active learning with a misspecified prior.
- [Roponen et al., 2020] Roponen, J., Ríos Insua, D., and Salo, A. (2020). Adversarial risk analysis under partial information. *European Journal of Operational Research*, 287(1):306–316.
- [Rosenthal, 1973] Rosenthal, R. W. (1973). A class of games possessing pure-strategy nash equilibria. *International Journal of Game Theory*, 2:65–67.
- [Roughgarden, 2003] Roughgarden, T. (2003). The price of anarchy is independent of the network topology. *Journal of Computer and System Sciences*, 64(2):341–364.
- [Roughgarden and Tardos, 2002] Roughgarden, T. and Tardos, E. (2002). How bad is selfish routing? *J. ACM*, 49(2):236–259.
- [Sasaki, 2014] Sasaki, Y. (2014). Subjective rationalizability in hypergames. *Adv. Decis. Sci.*, 2014:263615:1–263615:7.
- [Sasaki, 2017] Sasaki, Y. (2017). Generalized Nash equilibrium with stable belief hierarchies in static games with unawareness. *Annals of Operations Research*, 256(2):271–284.

- [Sasaki and Kijima, 2012] Sasaki, Y. and Kijima, K. (2012). Hypergames and bayesian games: A theoretical comparison of the models of games with incomplete information. *Journal of Systems Science and Complexity*, 25(4):720–735.
- [Schipper, 2014] Schipper, B. C. (2014). Unawareness - A gentle introduction to both the literature and the special issue. *Mathematical Social Sciences*, 70:1–9.
- [Schipper, 2017] Schipper, B. C. (2017). Self-confirming games: Unawareness, discovery, and equilibrium. In *TARK*.
- [Schipper, 2018] Schipper, B. C. (2018). Discovery and Equilibrium in Games with Unawareness. Mpra paper, University Library of Munich, Germany.
- [Schmeidler, 1973] Schmeidler, D. (1973). Equilibrium points of nonatomic games. *Journal of Statistical Physics*, 7(4):295–300.
- [Seregina et al., 2017] Seregina, T., Brun, O., Azouzi, R. E., and Prabhu, B. J. (2017). On the design of a reward-based incentive mechanism for delay tolerant networks. *IEEE Transactions on Mobile Computing*, 16(2):453–465.
- [Shalizi, 2009] Shalizi, C. (2009). Dynamics of bayesian updating with dependent data and misspecified models. *arXiv: Statistics Theory*.
- [Shiryayev, 1984] Shiryayev, A. (1984). *Mathematical Foundations of Probability Theory*. Springer, New York, NY.
- [Shoham and Leyton-Brown, 2008] Shoham, Y. and Leyton-Brown, K. (2008). *Multiagent Systems: Algorithmic, Game-Theoretic, and Logical Foundations*. Cambridge University Press, USA.
- [Spiegler, 2016] Spiegler, R. (2016). Bayesian Networks and Boundedly Rational Expectations *. *The Quarterly Journal of Economics*, 131(3):1243–1290.
- [Stanford, 1996] Stanford, W. (1996). The limit distribution of pure strategy nash equilibria in symmetric bimatrix games. *Math. Oper. Res.*, 21(3):726–733.
- [Strang, 2009] Strang, G. (2009). *Introduction to Linear Algebra*. Wellesley-Cambridge Press.
- [Syrkkanis et al., 2015] Syrgkanis, V., Agarwal, A., Luo, H., and Schapire, R. E. (2015). Fast convergence of regularized learning in games. In *Advances in Neural Information Processing Systems 28: Annual Conference on Neural Information Processing Systems 2015*, pages 2989–2997.
- [Takahashi, 2008] Takahashi, S. (2008). The number of pure nash equilibria in a random game with nondecreasing best responses. *Games Econ. Behav.*, 63(1):328–340.

- [Teneketzis and Castañón, 1983] Teneketzis, D. and Castañón, D. (1983). Informational aspects of a class of subjective games of incomplete information: Static case. *Journal of Optimization Theory and Applications*, 54:413–422.
- [Thadden and Zhao, 2014] Thadden, E.-L. and Zhao, X. (2014). Multi-task agency with unawareness. *Theory and Decision*, 77(2):197–222.
- [Turrini, 2016] Turrini, P. (2016). Endogenous games with goals: side-payments among goal-directed agents. *Auton. Agents Multi Agent Syst.*, 30(5):765–792.
- [Tuyls et al., 2018] Tuyls, K., Pérolat, J., Lanctot, M., Leibo, J. Z., and Graepel, T. (2018). A generalised method for empirical game theoretic analysis. In *AAMAS*.
- [van Damme, 1991] van Damme, E. (1991). *Stability and Perfection of Nash Equilibria*. Springer-Verlag.
- [Vane and Lehner, 2002] Vane, R. and Lehner, P. (2002). Using hypergames to increase planned payoff and reduce risk. *Autonomous Agents and Multi-Agent Systems*, 5(3):365–380.
- [Varsos et al., 2022] Varsos, C., Fasoulakis, M., Flouris, G., and Bitsaki, M. (2022). Coordination mechanisms with misinformation. In *Proceedings of the 14th International Conference on Agents and Artificial Intelligence - Volume 1: ICAART*,, pages 237–244. SciTePress.
- [Varsos et al., 2021] Varsos, C., Flouris, G., Bitsaki, M., and Fasoulakis, M. (2021). A study of misinformation games. In *PRICAI 2021: Trends in Artificial Intelligence*, pages 76–87. Springer International Publishing.
- [von Neumann, 1928] von Neumann, J. (1928). Zur theorie der gesellschaftsspiele. *Mathematische Annalen*, 100:295–320.
- [Wardrop, 1952] Wardrop, J. G. (1952). Some theoretical aspects of road traffic research. In *Institute of Civil Engineers, Part II, 1*, pages 325–378.
- [Wei et al., 2021] Wei, C.-Y., Lee, C.-W., Zhang, M., and Luo, H. (2021). Linear last-iterate convergence in constrained saddle-point optimization. In *9th International Conference on Learning Representations ICLR '21*.
- [Wu et al., 2020] Wu, M., Amin, S., and Ozdaglar, A. E. (2020). Value of information in bayesian routing games.
- [Zamir, 2009] Zamir, S. (2009). *Bayesian Games: Games with Incomplete Information*. Springer New York.

Appendix A

Publications

The research activity related to this thesis has so far produced the following publications:

- (1) Forward Looking Best-Response Multiplicative Weights Update Methods for Bilinear Zero-sum Games, [AISTATS 2022]
<https://arxiv.org/abs/2106.03579>
- (2) A Study of Misinformation Games, [PRICAI 2021]
https://link.springer.com/chapter/10.1007/978-3-030-89188-6_6
- (3) Adaptation Procedure in Misinformation Games [preprint]
- (4) A Study on Noisy Games [preprint]
- (5) Coordination mechanisms with misinformation [ICAART 2022]
<https://www.scitepress.org/ProceedingsDetails.aspx?ID=PKHqtxzij4c=&t=1>
- (6) Epistemic Adaptive Evolution in Misinformation Games [preprint]

Appendix B

Supplementary material

B.1 Basic Probability theory

We provide here some basic knowledge on probability theory that will be useful in the following sections. The interested reader is referred to [Shiryayev, 1984] for further details.

A random variable X is characterized by its *probability density function (pdf)*, denoted by f_X , which represents the “intensity” of the probability in each given point. The pdf can be used to compute the probability that X falls within a given range, say $[a, b]$, for any $a \leq b$. Formally, f_X is such that:

$$\mathcal{P}[a \leq X \leq b] = \int_a^b f_X(x) dx$$

We denote by F_X the *cumulative distribution function (CDF)* of a random variable X , which equals the probability that the value of X is at most x . Formally:

$$F_X(x) = \int_{-\infty}^x f_X(t) dt = \mathcal{P}[X \leq x]$$

In this document, we focus on random variables X following the normal distribution, denoted by $X \sim \mathcal{N}(\mu, d^2)$ (for some mean $\mu \in \mathbb{R}$ and standard deviation* $d > 0$). For the special case where $\mu = 0, d = 1$ (i.e., when $X \sim \mathcal{N}(0, 1)$), we get the *standard normal distribution*, with the following pdf (ϕ) and cdf (Φ):

$$\begin{aligned} \phi(x) &= \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \\ \Phi(x) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt \end{aligned} \tag{B.1}$$

*In probability theory, the standard deviation is typically denoted by σ . To avoid confusion with the strategies of normal form games which use the same symbol (see Subsection ??), we use d as a symbol for standard deviation in the thesis.

For the general case, where $X \sim \mathcal{N}(\mu, d^2)$, the pdf and cdf are:

$$\begin{aligned} f_X(x) &= \frac{1}{d} \phi\left(\frac{x-\mu}{d}\right) = \frac{1}{d\sqrt{2\pi}} e^{-\left(\frac{x-\mu}{d\sqrt{2}}\right)^2} \\ F_X(x) &= \Phi\left(\frac{x-\mu}{d}\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x-\mu}{d}} e^{-\frac{t^2}{2}} dt \end{aligned} \quad (\text{B.2})$$

It has been shown that, if $X_i \sim \mathcal{N}(\mu_i, d_i^2)$, $c_0, c_i \in \mathbb{R}$, then:

$$c_0 + \sum c_i X_i \sim \mathcal{N}\left(c_0 + \sum c_i \mu_i, \sum c_i^2 d_i^2\right) \quad (\text{B.3})$$

Given two events A, B , the symbol $\mathcal{P}[A|B]$ denotes the *conditional probability of A given B*, which amounts to the probability that A is true under the condition that B is true.

When combining events, the following are true:

$$\begin{aligned} \text{General Conjunction Rule: } \mathcal{P}[A \wedge B] &= \mathcal{P}[B] \mathcal{P}[A|B] = \mathcal{P}[A] \mathcal{P}[B|A] \\ \text{Restricted Conjunction Rule: } \mathcal{P}[A \wedge B] &= \mathcal{P}[A] \cdot \mathcal{P}[B] \\ &\quad (\text{when } A, B \text{ are independent}) \\ \text{General Disjunction Rule: } \mathcal{P}[A \vee B] &= \mathcal{P}[A] + \mathcal{P}[B] - \mathcal{P}[A \wedge B] \\ \text{Restricted Disjunction Rule: } \mathcal{P}[A \vee B] &= \mathcal{P}[A] + \mathcal{P}[B] \\ &\quad (\text{when } A, B \text{ are mutually exclusive}) \end{aligned} \quad (\text{B.4})$$

B.2 Kullback-Leibler divergence

To measure the difference between two probability distributions over the same variable x , we use a measure, called the *Kullback-Leibler divergence*, or simply, the *KL divergence*, (see p.361 [Boyd and Vandenberghe, 2004]). The KL divergence, closely related to relative entropy, information divergence, and information for discrimination, is a non-symmetric measure of the difference between two probability distributions $p(x)$ and $q(x)$. Formally,

Definition 55. Let $p(x)$ and $q(x)$ are two probability distributions of a discrete random variable x . The Kullback-Leibler (KL) divergence of $q(x)$ from $p(x)$, denoted $D_{KL}(p(x), q(x))$, is a measure of the information lost when $q(x)$ is used to approximate $p(x)$. $D_{KL}(p(x), q(x))$ is defined as

$$D_{KL}(p(x), q(x)) = \sum_{x \in X} p(x) \ln \frac{p(x)}{q(x)} \quad (\text{B.5})$$

B.3 Auxiliary definitions and results

Definition 56. *The Hamming distance between two equal-length strings of symbols is the number of positions at which the corresponding symbols are different.*

Theorem 9. *(Theorem 2 of [Schmeidler, 1973]) Consider a non-atomic congestion game. Suppose that the cost functions satisfy: $c_r^i(\cdot)$ is continuous for every arc r and user i , and $i \mapsto c_r^i(x)$ is measurable for every $x \in [0, r]$. Then there exists a Nash equilibrium.*

Lemma 18. *(Lemma C.3 of [Mertikopoulos et al., 2018]) Let $R \in \mathbb{R}^{m \times n}$ be a game matrix for a two-player zero-sum game with value ρ . Then exists a Nash equilibrium (x^*, y^*) such that*

$$\begin{aligned} (Ry^*)i &= \rho & \forall i \in \text{supp}(x^*) \\ (Ry^*)i &< \rho & \forall i \notin \text{supp}(x^*) \\ (R^T x^*)i &= \rho & \forall i \in \text{supp}(y^*) \\ (R^T x^*)i &> \rho & \forall i \notin \text{supp}(y^*) \end{aligned}$$

Example B.1. *(Water-filling solution. Example 5.2 in [Boyd and Vandenberghe, 2004]) We consider the convex optimization problem*

$$\begin{aligned} \text{minimize} \quad & -\sum_{i=1}^n \log(\alpha_i + x_i) \\ \text{subject to} \quad & x \geq 0, \mathbf{1}^T x = 1, \end{aligned}$$

where $\alpha_i > 0$. This problem arises in information theory, in allocating power to a set of n communication channels. The variable x_i represents the transmitter power allocated to the i^{th} channel, and $\log(\alpha_i + x_i)$ gives the capacity or communication rate of the channel, so the problem is to allocate a total power of one to the channels, in order to maximize the total communication rate.

Introducing Lagrange multipliers $\lambda^* \in \mathbb{R}^n$ for the inequality constraints $x^* \geq 0$, and a multiplier $\nu^* \in \mathbb{R}$ for the equality constraint $\mathbf{1}^T x = 1$, we obtain the KKT conditions

$$x^* \geq 0, \mathbf{1}^T x = 1, \lambda^* \geq 0, \lambda^* x_i^* = 0 \quad i = 1, \dots, n$$

and

$$-\frac{1}{\alpha_i + x_i^*} - \lambda_i^* + \nu^* = 0, \quad i = 1, \dots, n$$

We can directly solve these equations to find x^* , λ^* , and ν^* . We start by noting that λ^* acts

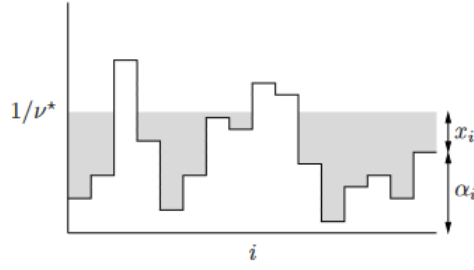


Figure B.1: Illustration of water-filling algorithm. The height of each patch is given by α_i . The region is flooded to a level $1/\nu^*$ which uses a total quantity of water equal to one. The height of the water (shown shaded) above each patch is the optimal value of x_i^* .

as a slack variable in the last equation, so it can be eliminated, leaving

$$x^* \geq 0, \quad \mathbf{1}^T x = 1, \quad x_i^* \left(\nu^* - \frac{1}{\alpha_i + x_i^*} \right) = 0, \quad i = 1, \dots, n$$

Thus,

$$\nu^* \geq \frac{1}{\alpha_i + x_i^*}, \quad i = 1, \dots, n$$

If $\nu^* \geq \frac{1}{\alpha_i}$, this last condition can only hold if $x_i^* > 0$, which by the third condition implies that $\nu^* = \frac{1}{\alpha_i + x_i^*}$. Solving for x_i^* , we conclude that $x_i^* = \frac{1}{\nu^*} - \alpha_i$ if $\nu^* < \frac{1}{\alpha_i}$. If $\nu^* \geq \frac{1}{\alpha_i}$, then $x_i^* > 0$ is impossible, because it would imply $\nu^* \geq \frac{1}{\alpha_i} > \frac{1}{\alpha_i + x_i^*}$, which violates the complementary slackness condition. Therefore, $x_i^* = 0$ if $\nu^* \geq \frac{1}{\alpha_i}$. Thus we have

$$x_i^* = \begin{cases} \frac{1}{\nu^*} - \alpha_i, & \nu^* < \frac{1}{\alpha_i} \\ 0, & \nu^* \geq \frac{1}{\alpha_i} \end{cases}$$

or, put more simply, $x_i^* = \max\{0, 1/\nu^* - \alpha_i\}$. Substituting this expression for x_i^* into the condition $\mathbf{1}^T x = 1$ we obtain

$$\sum_{i=1}^n \max\{0, 1/\nu^* - \alpha_i\} = 1$$

The lefthand side is a piecewise-linear increasing function of $1/\nu^*$, with breakpoints at α_i , so the equation has a unique solution which is readily determined. This solution method is called water-filling for the following reason. We think of α_i as the ground level above patch i , and then flood the region with water to a depth $1/\nu$, as illustrated in Figure B.1. The total

amount of water used is then

$$\sum_{i=1}^n \max\{0, \frac{1}{\nu^*} - \alpha_i\}.$$

We then increase the flood level until we have used a total amount of water equal to one. The depth of water above patch i is then the optimal value x_i^* . \square

B.4 Jacobian matrix

Recall the form of the Jacobian of our dynamical system in Equation (6.6).

We compute the form of each entry of J at the point (x, y) . Let $Q_x = \sum_{\ell} x_{\ell} e^{\eta e_{\ell}^T R f(x, y)}$, $Q_y = \sum_{\ell} y_{\ell} e^{-\eta e_{\ell}^T R^T h(x, y)}$, $S_x = \sum_{\ell} x_{\ell} e^{\xi e_{\ell}^T R y}$, and $S_y = \sum_{\ell} y_{\ell} e^{-\xi e_{\ell}^T R^T x}$.

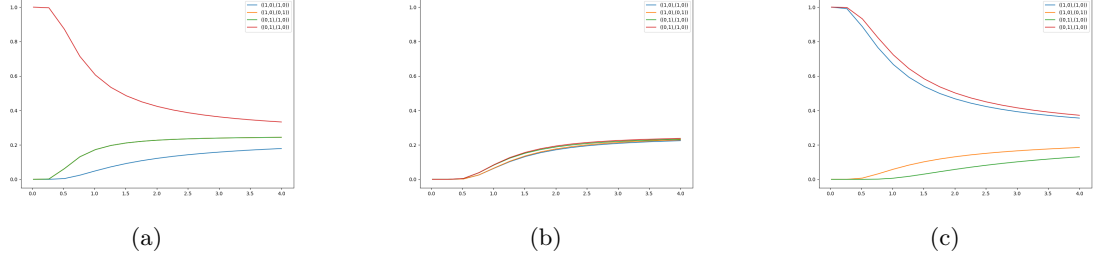
$$\begin{aligned} \frac{\partial \varphi_{1,i}}{\partial x_i} &= e^{\eta e_i^T R f(x, y)} \frac{Q_x \left(1 + \eta x_i \frac{\partial}{\partial x_i} (e_i^T R f(x, y)) \right) - x_i \frac{\partial}{\partial x_i} Q_x}{Q_x^2}, \quad i \in [n], \\ \frac{\partial \varphi_{1,i}}{\partial x_j} &= x_i e^{\eta e_i^T R f(x, y)} \frac{\eta Q_x \frac{\partial}{\partial x_j} (e_i^T R f(x, y)) - \frac{\partial}{\partial x_j} Q_x}{Q_x^2}, \quad i, j \in [n] \text{ and } i \neq j, \\ \frac{\partial \varphi_{1,i}}{\partial y_j} &= x_i e^{\eta e_i^T R f(x, y)} \frac{\eta Q_x \frac{\partial}{\partial y_j} (e_i^T R f(x, y)) - \frac{\partial}{\partial y_j} Q_x}{Q_x^2}, \quad i, j \in [n], \\ \frac{\partial \varphi_{2,i}}{\partial x_j} &= y_i e^{-\eta e_i^T R^T h(x, y)} \frac{-\eta Q_y \frac{\partial}{\partial x_j} (e_i^T R^T h(x, y)) - \frac{\partial}{\partial x_j} Q_y}{Q_y^2}, \quad i, j \in [n], \\ \frac{\partial \varphi_{2,i}}{\partial y_i} &= e^{-\eta e_i^T R^T h(x, y)} \frac{Q_y \left(1 - \eta y_i \frac{\partial}{\partial y_i} (e_i^T R^T h(x, y)) \right) - y_i \frac{\partial}{\partial y_i} Q_y}{Q_y^2}, \quad i \in [n], \\ \frac{\partial \varphi_{2,i}}{\partial y_j} &= y_i e^{-\eta e_i^T R^T h(x, y)} \frac{-\eta Q_y \frac{\partial}{\partial y_j} (e_i^T R^T h(x, y)) - \frac{\partial}{\partial y_j} Q_y}{Q_y^2}, \quad i, j \in [n] \text{ and } i \neq j. \end{aligned} \tag{B.6}$$

At the point (x^*, y^*) , after exploiting the fact that this is an equilibrium profile, and simplifying some of the calculations, we obtain the following forms.

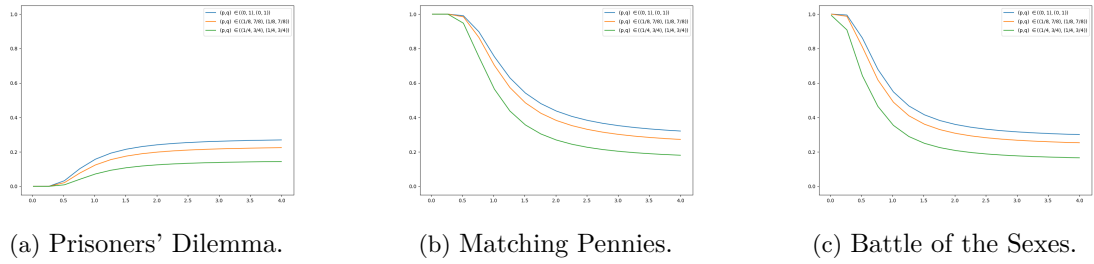
$$\begin{aligned}
\frac{\partial \varphi_{1,i}}{\partial x_i} &= 1 - x_i^* \left(\eta \xi \left(\sum_k R_{ik}^2 y_k^* - \sum_k x_k^* \sum_l R_{kl} y_l^* R_{li}^T \right) + 1 \right), \quad i \in \text{supp}(x^*), \\
\frac{\partial \varphi_{1,i}}{\partial x_i} &= \frac{e^{\eta e_i^T R y^*}}{e^{\eta v}}, \quad i \notin \text{supp}(x^*), \\
\frac{\partial \varphi_{1,i}}{\partial x_j} &= -x_i^* \left(\eta \xi \left(\sum_k R_{ik} y_k^* R_{kj}^T - \sum_k x_k^* \sum_l R_{kl} y_l^* R_{lj}^T \right) + 1 \right), \quad i \in \text{supp}(x^*), i \neq j, \\
\frac{\partial \varphi_{1,i}}{\partial x_j} &= 0, \quad i \notin \text{supp}(x^*) \text{ and } i \neq j, \\
\frac{\partial \varphi_{1,i}}{\partial y_j} &= x_i^* \eta (R_{ij} - e_j^T R^T x^*) \frac{e^{-\xi e_j^T R^T x^*}}{e^{-\xi v}}, \quad \text{for all } i \in \text{supp}(x^*), \\
\frac{\partial \varphi_{1,i}}{\partial y_j} &= 0, \quad i \notin \text{supp}(x^*), \\
\frac{\partial \varphi_{2,i}}{\partial x_j} &= -y_i^* \eta (R_{ij}^T - e_j^T R y^*) \frac{e^{\xi e_j^T R y^*}}{e^{\xi v}}, \quad \text{for all } i \in \text{supp}(y^*), \\
\frac{\partial \varphi_{2,i}}{\partial x_j} &= 0, \quad i \notin \text{supp}(y^*), \\
\frac{\partial \varphi_{2,i}}{\partial y_i} &= 1 - y_i^* \left(\eta \xi \left(\sum_k (R_{ik}^T)^2 x_k^* - \sum_k y_k^* \sum_l R_{kl}^T x_l^* R_{li} \right) + 1 \right), \quad i \in \text{supp}(y^*), \\
\frac{\partial \varphi_{2,i}}{\partial y_i} &= \frac{e^{-\eta e_i^T R^T x^*}}{e^{-\eta v}}, \quad i \notin \text{supp}(y^*), \\
\frac{\partial \varphi_{2,i}}{\partial y_j} &= -y_i^* \left(\eta \xi \left(\sum_k R_{kj} R_{ik}^T x_k^* - \sum_k y_k^* \sum_l R_{kl}^T x_l^* R_{lj} \right) + 1 \right), \quad i \in \text{supp}(y^*), i \neq j, \\
\frac{\partial \varphi_{2,i}}{\partial y_j} &= 0, \quad i \notin \text{supp}(y^*) \text{ and } i \neq j,
\end{aligned} \tag{B.7}$$

B.5 Numerical experiments in noisy games

In this section, we demonstrate further numerical results regarding the formulas in noisy games. Particularly, we consider the case where: i) a player has only pure equilibrium strategy, ii) a player has only a mixed equilibrium strategy, and iii) a player has only mixed equilibrium strategy within a tolerance range, predicates OP_x^G , OM_x^G , and ROM_x^G respectively, see Section 7.3. We examine these settings for the benchmark games in Figure 7.3. As Prisoner's Dilemma and Win-Win are both games with dominated pure strategies, their experimental behavior is similar, thus we provide results only for Prisoner's Dilemma game.

B.5.1 Case OP_x^G Figure B.2: Cases OP_x^G : (a) Prisoners' Dilemma, (b) Matching Pennies, and (c) Battle of the Sexes.

Observe that in the case where $D < 0.5$ the strategic behavior of the players are similar enough to their behavior in the games without noise. On the other hand, as D increases players start to deviate from the un-misinform behavior, meaning that the noise affects their choices. Finally, for large enough D players play randomly.

B.5.2 Case OM_x^G Figure B.3: Cases OM_x^G : (a) Prisoners' Dilemma, (b) Matching Pennies, and (c) Battle of the Sexes.

In Figure B.4a the mixed strategy profiles have small probability to occur for $D < 0.5$, as initially, the distortion provided by the noise does not alternate the strategic behavior of the players. Nevertheless, as D increases players may end up deciding on a mixed strategy profile.

In Figure B.4b, for small values of D , there is a high probability for the players to end up in a mixed strategy profile. As D increases, these probabilities decrease. Thus, the noise made less possible the occurrence of mixed strategy profiles, in the case of the Matching

Pennies game.

Finally, in Figure B.4c, the strategic behaviour of the player is affected similarly as in the Matching Pennies case. This happens because in the Battle of the Sexes a mixed strategy profile is “by default”, without noise, in the ease of the players.

B.5.3 Case ROM_x^G

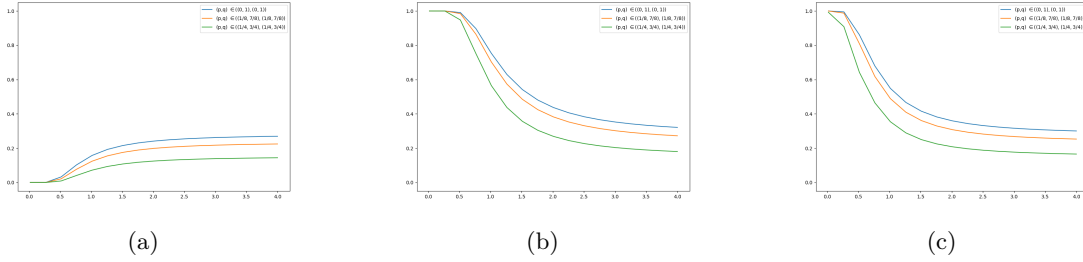


Figure B.4: Cases ROM_x^G : (a) Prisoners' Dilemma, (b) Matching Pennies, and (c) Battle of the Sexes.

Observe that in the case where $D < 0.5$ the strategic behavior of the players are similar enough to their behavior in the games without noise. On the other hand, as D increases players start to deviate from the un-misinformation behavior, meaning that the noise affects their choices. Finally, for large enough D players play randomly.

Index

- VA-replacement, 65
- ε -closeness, 96
- ε -close, 97
- k -neighbor, 54
- Adaptation Procedure, 43, 46
 - length, 47
- adaptive evolution, 60, 65
- adaptive update, 56
- branch, 51
- characteristic strategy set, 45
- Epistemic Adaptive Evolution, 44
- epistemic adaptive evolution, 67
- epistemic evolution, 56, 60, 63
- epistemic revision, 62
- equilibrium
 - ε -Nash, 22
 - epistemic natural, 56, 60
 - minmax, 22
 - Nash, 20
 - natural misinformed, 30, 39
 - pseudo misinformed, 31
 - stable epistemic natural, 56, 67
 - stable misinformed, 44, 48
- function
 - payoff, 19, 30
 - social welfare, 20
- game
 - actual, 28
 - Battle of the Sexes, 32, 131
 - Bayesian, 3
 - bimatrix, 20
 - epistemic, 59
 - epistemic misinformation, 59
 - load balancing (LBG), 25
 - Matching Pennies, 131
 - merged, 31
 - misinformation, 28
 - canonical, 28
 - non-canonical, 29
 - normal form, 19
 - Prisoner's Dilemma, 131
 - Prisoners' Dilemma, 27
 - single-commodity non-atomic congestion (NACG), 25
 - subjective epistemic, 59
 - value, 21
 - Win-Win, 131
 - zero-sum, 21
- Graph
 - AP, 48
 - API, 49
- Hamming distance, 171
- inflation, 29
- Kullback-Leibler divergence, 71, 170
- method
 - Extra-gradient, 23
 - Multiplicative Weights Update, 23
 - Optimistic Mirror Descent, 15

- Optimistic Multiplicative Weights Update, 23
- metric
 - Biased Price of Anarchy, 40
 - Price of Anarchy, 21
 - Price of Misinformation, 32, 39
- misinformation
 - game, 28
 - normal-form game, 28
- misinformed
 - equilibrium strategy, 30
 - Prisoners' Dilemma, 27
 - strategy, 29
 - strategy profile, 30
- multi-agent systems, 1
- neighbourhood, 54
- neighbours, 54
- position, 19
 - impossible, 54
 - possible, 54
- Predicate
 - Infinite-Nash, 101
 - Only-mixed, 101
 - Only-pure, 101
 - Pure-and-mixed, 101
 - Ranged-only-mixed, 101
 - Ranged-pure-and-mixed, 101
- utility gain, 99
- water-filling, 172

