Εθνικό Μετσόβιο Πολυτεχνείο.
Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών.
Τομέας Σημάτων, Ελέγχου και Ρομποτικής.
Εργαστήριο Συστημάτων Αυτομάτου Ελέγχου

Δυναμικά Παίγνια με Μεγάλο Αριθμό Παικτών: Τυχαία Είσοδος Παικτών και Τυχαίες Αλληλεπιδράσεις

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

του

Ιωάννη Κορδώνη

Επιβλέπων Καθηγητής: Γ. Π. Παπαβασιλόπουλος

Αθήνα, Ιούνιος 2015

Εθνικό Μετσόβιο Πολυτεχνείο.

Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών.

Τομέας Σημάτων, Ελέγχου και Ρομποτικής.

Εργαστήριο Συστημάτων Αυτομάτου Ελέγχου

Δυναμικά Παίγνια με Μεγάλο Αριθμό Παικτών: Τυχαία Είσοδος Παικτών και Τυχαίες Αλληλεπιδράσεις

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ του Ιωάννη Κορδώνη

Επταμελής Επιτροπή:

Γ. Π. Παπαβασιλόπουλος (Επιβλέπων),

Π. Μαραγκός (μέλος τριμελούς),

Ν. Καρκανιάς (μέλος τριμελούς),

Ι. Τσινιάς,

Κ. Βουρνάς

Τ. Κουσιουρής

Β. Παπανικλάου

Αθήνα, Ιούνιος 2015

Περίληψη

Η διατριβή αυτή εξετάζει ορισμένα θεωρητικά θέματα στη θεωρία των Δυναμικών Παιγνίων, έχοντας ως κίνητρο και πιθανή περιοχή εφαρμογής τη μοντελοποίηση των αγορών ηλεκτρικής ενέργειας και των ευφυών δικτύων. Η διατριβή χωρίζεται σε τρία μέρη.

Πρώτο Μέρος: Αρχικά παρουσιάζονται κάποια αποτελέσματα για συστήματα με γραμμική δυναμική και Μαρκοβιανά άλματα (MJLS), στα οποία η αλυσίδα Markov έχει ένα γενικό χώρο κατάστασης, επεκτείνοντας τα αντίστοιχα αποτελέσματα της βιβλιογραφίας για διακριτό χώρο κατάστασης (πεπερασμένο ή αριθμήσιμα άπειρο). Συγκεκριμένα, χαρακτηρίζεται η μέση τετραγωνική ευστάθεια των συστημάτων αυτών και επιλύεται το πρόβλημα του Γραμμικού Τετραγωνικού (LQ) ελέγχου σε πεπερασμένο και άπειρο χρονικό ορίζοντα.

Στη συνέχεια, αναλύονται μοντέλα Δυναμικών Παιγνίων στα οποία παρουσιάζεται τυχαία είσοδος παικτών. Συγκεκριμένα, θεωρούμε ένα παίκτη με άπειρο χρονικό ορίζοντα (major player) και σε κάθε χρονική στιγμή ένα τυχαίο αριθμό από παίκτες με πεπερασμένους χρονικούς ορίζοντες (minor players), των οποίων η είσοδος περιγράφεται από μια αλυσίδα Markov. Η ανάλυση γίνεται σε ένα Γραμμικό Τετραγωνικό πλαίσιο. Μελετάται και χαρακτηρίζεται η ισορροπία Nash με χρήση συζευγμένων εξισώσεων τύπου Riccati για MJLS, ενώ εμφαση δίνεται στην περίπτωση παιγνίων με πολύ μεγάλο αριθμό minor players.

Δεύτερο μέρος: Στο μέρος αυτό αναλύονται Στατικά και Δυναμικά παιχνίδια στα οποία οι συμμετέχοντες αλληλεπιδρούν πάνω σε κάποιο μεγάλο Δίκτυο. Θεωρούμε ότι οι παίκτες δεν έχουν πλήρη γνώση των χαρακτηριστικών του δικτύου των αλληλεπιδράσεων, ούτε των προτιμήσεων των παικτών που συμμετέχουν. Αντ' αυτού θεωρούμε ότι διαθέτουν στατιστικές πληροφορίες για το δίκτυο αλληλεπιδράσεων, καθώς και κάποιες τοπικές πληροφορίες. Ορισμένες έννοιες από τη Στατιστική Φυσική χρησιμοποιούνται για να οριστεί μια έννοια κατά πιθανότητα προσεγγιστικής ισορροπίας σε παιχνίδια με μεγάλο αριθμό παικτών, ενώ ορίζεται και μια έννοια πληροφοριακής πολυπλοκότητας. Αναλύονται, τέλος, διάφορα παραδείγματα παιγνίων με αλληλεπιδράσεις σε μεγάλα δίκτυα, όπως Στατικά και Γραμμικά Τετραγωνικά Δυναμικά Παιχνίδια σε τυχαία γραφήματα τύπου Erdos-Renyi, Στατικά Τετραγωνικά Παιχνίδια σε πλέγματα και Γραμμικά Τετραγωνικά παιχνίδια σε δακτυλίους.

Τρίτο μέρος: Στο τελευταίο μέρος της διατριβής μελετάται η δυνατότητα εξαπάτησης σε καταστάσεις στρατηγικής δυναμικής αλληλεπίδρασης χωρίς πλήρη δομική πληροφορία, στην περίπτωση που οι παίκτες χρησιμοποιούν στρατηγικές προσαρμογής/εκμάθησης. Παράδειγμα τέτοιας αλληλεπίδρασης είναι ο ανταγωνισμός μεταξύ παραγωγών ηλεκτρικής ενέργειας, όπου ο κάθε παραγωγός γνωρίζει το δικό του κόστος, ενώ δεν γνωρίζει το κόστος των άλλων παραγωγών.

Αρχικά, διατυπώνονται κριτήρια για την αξιολόγηση των δυναμικών κανόνων. Στη συνέχεια, επικεντρώνουμε σε μια υποκατηγορία στρατηγικών εξαπάτησης τις οποίες καλούμε στρατηγικές υποκρισίας και αναλύουμε κάποια πιθανά αποτελέσματα της στρατηγικής αλληλεπίδρασης, όταν ένας ή περισσότεροι παίκτες υποκρίνονται. Προκύπτει ότι στην περίπτωση που ένας μόνο παίκτης υποκρίνεται και υπάρχει αρκετή αβεβαιότητα, τότε το αποτέλεσμα είναι ίδιο με την περίπτωση που ο παίκτης που υποκρίνεται ήταν Stackelberg αρχηγός. Επίσης προκύπτει ότι σε παιχνίδια με πολλούς, περίπου ισοδύναμους, παίκτες το όφελος από την υποκρισία είναι μικρό. Τέλος, μελετάμε εφαρμογές σε μοντέλα αγορών ηλεκτρικής ενέργειας και προσδιορίζονται περιπτώσεις στις οποίες η εξαπάτηση ενισχύει τη συνεργασία μεταξύ των παικτών και άλλες στις οποίες ενισχύει τον ανταγωνισμό.

Abstract

We study some theoretical topics on the theory of Dynamic Games, having as motivation and possible application area the modeling of Electricity Markets and the Smart Grid. The thesis is divided into three parts.

First Part: At first, some results on the theory of Markov Jump Linear Systems (MJLS), in which the Markov chain has a general state space are presented, extending the existing literature for discrete (finite or countably infinite) state space. Particularly, the mean square stability of the MJLS is characterized and the Linear Quadratic (LQ) control problems for the finite and infinite time horizon are solved, using appropriate Riccati type equations.

We then analyze Dynamic Games in which there is a random entrance of players. Particularly, we consider an infinite time horizon player called the major player interacting with a random number of minor players having finite time horizons, the entrance of whom is governed by a Markov chain. The analysis is made in a LQ framework. The Nash equilibria are characterized using a set of coupled Riccati type equations for MJLS. An emphasis is paid on the large number of players case, in which the Mean Field (MF) approximation is used.

Second Part: In this part, Static and Dynamic games involving agents interacting on a large graph are studied. We assume that the players do not know the graph of interactions precisely nor the other players preferences. Instead, we assume that each player possesses statistical information about the network of interactions, as well as some local information. Some notions from the Statistical Physics domain are modified to define a Probabilistic Approximate Nash (PAN) equilibrium concept. Furthermore, we define an informational complexity notion. Some special cases are then analyzed, involving Static and LQ games on Erdos-Renyi Random Graphs or Small World Networks, Static Quadratic games on Lattices and LQ games on rings.

Third Part: In the last part of the thesis, the possibility of cheating Dynamic rules (such as learning or adaptation), when applied to Repeated or Dynamic Game situations with incomplete structural information, is studied. An example of such a game situation is the repeated reaction of the energy producing firms, where each one does not know precisely the production cost of its opponents.

At first, two criteria to assess the Dynamic rules are stated. Then, we concentrate to a subclass of cheating strategies, called pretenders strategies and study some possible outcomes, when a player or all the players are pretending. If only one player pretends and there is enough uncertainty the outcome would be the same as if the pretending player were the Stackelberg leader. Furthermore, in games with a large number of equivalent players, the gain from pretending is small and the optimal pretended values are close to the actual. Finally, we study applications to Electricity Market models. Cases where pretending enhances cooperation or competition are identified.

Ευχαριστίες

Θέλω να ευχαριστήσω πάρα πολύ τον επιβλέποντα καθηγητή μου κ. Γιώργο Παπαβασιλόπουλο για δύο λόγους. Αφενός, μου προσέφερε ένα περιβάλλον μεγάλης ακαδημαϊκής ελευθερίας το οποίο συνέβαλλε πολύ στη διαμόρφωση της παρούσας εργασίας και μου επέτρεψε να έχω τη χαρά της δημιουργίας και αφετέρου για τις ιδέες που πήρα από αυτόν τα τελευταία πολλά χρόνια. Θα ήθελα επίσης να ευχαριστήσω τα μέλη της Τριμελούς και Επταμελούς επιτροπής της διατριβής αυτής.

Ιδιαίτερες ευχαριστίες οφείλω στο Δρ. Αλέξανδρο Χαλαλαμπίδη για την πολύτιμη βοήθειά, τις συμβουλές, τη συνεργασία στο επικουρικό έργο, τις πολλές συζητήσεις που είχαμε καθώς και τη συνεισφορά του στη μελέτη του τελευταίου θέματος της παρούσας διατριβής. Στο τελευταίο μέρος της διατριβής συνεισέφεραν επίσης και οι υπ. Δρ. Παναγιώτης Κοντογιώργος και Νίκος Χρυσανθόπουλος, τους οποίους επίσης θα ήθελα να ευχαριστήσω.

Θα ήθελα επίσης να ευχαριστήσω τις υπ. Δρ. Ντόρα Καρβουνιάρη και Νεφέλη Κακαβά για τη βοήθεια που πολλές φορές μου προσέφεραν κατά τη διάρκεια της εκπόνησης της διατριβής αυτής.

Σχεδόν αποκλειστική πηγή χρηματοδότησης κατά τη διάρκεια της εκπόνησης της διατριβής αυτής αποτέλεσε υποτροφία από τον Ειδικό Λογαριασμό Κονδυλίων Έρευνας του Ιδρύματος.

Περιεχόμενα

Κεφάλαιο 1

Εισαγωγή

Η διατριβή αυτή μελετά ορισμένα θεωρητικά θέματα στη θεωρία των Δυναμικών Παιγνίων. Κίνητρο και πιθανή περιοχή εφαρμογής αποτελεί η μοντελοποίηση των αγορών ηλεκτρικής ενέργειας και των ευφυών δικτύων. Στο πρώτο μέρος της εργασίας, αναλύονται Δυναμικά Παίγνια στα οποία έχουμε τυχαία είσοδο παικτών, ενώ επιλύονται και ορισμένα συναφή προβλήματα Στοχαστικού Ελέγχου. Στο δεύτερο μέρος, εξετάζονται παιχνίδια στα οποία οι παίκτες αλληλεπιδρούν πάνω σε ένα μεγάλο Δίκτυο και μελετώνται έννοιες προσεγγιστικής ισορροπίας, πληροφορίας και πολυπλοκότητας. Στο τρίτο μέρος εξετάζεται η δυνατότητα εξαπάτησης σε καταστάσεις παιγνίου, στην περίπτωση που οι παίκτες χρησιμοποιούν δυναμικούς κανόνες προσαρμογής/εκμάθησης.

Το Κεφάλαιο 2 μελετά συστήματα με γραμμική δυναμική και Μαρκοβιανά άλματα (MJLS), στα οποία η αλυσίδα Markov έχει ένα γενικό χώρο κατάστασης, επεκτείνοντας τα αντίστοιχα αποτελέσματα της βιβλιογραφίας για διακριτό χώρο κατάστασης (πεπερασμένο ή αριθμήσιμα άπειρο). Κίνητρο για την μελέτη τέτοιων συστημάτων αποτέλεσε ο χαρακτηρισμός της ισορροπίας σε γραμμικά τετραγωνικά παίγνια με τυχαία είσοδο παικτών. Αρχικά, μελετάται το πρόβλημα της μέσης τετραγωνικής ευστάθειας και επιλύεται το πρόβλημα του Γραμμικού Τετραγωνικού (LQ) ελέγχου σε πεπερασμένο και άπειρο χρονικό ορίζοντα, με τη βοήθεια κατάλληλων εξισώσεων τύπου Riccati. Τα αποτελέσματα αυτά, αν και τεχνικού χαρακτήρα, επεκτείνουν τις εφαρμογές των MJLS. Σαν παράδειγμα, παρουσιάζεται μια εφαρμογή των MJLS με γενικό χώρο κατάστασης σε γραμμικά συστήματα ελεγχόμενα μέσω ενός δικτύου επικοινωνιών.

Στο Κεφάλαιο 3, αναλύονται μοντέλα Δυναμικών Παιγνίων με τυχαία είσοδο παικτών. Συγκεκριμένα, θεωρούμε ένα παίκτη με άπειρο χρονικό ορίζοντα (major player) και σε κάθε χρονική στιγμή ένα τυχαίο αριθμό από παίκτες με πεπερασμένους χρονικούς ορίζοντες (minor players), των οποίων η είσοδος περιγράφεται από μια αλυσίδα Markov. Κίνητρο για τη μελέτη τέτοιων χρονικών αλληλεπιδράσεων σε παίγνια είναι διάφορες πρακτικές καταστάσεις όπως η αλληλεπίδραση ενός δημόσιου φορέα παραγωγής ενέργειας με τους παραγωγούς ανανεώσιμης ενέργειας και η αλληλεπίδραση μιας τράπεζας με τους δανειολήπτες. Η ανάλυση γίνεται σε ένα Γραμμικό Τετραγωνικό πλαίσιο. Έμφαση δίνεται και στην περίπτωση παιγνίων με πολύ μεγάλο αριθμό minor players. Η περίπτωση αυτή αναλύεται με χρήση της προσέγγισης Μέσου Πεδίου (MF) και παρουσιάζονται κάποια αποτελέσματα ύπαρξης και προσεγγιστικής ασυμπτωτικής ισορροπίας.

Στο κεφάλαιο 4 αναλύονται Στατικά και Δυναμικά παιχνίδια στα οποία οι συμμετέχοντες αλληλεπιδρούν πάνω σε κάποιο μεγάλο Δίκτυο. Κίνητρο για τη μελέτη αυτή είναι διάφορες πρακτικές περιπτώσεις στρατηγικής αλληλεπίδρασης πάνω σε μεγάλα δίκτυα, όπως η αλληλεπίδραση των διάφορων παραγωγών και καταναλωτών ή των ευφυών μικρό-δικτύων (smart micro-grids) πάνω στο ηλεκτρικό δίκτυο ή οι στρατηγικές αλληλεπιδράσεις σε κοινωνικά δίκτυα.

Θεωρούμε ότι οι παίκτες δεν έχουν πλήρη γνώση των χαρακτηριστικών του δικτύου των αλληλεπιδράσεων, ούτε των προτιμήσεων των παικτών που συμμετέχουν. Αντ' αυτού θεωρούμε ότι διαθέτουν στατιστικές πληροφορίες για το δίκτυο αλληλεπιδράσεων, καθώς και κάποιες τοπικές πληροφορίες. Κάποιες έννοιες προσεγγιστικής περιγραφής μακροσκοπικών συστημάτων από τη Στατιστική Φυσική χρησιμοποιούνται για να οριστεί μια έννοια προσεγγιστικής ισορροπίας σε παιχνίδια με μεγάλο αριθμό παικτών. Στη συνέχεια ορίζεται μια έννοια πληροφοριακής πολυπλοκότητας, ως η ελάχιστη ποσότητα πληροφορίας που είναι αναγκαία για την ύπαρξη προσεγγιστικής ισορροπίας.

Αναλύονται διάφορα παραδείγματα παιγνίων με αλληλεπιδράσεις σε μεγάλα δίκτυα, όπως Στατικά και Γραμμικά Τετραγωνικά Δυναμικά Παιχνίδια σε τυχαία γραφήματα τύπου Erdos-Renyi, Στατικά Τετραγωνικά Παιχνίδια σε πλέγματα και Γραμμικά Τετραγωνικά παιχνίδια σε δακτυλίους. Οι βασικοί λόγοι για την ύπαρξη χαμηλής πολυπλοκότητας στα παραδείγματα που αναλύθηκαν είναι κάποιοι νόμοι μεγάλων αριθμών, η συστολικότητα της απεικόνισης της βέλτιστης απόκρισης (best response), η συνεργασία μεταξύ των παικτών ή μικρά κέρδη σε απομακρυσμένους παίκτες.

Στο κεφάλαιο 5 μελετώνται καταστάσεις στρατηγικής δυναμικής αλληλεπίδρασης χωρίς πλήρη δομική πληροφορία. Παράδειγμα τέτοιας αλληλεπίδρασης είναι ο ανταγωνισμός μεταξύ παραγωγών ηλεκτρικής ενέργειας, όπου ο κάθε παραγωγός γνωρίζει το δικό του κόστος ενώ δεν γνωρίζει το κόστος των άλλων παραγωγών. Μελετώνται κάποιες στρατηγικές προσαρμογής/εκμάθησης. Τέτοιες στρατηγικές εν γένει δεν είναι σε ισορροπία Nash. Μελετάται, συνεπώς, η δυνατότητα εξαπάτησης, δηλαδή τη δυνατότητα εκμετάλλευσης της γνώσης της στρατηγικής του ενός παίκτη από τους άλλους και προτείνονται κάποια κριτήρια για την αξιολόγηση προσαρμοστικών στρατηγικών.

Στη συνέχεια επικεντρώνουμε σε μια υποκατηγορία στρατηγικών εξαπάτησης τις οποίες καλούμε στρατηγικές υποκρισίας και αναλύουμε κάποια πιθανά αποτελέσματα της στρατηγικής αλληλεπίδρασης, όταν ένας ή περισσότεροι παίκτες υποκρίνονται. Προκύπτει ότι στην περίπτωση που ένας μόνο παίκτης υποκρίνεται και υπάρχει αρκετή αβεβαιότητα, τότε το αποτέλεσμα είναι ίδιο με την περίπτωση που ο παίκτης που υποκρίνεται ήταν Stackelberg αρχηγός. Επίσης προκύπτει ότι σε παιχνίδια με πολλούς, περίπου ισοδύναμους, παίκτες το όφελος από την υποκρισία είναι μικρό. Τέλος, μελετάμε εφαρμογές σε μοντέλα αγορών ηλεκτρικής ενέργειας και προσδιορίζονται περιπτώσεις στις οποίες η εξαπάτηση ενισχύει τη συνεργασία μεταξύ των παικτών και άλλες στις οποίες ενισχύει τον ανταγωνισμό.

Τα αποτέλεσμα της έρευνας που παρουσιάζονται, έχουν δημοσιευτεί στα [KP14a] (Κεφάλαιο 2) και [KP15] (Κεφάλαιο 3), [KP][1](Κεφάλαιο 4). Προκαταρκτικές μορφές των παραπάνω εργασιών ήταν οι [KP11] , [KP12], [KP13b], [KP13a], [KP14b].

Το υπόλοιπο του κεφαλαίου οργανώνεται ακολούθως: Στο Εδάφιο 1.1, γίνεται μια συνοπτική επισκόπηση των βασικών εννοιών της θεωρίας παιγνίων με έμφαση στα δυναμικά παίγνια. Στη συνέχεια, παρουσιάζονται ορισμένες πρόσφατες εξελίξεις στη θεωρία των παιγνίων με μεγάλο αριθμό παικτών, στο Εδάφιο 1.2. Στο Εδάφιο 1.3, παρουσιάζεται η χρήση δυναμικών κανόνων εκμάθησης σε καταστάσεις παιγνίου. Τέλος, στο Εδάφιο 1.4 παρουσιάζονται τα προβλήματα με τα οποία ασχολείται η παρούσα διατριβή, καθώς και συνοπτικά τα κύρια αποτελέσματα.

1.1 Δυναμικά Παίγνια

1.1.1 Περιγραφή

Η θεωρία παιγνίων μελετά μαθηματικά μοντέλα της λήψης αποφάσεων σε καταστάσεις στρατηγικής αλληλεπίδρασης ανάμεσα σε ανεξάρτητες οντότητες (παίκτες, decision makers) [BO99], [Mye13], [Owe69], [OR94], [FT91a], [SLB08]. Η βασική υπόθεση που γίνεται είναι ότι οι παίκτες είναι πλήρως ή μερικά ορθολογιστικοί (rational), δηλαδή η δράση τους μπορεί να περιγραφεί με βάση τη βελτιστοποίηση ή την απόπειρα βελτιστοποίησης κάποιου κριτηρίου (cost function, payoff function, or utility). Έμφαση δίνεται στα δυναμικά παίγνια, δηλαδή στα παίγνια στα οποία οι παίκτες λαμβάνουν αποφάσεις σε διάφορα χρονικά σημεία και η σειρά με την οποία λαμβάνονται οι αποφάσεις είναι σημαντική ([BO99], [SH69b], [SH69a], [SCJ73b],[SCJ73a]).

Τα βασικά στοιχεία για την περιγραφή ενός παιχνιδιού είναι: Το σύνολο των παικτών, το σύνολο των πιθανών δράσεων του κάθε παίκτη, οι συναρτήσεις κόστους, η δυναμική, ο ορίζοντας (ή οι ορίζοντες), δηλαδή το διάστημα στο οποίο οι παίκτες συμμετέχουν στο παιχνίδι και η διαθέσιμη πληροφορία σε κάθε παίκτη.

Στη συνέχεια δίνεται ο ορισμός των ντετερμινιστικών δυναμικών παιγνίων σε διακριτό χρόνο.

Ορισμός 1: Ένα δυναμικό, ντετερμινιστικό παιχνίδι με N παίκτες αποτελείται από:

(i) Ένα σύνολο παικτών $\mathcal{N} = \{1, \ldots, N\}$.

(ii) Ένα σύνολο χρονικών σημείων $\mathbb{T} = \{1, \ldots, T\}$, όπου το T είναι ο ορίζοντας του παιχνιδιού και μπορεί να έχει πεπερασμένη ή άπειρη τιμή.

[1]Under Review

(iii) Ένα χώρο κατάστασης X στον οποίο ανήκει η κατάσταση x_k του παιχνιδιού, κάθε στιγμή k.

(iv) Ένα σύνολο δράσεων U_k^i (πεπερασμένο ή άπειρο) που αποτελείται από τις επιτρεπτές δράσεις του κάθε παίκτη, σε κάθε χρονική στιγμή. Η δράση του παίκτη i τη χρονική στιγμή k συμβολίζεται u_k^i.

(v) Μια συνάρτηση $f_k : X \times U_k^1 \times \cdots \times U_k^N \to X$, έτσι ώστε:

$$x_{k+1} = f_k(x_k, u_k^1, \ldots, u_k^N),$$

η οποία εκφράζει τη δυναμική του παιχνιδιού και μια αρχική συνθήκη $x_1 \in X$.

(vi) Σύνολα Y_k^i που εκφράζουν τις πιθανές παρατηρήσεις του παίκτη i τη χρονική στιγμή k και συναρτήσεις $h_k^i : X \to Y_k^i$ έτσι ώστε η παρατήρηση y_k^i του παίκτη i τη χρονική στιγμή k να δίνεται από:

$$y_k^i = h_k^i(x_k).$$

(vii) Ένα υποσύνολο I_k^i του $\{y_1^1, \ldots, y_k^1, \ldots, y_1^N, \ldots, y_k^N, u_1^1, \ldots, u_{k-1}^1, , \ldots, u_1^N, \ldots, u_{k-1}^N\}$ που εκφράζει τη διαθέσιμη στον παίκτη i πληροφορία τη χρονική στιγμή k.

(viii) Μια κλάση Γ_k^i από επιτρεπτές συναρτήσεις $\gamma_k^i : I_k^i \mapsto u_k^i \in U_k^i$. Η απεικόνιση $\gamma^i = \{\gamma_1^i, \ldots, \gamma_T^i\}$ θα καλείται η στρατηγική του παίκτη i.

(ix) Συναρτήσεις κόστους $J^i : \prod_{k<T}(X \times U_k^1 \times \cdots \times U_k^N) \to \mathbb{R}$.

Η έννοια του μοντέλου αυτού είναι ότι ο κάθε παίκτης i συμπεριφέρεται με τρόπο τέτοιο ώστε να επιχειρεί να ελαχιστοποιήσει το κόστος J^i, επιλέγοντας μια επιτρεπτή στρατηγική γ^i. Υπάρχουν διάφοροι τρόποι να γίνει αυτό, οι οποίοι οδηγούν σε διαφορετικές έννοιες λύσης, κάποιες από τις οποίες παρουσιάζονται στη συνέχεια.

Τα κόστη των παικτών μπορούν να εκφραστούν συναρτήσει των στρατηγικών, αντικαθιστώντας τις τιμές του ελέγχου που προκύπτουν από τις στρατηγικές γ^i στη δυναμική εξίσωση. Σε αυτή την περίπτωση το παιχνίδι έρχεται στη λεγόμενη κανονική μορφή.

Το μοντέλο αυτό μπορεί να επεκταθεί και στη στοχαστική περίπτωση, στην περίπτωση δηλαδή που η δυναμική και τα κόστη εξαρτώνται από κάποιες τυχαίες μεταβλητές. Ένας τρόπος να γίνει αυτό είναι να θεωρηθεί ένας επιπλέον παίκτης, καλούμενος «η φύση», οι δράσεις του οποίου μπορούν να περιγραφούν από κάποιες γνωστές στους παίκτες κατανομές. Τα κόστη, στην περίπτωση αυτή, μπορούν να αντικατασταθούν από τις εκτιμώμενες τιμές τους.

Στην εργασία αυτή θα ασχοληθούμε με προσθετικά κόστη στη μορφή:

$$J^i = \sum_{k<T} g_k^i(x_k, u_k^1, \ldots, u_k^N, x_{k+1})$$

1.1.2 Έννοιες Λύσης

Υπάρχουν διάφορες έννοιες που επιχειρούν να περιγράψουν/προβλέψουν το αποτέλεσμα κάποιας στρατηγικής αλληλεπίδρασης (κατάστασης παιγνίου). Η συνηθέστερη από αυτές είναι η ισορροπία Nash. Προς χάριν απλότητας, στην υποπαράγραφο αυτή δίνονται οι ορισμοί για παίγνια με δύο παίκτες.

Ορισμός 2 (Βέλτιστη Απόκριση): Μια στρατηγική γ^1 αποτελεί βέλτιστη απόκριση σε μια στρατηγική γ^2 του παίκτη 2 αν $\gamma^1 \in \arg\min_{\gamma^1 \in \Gamma^1} J(\gamma^1, \gamma^2)$. Το σύνολο των στρατηγικών του παίκτη 1 που αποτελούν βέλτιστη απόκριση στη στρατηγική γ^1 θα συμβολίζεται με $BR(\gamma^2)$.

Ορισμός 3 (Ισορροπία Nash): Ένα ζευγάρι στρατηγικών $\gamma^{1,N}, \gamma^{2,N}$ είναι σε ισορροπία Nash, αν η καθεμία τους είναι μια βέλτιστη απόκριση στην άλλη δηλαδή, $\gamma^{1,N} \in BR(\gamma^{2,N})$ και $\gamma^{2,N} \in BR(\gamma^{1,N})$.

Σε ένα σύνολο στρατηγικών που αποτελούν ισορροπία Nash, κανένας παίκτης δεν έχει κίνητρο να αλλάξει μονομερώς στρατηγική.

Παρατήρηση 1: Η ισορροπία Nash είναι μια αρκετά φυσική έννοια λύσης. Η ύπαρξη της ισορροπίας Nash έχει αποδειχθεί σε πεπερασμένα παιχνίδια (επεκτείνοντας το χώρο των στρατηγικών για να περιλαμβάνει τις μεικτές στρατηγικές) [N+50] και σε κυρτά παιχνίδια [Ros65]. Στα δυναμικά παιχνίδια η ύπαρξη της ισορροπίας Nash δεν έχει χαρακτηριστεί πλήρως, ενώ υπάρχουν διάφορα μερικά αποτελέσματα.

Σε πολλές περιπτώσεις, έχει ενδιαφέρον να μελετηθούν οι προσεγγιστικές ισορροπίες Nash.

Ορισμός 4: Ένα ζευγάρι στρατηγικών γ^1, γ^2 αποτελούν μια ε-Nash ισορροπία αν ισχύει $J^1(\gamma^1, \gamma^2) < J^1(\tilde{\gamma}^1, \gamma^2) + \varepsilon$ και $J^2(\gamma^1, \gamma^2) < J^2(\gamma^1, \tilde{\gamma}^2) + \varepsilon$, για οποιεσδήποτε στρατηγικές $\tilde{\gamma}^1, \tilde{\gamma}^2$.

Στην ισορροπία Nash, οι ρόλοι των παικτών είναι συμμετρικοί. Σε διάφορες περιπτώσεις, οι παίκτες έχουν ασύμμετρη επιρροή, ή κάποιος είναι ικανός να δεσμευτεί να έχει κάποια μελλοντική συμπεριφορά, ενώ σε κάποιες άλλες περιπτώσεις ο ένας παίκτης δρα πριν από τον άλλο. Σ' αυτές τις περιπτώσεις ο παίκτης που ανακοινώνει πρώτος τη στρατηγική του έχει ένα πλεονέκτημα σε σχέση με την περίπτωση που ανακοινώνουν ταυτόχρονα τις στρατηγικές τους. Ο παίκτης αυτός θα ονομάζεται αρχηγός Stackelberg, ενώ ο άλλος παίκτης ακόλουθος. Στη συνέχεια παρουσιάζεται η έννοια της ισορροπίας Stackelberg, η οποία δίνει μια έννοια λύσης για την πρόβλεψη του αποτελέσματος τέτοιων καταστάσεων στρατηγικής αλληλεπίδρασης.

Ορισμός 5 (Ισορροπία Stackelberg): Ένα ζευγάρι στρατηγικών $\gamma^{1,S_L}, \gamma^{2,S_F}$ αποτελεί ισορροπία Stackelberg , με τον παίκτη 1 αρχηγό, αν $\gamma^{2,S_F} \in BR(\gamma^{1,S_L})$ και $\gamma^{1,S_L} \in \arg\min_{\gamma^1} \{\max_{\gamma^2 \in BR(\gamma^1)} J(\gamma^1, \gamma^2)\}$.

Παρατήρηση 2: Σε όλες τις περιπτώσεις ο αρχηγός στην ισορροπία Stackelberg έχει μικρότερο κόστος σε σχέση με την ισορροπία Nash. Ο ακόλουθος σε κάποιες περιπτώσεις ωφελείται σε σχέση με την ισορροπία Nash, ενώ σε άλλες βλάπτεται. Επιπλέον, είναι πιθανόν ένας παίκτης να προτιμά να είναι ακόλουθος από αρχηγός [Bas73].

Τα αποτελέσματα στις ισορροπίες Nash και Stackelberg, δεν είναι πάντα αποδοτικά. Σε διάφορες περιπτώσεις μπορούν όλοι οι παίκτες, συντονίζοντας τη δράσεις τους, να βελτιώσουν την απόδοσή τους σε σχέση με κάποια ισορροπία. Μια ενδιαφέρουσα έννοια λύσης είναι το σύνολο των σημείων στα οποία δεν είναι δυνατόν να βρεθεί κάποιος συνδυασμός δράσεων έτσι ώστε να μειώνεται το κόστος όλων των παικτών.Το σύνολο των σημείων αυτών λέγεται σύνολο Pareto.

Ορισμός 6 (Σύνολο Pareto): Ένα ζευγάρι στρατηγικών γ^1, γ^2 ανήκει στο σύνολο Pareto αν δεν υπάρχουν $\bar{\gamma}^1, \bar{\gamma}^2$ έτσι ώστε $J^1(\bar{\gamma}^1, \bar{\gamma}^2) \leq J^1(\gamma^1, \gamma^2)$, $J^2(\bar{\gamma}^1, \bar{\gamma}^2) \leq J^2(\gamma^1, \gamma^2)$ και τουλάχιστον μία από τις δύο ανισότητες να ισχύει αυστηρά.

1.1.3 Πληροφορία και Χρονική Συνέπεια

Υπάρχουν διάφοροι πιθανοί τρόποι με τους οποίους οι παίκτες λαμβάνουν πληροφορία. Μια πιθανότητα είναι να γνωρίζουν μόνο την αρχική κατάσταση. Εναλλακτικά, μπορούν να μετρούν το διάνυσμα κατάστασης ή ο κάθε παίκτης να έχει μια διαφορετική μέτρηση μιας συνάρτησης του διανύσματος κατάστασης. Οι διαφορετικές δομές πληροφορίας οδηγούν σε διαφορετικά σύνολα στρατηγικών και συνεπώς σε πιθανόν διαφορετικές έννοιες λύσης. Στον ακόλουθο ορισμό παρουσιάζονται κάποιες δομές πληροφορίας.

Ορισμός 7: Σ' ένα δυναμικό παιχνίδι σε διακριτό χρόνο η δομή πληροφορίας του παίκτη i είναι:

(i) Ανοικτού βρόχου (open loop) αν $I_k^i = \{x_1\}$.

(ii) Κλειστού βρόχου (closed loop) αν $I_k^i = \{x_1, \ldots, x_k\}$.

(iii) Ανάδρασης χωρίς μνήμη (feedback no memory) αν $I_k^i = \{x_k\}$.

(iv) Κλειστού βρόχου με μέτρηση εξόδου (closed loop imperfect state information) αν $I_k^i = \{y_1^i, \ldots, y_k^i\}$.

Διαφορετικές τεχνικές έχουν αναπτυχθεί για τον υπολογισμό των ισορροπιών σε παιχνίδια με πληροφορία ανοικτού ή κλειστού βρόχου (π.χ. [SH69b]). Για παράδειγμα, σε παιχνίδια με πληροφορία ανοικτού βρόχου η αρχή του Pontryagin μπορεί να χρησιμοποιηθεί ενώ για πληροφορία κλειστού βρόχου, ο δυναμικός προγραμματισμός (εξισώσεις HJB για το συνεχή χρόνο).

Όταν χρησιμοποιούνται στρατηγικές που εξαρτώνται από παλαιότερες τιμές του διανύσματος κατάστασης, μπορεί να υπάρξει μια πλειάδα από ισορροπίες ή προσεγγιστικές ισορροπίες (folk theorems ex. [FT91a] ch. 5, [PCJ80], [Bas74], [Rad80], [Pap89a]). Στις εργασίες αυτές η χρήση των προηγούμενων τιμών του διανύσματος κατάστασης ή των δράσεων των άλλων παικτών επιτρέπει την ύπαρξη ενός πολύ μεγάλου συνόλου λύσεων. Σε κάποιες από τις λύσεις ενισχύεται η συνεργασία μεταξύ των παικτών ενώ σε άλλες επιβάλλεται μια λύση που είναι στο απόλυτο συμφέρον κάποιου παίκτη.

Ένα κριτήριο για τις στρατηγικές στα δυναμικά παίγνια είναι η «χρονική συνέπεια». Στον επόμενο ορισμό εισάγονται οι έννοιες της ασθενούς και ισχυρής χρονικής συνέπειας.

Ορισμός 8: Ένα σύνολο στρατηγικών γ^i που αποτελούν ισορροπία σε ένα παιχνίδι είναι:

(i) Ασθενώς χρονικά συνεπές (weakly time consistent) αν για οποιαδήποτε στιγμή $t < T$ οι στρατηγικές γ^i αποτελούν ισορροπία για το παιχνίδι που ξεκινάει τη χρονική στιγμή t, δεδομένου ότι μέχρι τη στιγμή εκείνη είχαν εφαρμοστεί οι στρατηγικές γ^i.

(ii) Ισχυρά χρονικά συνεπές (strongly time consistent) αν για οποιαδήποτε στιγμή $t < T$ οι στρατηγικές γ^i αποτελούν ισορροπία για το παιχνίδι που ξεκινάει τη χρονική στιγμή t, οποιεσδήποτε στρατηγικές και να έχουν εφαρμοστεί μέχρι τη στιγμή t.

Στην εργασία αυτή ασχολούμαστε στρατηγικές που ικανοποιούν το δυναμικό προγραμματισμό και είναι ισχυρά χρονικά συνεπείς. Οι ισορροπίες, αποτελούμενες από τέτοιες στρατηγικές, θα καλούνται στο εξής και «τέλειες» (perfect).

1.2 Παίγνια με μεγάλο αριθμό παικτών : $MFGs$

Στην παράγραφο αυτή γίνεται μια σύντομη επισκόπηση της θεωρίας των παιγνίων με μεγάλο αριθμό παικτών. Το ενδιαφέρον για τέτοια παιχνίδια είναι αρκετά παλιό. Ένα μοντέλο για ισορροπία σε δίκτυα κυκλοφορίας με ένα συνεχές από χρήστες διατυπώθηκε στο [War52], ενώ ένα μοντέλο για αγορές, θεωρώντας ένα συνεχές από παίκτες, παρουσιάστηκε στο [Aum64] (επίσης [Owe69], Ch. X). Στη συνεργατική θεωρία παιγνίων, μελετήθηκαν τα oceanic games [MS78], παιχνίδια με ένα συνεχές από παίκτες και υπολογίστηκαν οι Shapley τιμές τους. Δυναμικά παίγνια με ένα συνεχές από παίκτες σε διακριτό χρόνο μελετήθηκαν στα [JR88] και [BB92]. Η προσέγγιση μέσου πεδίου χρησιμοποιήθηκε για την ανάλυση διακριτών δυναμικών παιγνίων στο [WBVR05] και ορίστηκε η έννοια του Oblivious Equilibrium.

Τα τελευταία χρόνια υπάρχει έντονη ερευνητική δραστηριότητα στα στοχαστικά δυναμικά παίγνια με μεγάλο αριθμό παικτών και έχουν προταθεί δύο αρκετά συναφείς προσεγγίσεις. Η μια ονομάζεται παίγνια μέσου πεδίου (Mean Field Games (MGF)) [LL07], [GLL10] ενώ η δεύτερη Nash Certain Equivalence (NCE) [HMC05], [HMC06].

Στις προσεγγίσεις αυτές ο κάθε παίκτης αλληλεπιδρά με τη μάζα των άλλων παικτών η οποία προσεγγίζεται από ένα συνεχές. Η ιδέα αυτή προέρχεται από την προσέγγιση μέσου πεδίου στη στατιστική φυσική. Κάτω από κατάλληλες προϋποθέσεις, παίρνοντας το όριο καθώς ο αριθμός των παικτών τείνει στο άπειρο, η συμπεριφορά της μάζας των παικτών γίνεται ασυμπτωτικά ντετερμινιστική. Διατυπώνεται, συνεπώς, ένα ζευγάρι από συζευγμένες εξισώσεις που περιγράφουν τη συμπεριφορά των παικτών. Η πρώτη από αυτές εκφράζει τη το γεγονός ότι οι στρατηγική του κάθε παίκτη είναι βέλτιστη, δοσμένης της εξέλιξης του μέσου πεδίου των διανυσμάτων κατάστασης των άλλων παικτών. Στο συνεχή χρόνο είμαι μια εξίσωση Hamilton-Jacobi-Bellman και λύνεται από το τέλος προς την αρχή. Η άλλη εξίσωση περιγράφει την εξέλιξη του μέσου πεδίου των διανυσμάτων κατάστασης των παικτών και υπό κατάλληλες συνθήκες είναι μια εξίσωση Kolmogorov-Focker-Plank.

Η ανάλυση των παιγνίων με μεγάλο αριθμό παικτών με τη χρήση της προσέγγισης του μέσου πεδίου, σε πολλές περιπτώσεις, απλοποιεί σημαντικά την κατάσταση, καθώς κάνει δυνατή την απόδειξη αποτελεσμάτων ύπαρξης και σε κάποιες περιπτώσεις τον υπολογισμό των στρατηγικών ισορροπίας. Στη συνέχεια, μελετώνται οι στρατηγικές που έχουν εξαχθεί με την προσέγγιση του μέσου πεδίου, όταν εφαρμόζονται σε ένα παιχνίδι με πολλούς αλλά πεπερασμένους παίκτες. Συχνά, μπορεί να αποδειχθεί ότι οι στρατηγικές αυτές αποτελούν ε-ισορροπία, με $\varepsilon \to 0$ καθώς ο αριθμός των παικτών τείνει στο άπειρο, γεγονός που δικαιολογεί τη χρήση της προσέγγισης.

Στη συνέχεια αναφέρονται κάποιες εργασίες που σχετίζονται άμεσα με τη διατριβή αυτή, καθώς και κάποιες επεκτάσεις των παιγνίων μέσου πεδίου. Γραμμικά τετραγωνικά παίγνια με μεγάλο αριθμό παικτών μελετώνται στις εργασίες: [HCM07],[HCM12], [Bar12], [BSYY14] και [Pap14]. Παιχνίδια με risk-sensitive κριτήρια κόστους, καθώς και σθεναρά παιχνίδια αναλύονται στις εργασίες: [TZB$^+$11], [BTB$^+$12]. Μια αρκετά γενική προσέγγιση βασισμένη στη θεωρία τελεστών, παρουσιάζεται στην εργασία [KLY11]. Παίγνια που περιλαμβάνουν ένα σημαντικό παίκτη (major player), παρουσιάζονται στις [Hua10], [NH12], [NC13], [WZ12]. Στην εργασία [Tem11a] παρουσιάζονται διάφορες επεκτάσεις της θεωρίας των παιγνίων μέσου πεδίου όπως η σύνδεση των προβλημάτων συνεχούς και διακριτού χρόνου.

1.3 Δυναμικοί Κανόνες Εκμάθησης/Προσαρμογής

Στο εδάφιο 1.1.2, παρουσιάστηκαν κάποιες έννοιες λύσης βασισμένες στην ιδέα της ισορροπίας. Η ισορροπία προϋποθέτει ότι οι παίκτες έχουν κάποια «κοινή γνώση» (common knowledge). Για παράδειγμα θα πρέπει να συμφωνούν στο είδος της ισορροπίας που παίζουν (π.χ. αν είναι Stackelberg ποιος είναι αρχηγός) και θα πρέπει ο καθένας να γνωρίζει το κόστος του άλλου παίκτη ή στη στοχαστική περίπτωση, την κατανομή από την οποία έχει προέλθει. Επιπροσθέτως, σε δυναμικά παιχνίδια είναι πολλές φορές πολύ δύσκολο να υπολογιστεί η ισορροπία, ιδίως σε προβλήματα με ελλιπή πληροφορία.

Ένας εναλλακτικός τρόπος πρόβλεψης/περιγραφής της συμπεριφοράς των παικτών σε επαναλαμβανόμενα ή δυναμικά παιχνίδια είναι η διατύπωση κάποιων ντετερμινιστικών ή στοχαστικών δυναμικών κανόνων για τη συμπεριφορά των παικτών. Κίνητρο για τη διατύπωση τέτοιων κανόνων είναι συνήθως η προσπάθεια περιγραφής της εκμάθησης των παικτών, της προσαρμογής των στρατηγικών ή της διαδικασίας της εξέλιξης. Πολλά από αυτά τα μοντέλα αντανακλούν την «περιορισμένη ορθολογικότητα» (bounded rationality) των παικτών, δηλαδή την περιορισμένη ικανότητά τους να λύνουν δύσκολα προβλήματα [Sim72], [Sel01]. Οι διάφοροι δυναμικοί κανόνες διαφέρουν στην πολυπλοκότητά τους καθώς και στην πληροφορία που απαιτείται για την υλοποίησή τους. Η μελέτη των δυναμικών κανόνων συνήθως εστιάζει στην ασυμπτωτική τους συμπεριφορά. Στη συνέχεια περιγράφονται συνοπτικά μερικοί τέτοιοι κανόνες. Η παρούσα σύνοψη βασίζεται στις αναφορές [FL98], [FL09], [You04], [SLB08] ch. 7 και [LT11].

Ο παλιότερος και ίσως ο πιο απλός δυναμικός κανόνας είναι η χρήση της βέλτιστης απόκρισης από κάθε παίκτη (best response map (Cournot, 1838)). Ο κάθε παίκτης επιλέγει την επόμενη δράση του θεωρώντας ότι οι υπόλοιποι πρόκειται να διατηρήσουν στο μέλλον τις τρέχουσες δράσεις τους. Η διαδικασία αυτή επαναλαμβάνεται σε κάθε βήμα, ταυτόχρονα από όλους του παίκτες. Διάφορες τροποποιήσεις έχουν προταθεί, όπως η ασύγχρονη (πιθανόν στοχαστική) εφαρμογή των βημάτων από τους παίκτες, η χρήση μιας καλύτερης αντί της βέλτιστης απόκρισης, καθώς και η χρήση ορισμένων παλαιότερων τιμών των δράσεων στον υπολογισμό της βέλτιστης απόκρισης. Εάν υπάρχει σύγκλιση, τότε οι στρατηγικές στο όριο είναι σε ισορροπία Nash. Η σύγκλιση έχει αποδειχθεί για κάποιες περιπτώσεις πεπερασμένων παιγνίων (weakly acyclic [You04]), όταν οι απεικονίσεις της βέλτιστης απόκρισης είναι συστολικές, ενώ διάφορες παραλαγές της απεικόνισης βέλτιστης απόκρισης σε τετραγωνικά παιχνίδια έχει αποδειχθεί ότι συγκλίνουν ([Pap86]).

Ένας, επίσης κλασικός, δυναμικός κανόνας για επαναλαμβανόμενα πεπερασμένα παιχνίδια είναι το Fictitious Play. Ο κάθε παίκτης υποθέτει ότι οι άλλοι παίκτες ακολουθούν κάποια μεικτή στρατηγική σταθερή στο χρόνο. Σε κάθε βήμα, παρατηρεί τις εμπειρικές κατανομές των δράσεων των άλλων παικτών και επιλέγει τη βέλτιστη απόκριση στην κατανομή αυτή. Αν οι εμπειρικές κατανομές συγκλίνουν τότε στο όριο αποτελούν ισορροπία Nash. Παρόλα αυτά το fictitious play δε συγκλίνει πάντα, ενώ έχουν προσδιοριστεί κάποιες κατηγορίες παιχνιδιών που συγκλίνει (παιχνίδια μηδενικού αθροίσματος, potential games και πεπερασμένα παιχνίδια στα οποία ο ένας παίκτης έχει μόνο δύο στρατηγικές).

Μια άλλη κατηγορία δυναμικών κανόνων είναι οι στρατηγικές ενισχυτικής εκμάθησης (Reinforcement Learning). Σε αυτές τις στρατηγικές, οι παίκτες επιλέγουν κάποια από τις διάφορες δράσεις τους με κάποια πιθανότητα. Στη συνέχεια αυξάνουν ή μειώνουν την πιθανότητα να επιλέγουν την κάθε δράση με ανάλογα με το κόστος που είχαν όταν την επέλεξαν. Τέτοιοι κανόνες έχουν εφαρμοστεί σε επαναλαμβανόμενα παιχνίδια (π.χ. [Pap89b]) και σε στοχαστικά παιχνίδια (π.χ. [SPG03]). Αρκετά συναφείς είναι και οι κανόνες προσαρμογής σε δυναμικά παιχνίδια που βασίζονται στον προσαρμοστικό έλεγχο (π.χ. [Pap88], [YP94] [LZHMD14] ch. 6) και σε επαναλαμβανόμενα άπειρα παιχνίδια (π.χ. [FKB12]).

Μια αρκετά συνηθισμένη κατηγορία δυναμικών κανόνων στη θεωρία της εκμάθησης σε παιχνίδια βασίζεται στην έννοια του "regret". Για καθεμία δράση, ο κάθε παίκτης ορίζει το regret , δηλαδή το πόσο παραπάνω κόστος έχει σε σύγκριση με το να έπαιζε συνέχεια τη δράση αυτή. Υπάρχουν δυναμικοί κανόνες που εξασφαλίζουν το μέσο regret να τείνει στο μηδέν για μεγάλους χρόνους. για όλες τις διαθέσιμες δράσεις (no regret learning). Σε πεπερασμένα επαναλαμβανόμενα παιχνίδια, τέτοιοι κανόνες συγκλίνουν σε κάποιο είδος από συσχετισμένες ισορροπίες.

Σημαντική από θεωρητικής πλευράς είναι και η θεωρία του rational (Bayesian) learning [KL93]. Ο κάθε παίκτης σχηματίζει μια πεποίθηση (belief) για τις στρατηγικές των άλλων παικτών. Σε κάθε βήμα χρησιμοποιεί τον κανόνα του Bayes για να ανανεώσει τις την πεποίθησή του και επιλέγει τη βέλτιστη απόκριση στην πεποίθηση αυτή. Σε επαναλαμβανόμενα, πεπερασμένα παιχνίδια, κάτω από κάποιες τεχνικές συνθήκες (απόλυτη συνέχεια των κατανομών στις πιθανές ιστορίες που προκύπτουν αν οι πεποιθήσεις ήταν αληθείς ως προς τις πραγματικές κατανομές), μπορεί να αποδειχθεί ότι οι στρατηγικές συγκλίνουν στοχαστικά σε κάποια ισορροπία Nash του επαναλαμβανόμενου παιγνίου. Παρ' ότι η υπόθεση για απόλυτη συνέχεια είναι αρκετά περιοριστική και ότι το σύνολο των ισορροπιών Nash του επαναλαμβανόμενου παιγνίου είναι τεράστιο, η θεωρία του rational learning είναι σημαντική επειδή δείχνει σύγκλιση σε μια ισορροπία Nash σε προβλήματα με ελλιπή πληροφορία.

Σημαντικές από θεωρητικής πλευράς είναι και οι εργασίες [FY03], [FY$^+$06] οι οποίες εξετάζουν ορισμένους στοχαστικούς κανόνες εκμάθησης σε επαναλαμβανόμενα πεπερασμένα παιχνίδια. Το μοντέλο εκμάθησης περιλαμβάνει μια 'καθιερωμένη' δράση (status-quo) η οποία αναθεωρείται σε αραιά τυχαία χρονικά διαστήματα. Αν δεν θεωρηθεί επαρκώς καλή αυτή η δράση επιλέγεται μια άλλη κατά τύχη. Αποδεικνύεται η διαδικασία αυτή θα βρίσκεται με μεγάλη πιθανότητα πολύ κοντά σε μια ισορροπία του στατικού παιχνιδιού, για μεγάλους χρόνους.

Μια συναφής κατηγορία δυναμικών κανόνων εξετάζεται στην Εξελικτική Θεωρία Παιγνίων (Evolutionary Game Theory) [Smi82], [HS98]. Η Εξελικτική Θεωρία Παιγνίων εμπνέεται από τη θεωρία της Εξέλιξης και αρχικά διατυπώθηκες προκειμένου να περιγράψει τις συμπεριφορές σε πληθυσμούς διαφόρων ειδών σε καταστάσεις αλληλεπίδρασης, ενώ στη συνέχεια εφαρμόστηκε ευρύτερα και σε άλλες περιοχές. Το αριθμητικό κριτήριο ενός παιχνιδιού (κριτήριο κόστους) αντιστοιχίζεται με την εξελικτική καταλληλότητα (fitness) και συμπεριφορές (δράσεις ή στρατηγικές) οι οποίες ευνοούνται εξελικτικά επεκτείνονται. Ο πιο συνηθισμένος δυναμικός κανόνας είναι η «δυναμική της αντιγραφής» (Replicator Dynamics), ενώ μια σημαντική έννοια που διατυπώνεται είναι η εξελικτική ευστάθεια.

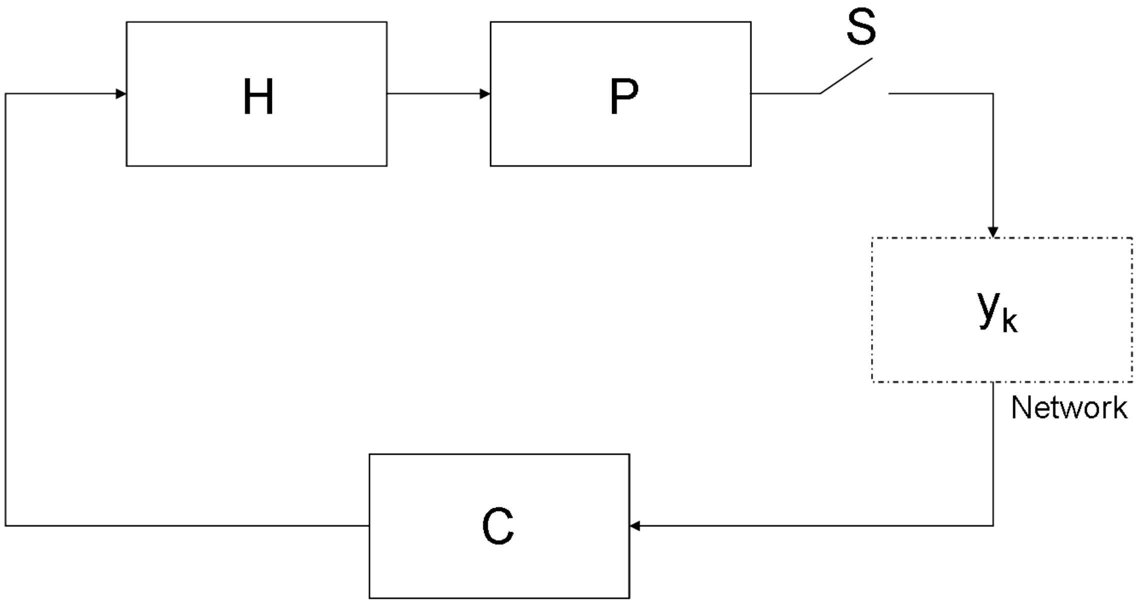
Στη συνέχεια παρουσιάζονται τα προβλήματα με τα οποία ασχολείται η παρούσα εργασία και συνοπτικά κάποια από τα αποτελέσματα.

1.4 Γραμμικά Συστήματα με Μαρκοβιανά Άλματα

Στο Κεφάλαιο 2 μελετώνται συστήματα στη μορφή:

$$\begin{aligned} &x_{k+1} = A(y_k)x_k + B(y_k)u_k + w_k, \\ &y_k : \quad \text{Αλυσίδα Markov}, \end{aligned} \tag{1.1}$$

θεωρώντας ότι η αλυσίδα Markov y_k έχει ένα γενικό χώρο κατάστασης D (πιθανόν υποσύνολο ενός Ευκλείδειου χώρου). Στη βιβλιογραφία έχει μελετηθεί η περίπτωση που ο χώρος κατάστασης D είναι πεπερασμένος ή αριθμήσιμα άπειρος.



Σχήμα 1.1: Σύστημα ελεγχόμενο μέσω δικτύου επικοινωνιών

Κίνητρο για τη μελέτη τέτοιων συστημάτων αποτέλεσε η ανάλυση γραμμικών τετραγωνικών παιχνιδιών με τυχαία είσοδο παικτών. Υπάρχουν όμως πολλές πιθανές εφαρμογές στις οποίες μπορεί να εμφανίζονται συστήματα στη μορφή (1.1). Για παράδειγμα ένα γραμμικό σύστημα που ελέγχεται μέσω ενός δικτύου επικοινωνιών, όπως της Εικόνας 1.1, όπου P (plant) είναι ένα γραμμικό σύστημα συνεχούς χρόνου, S ένας δειγματολήπτης, y_k είναι η καθυστέρηση που εισάγεται από το σύστημα επικοινωνιών, C ένας ελεγκτής και H ένας συγκρατητής μηδενικού βαθμού (zero order hold). Ένα άλλο παράδειγμα είναι ένα αβέβαιο γραμμικό σύστημα, όπου οι πίνακες ανήκουν σε ένα πολύεδρο. Η εισαγωγή μιας μαρκοβιανής δομής στην αβεβαιότητα οδηγεί σε MJLS με χώρο κατάστασης το πολύεδρο αυτό. Η ανάλυση πιθανόν να οδηγεί σε λιγότερο συντηρητικά αποτελέσματα ευστάθειας.

Τα βασικά αποτελέσματα αφορούν την μέση τετραγωνική (εκθετική) ευστάθεια τέτοιων συστημάτων, δηλαδή συνθήκες έτσι ώστε $E[x_k^T x_k] \to 0$ εκθετικά, καθώς $k \to \infty$ και το γραμμικό τετραγωνικό έλεγχο. Πιο συγκεκριμένα, αφού χαρακτηριστεί η μέση τετραγωνική εκθετική ευστάθεια μέσω της φασματικής ακτίνας ενός τελεστή, αποδεικνύονται ικανές και αναγκαίες συνθήκες που είναι πιο εύκολο να ελεγχθούν. Συγκεκριμένα:

<u>Πρόταση 1</u>: Για ένα σύστημα στη μορφή

$$x_{k+1} = A(y_k)x_k,$$

τα ακόλουθα είναι ισοδύναμα:

(i) Το σύστημα είναι εκθετικά ευσταθές υπό τη μέση τετραγωνική έννοια.

(ii) Υπάρχει σταθερά $a \in (0,1)$ και ακέραιος k_0 έτσι ώστε $E[x_{k_0}^T x_{k_0}] < a x_0^T x_0$, για όλα τα x_0.

(iii) Για ένα θετικά ορισμένο πίνακα από συναρτήσεις $Q(y)$ υπάρχει ένας θετικά ορισμένος πίνακας από συναρτήσεις M έτσι ώστε:

$$A^T(y)E[M(y_1)|y_0 = y]A(y) - M(y) = -Q(y),$$

για όλα τα $y \in D$.

Η συνθήκη (iii) περιλαμβάνει την επίλυση μιας γραμμικής ολοκληρωτικής εξίσωσης τύπου Fredholm και η συνθήκη (ii) μπορεί να εξεταστεί κάνοντας υπολογισμούς αναδρομικά στο k_0 και συνεπώς, είναι δυνατόν να ελεγχθούν αριθμητικά.

Στη συνέχεια μελετώνται προβλήματα γραμμικού τετραγωνικού ελέγχου. Στον άπειρο ορίζοντα το πρόβλημα είναι η ελαχιστοποίηση του κριτηρίου:

$$J = E\left[\sum_{k=0}^{\infty} a^k \left(x_k^T Q(y_k) x_k + u_k^T R(y_k) u_k\right)\right], \tag{1.2}$$

όπου $a \in (0,1)$ μια παράμετρος απόσβεσης.

Ο βέλτιστος ελεγκτής (εάν υπάρχει) μπορεί να εκφραστεί μέσω της λύσης της παρακάτω ολοκληρωτικής εξίσωσης Riccati.

$$\begin{aligned} K(y) = Q(y) + A^T(y)[a\Lambda(y) - a\Lambda(y)B(y)\cdot \\ \cdot (R(y)/a + B^T(y)\Lambda(y)B(y))^{-1}B(y)^T\Lambda(y)]A(y) \end{aligned} \tag{1.3}$$

όπου:

$$\Lambda(y) = E[K(y_{k+1})|y_k = y] \tag{1.4}$$

Μπορεί να αποδειχθεί ότι:

Πρόταση 2: Αν υπάρχει έλεγχος ανάδρασης που κάνει το J πεπερασμένο τότε το βέλτιστο κόστος μπορεί να γραφτεί ως:

$$J^*(x, y) = x^T K(y) x + c(y),$$

και ο βέλτιστος ελεγκτής δίνεται από:

$$u_k = L(y_k)x_k = -(B^T(y_k)\Lambda(y_k)B(y_k) + R(y_k)/a)^{-1}B^T(y_k)\Lambda(y_k)A(y_k)x_k,$$

όπου K, Λ αποτελούν λύση των (1.3),(1.4).

1.5 Παίγνια με Τυχαία Είσοδο Παικτών

Στο Κεφάλαιο 3 μελετώνται παιχνίδια με τυχαία είσοδο παικτών οι οποίοι συμμετέχουν στο παιχνίδι για διαφορετικά διαστήματα. Συγκεκριμένα, θεωρούμε ότι υπάρχει ένας παίκτης που μένει στο παιχνίδι για άπειρο χρονικό ορίζοντα τον οποίο καλούμε major player και πολλοί παίκτες με πεπερασμένους χρονικούς ορίζοντες που καλούνται minor players.

Υπάρχουν διάφορα πρακτικά παραδείγματα στα οποία υπάρχει ένας σημαντικός παίκτης ο οποίος έχει πολύ μεγάλο χρονικό ορίζοντα και αλληλεπιδρά με ένα αριθμό από παίκτες με μικρότερη επιρροή, ενώ ο χρόνος της αλληλεπίδρασης με τον καθένα από αυτούς τους παίκτες είναι σχετικά μικρός. Για παράδειγμα, σε μια απελευθερωμένη αγορά ενέργειας ο δημόσιος φορέας παραγωγής μπορεί να θεωρηθεί ο major player με άπειρο χρονικό ορίζοντα, ενώ οι επιχειρήσεις παραγωγής ανανεώσιμης ενέργειας, που παίρνουν άδεια να εισέλθουν στο σύστημα για ένα μικρό προκαθορισμένο χρονικό διάστημα, μπορούν να θεωρηθούν minor players. Ένα άλλο παράδειγμα είναι μια τράπεζα η οποία έχει πολύ μεγάλο χρονικό ορίζοντα και αλληλεπιδρά με πολλούς δανειολήπτες, οι οποίοι έχουν μικρότερο προκαθορισμένο χρονικό ορίζοντα. Ένα τρίτο παράδειγμα είναι ένα πανεπιστήμιο το οποίο έχει μεγάλο χρονικό ορίζοντα και αλληλεπιδρά με τους φοιτητές του κάθε εξαμήνου που έχουν σημαντικά μικρότερο χρονικό ορίζοντα.

Ένα παράδειγμα συμμετοχής των παικτών φαίνεται στην εικόνα 1.2. Η είσοδος των παικτών μπορεί να περιγραφεί από μια αλυσίδα Markov:

$$y_k = (N_k^0, \ldots, N_k^{T-1})/s_c, \tag{1.5}$$

όπου N_k^l είναι ο αριθμός των παικτών που εισήλθαν στο παιχνίδι τη χρονική στιγμή $k-l$, T ο χρονικός ορίζοντας των minor players και s_c ο μέγιστος πιθανός αριθμός των παικτών. Η δυναμική είναι γραμμική και τα κόστη τετραγωνικά.

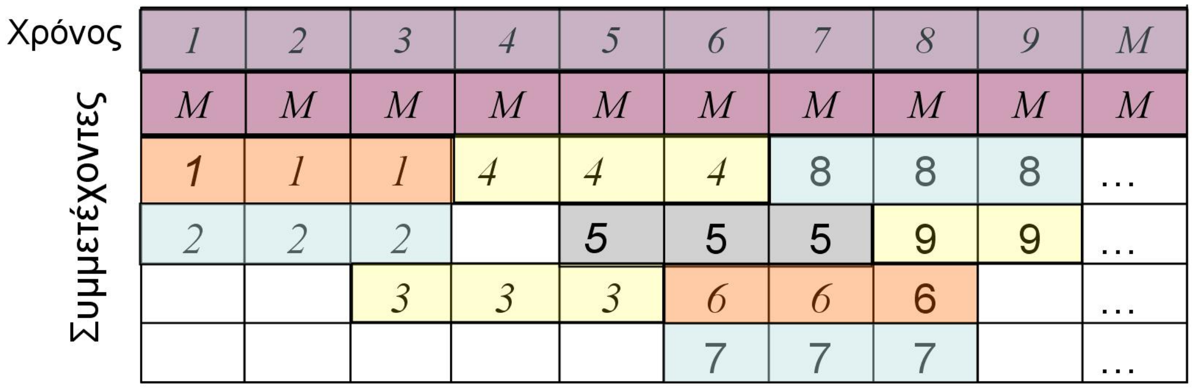
Το πρόβλημα της τυχαίας εισόδου παικτών μετασχηματίζεται σε συζευγμένα γραμμικά τετραγωνικά προβλήματα ελέγχου για MJLS. Οι ισορροπία Nash χαρακτηρίζεται συνεπώς από συζευγμένες εξισώσεις Riccati για MJLS.

Στη συνέχεια εξετάζεται η περίπτωση που υπάρχει ένας πολύ μεγάλος αριθμός από minor players. Η αλυσίδα Markov (1.5) η οποία έχει διακριτό αλλά πολύ μεγάλο χώρο κατάστασης, προσεγγίζεται από μια άλλη με συνεχή χώρο κατάστασης:

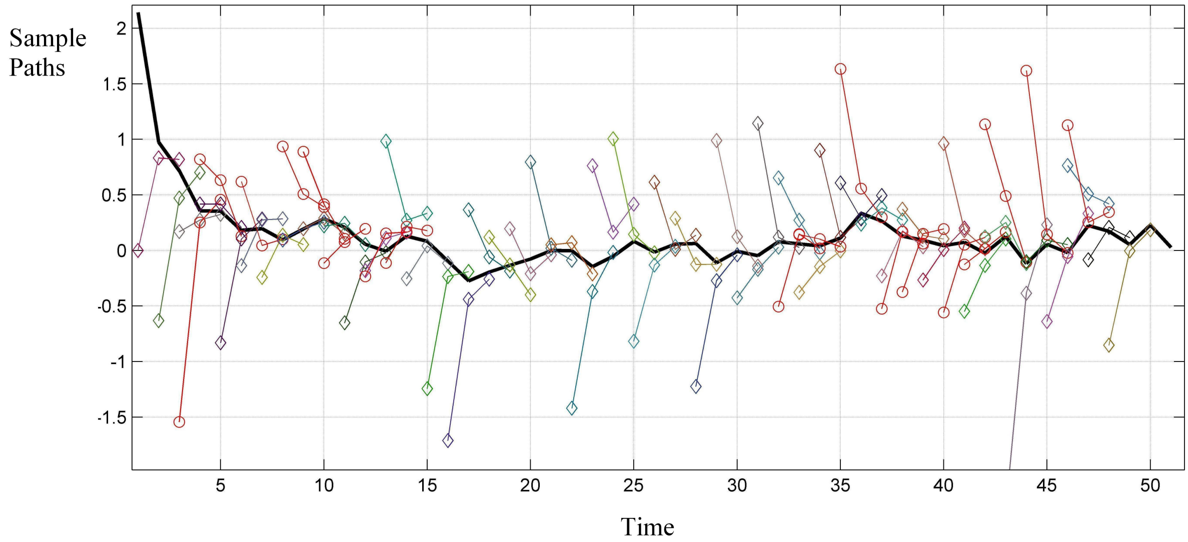
$$D = \left\{ (y^0, \ldots, y^{T-1}) \in \mathbf{R}^T : \sum_{i=0}^{T-1} y^i \leq 1, y^i \geq 0, \right\}.$$

Ακολούθως, αναλύεται το παίγνιο με την προσέγγιση του μέσου πεδίου και διατυπώνονται τα συζευγμένα προβλήματα βέλτιστου ελέγχου, στο όριο καθώς ο αριθμός των minor players τείνει στο άπειρο. Η λύση των προβλημάτων αυτών εφαρμόζεται σε παιχνίδια με μεγάλο (πεπερασμένο) αριθμό minor players και αποδεικνύεται ότι αποτελεί ε ισορροπία Nash, με $\varepsilon \to 0$ καθώς ο αριθμός των παικτών τείνει στο άπειρο. Η απόδειξη έχει ενδεχομένως και ανεξάρτητο ενδιαφέρον και παουσιάζεται στην Παράγραφο 3.8. Με την προσέγγιση του μέσου πεδίου είναι δυνατόν να αποδειχθούν και κάποια αποτελέσματα ύπαρξης.

Τέλος προτείνεται ένας αλγόριθμος για τον υπολογισμό της ισορροπίας Nash. Ένα παράδειγμα από sample paths παρουσιάζεται στην Εικόνα 1.3.



Σχήμα 1.2: Ένα παράδειγμα τυχαίας εισόδου παικτών

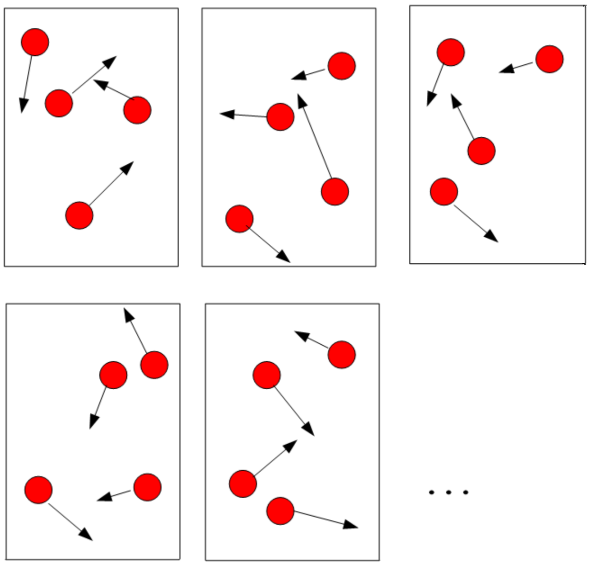


Σχήμα 1.3: Κάποιες τροχιές

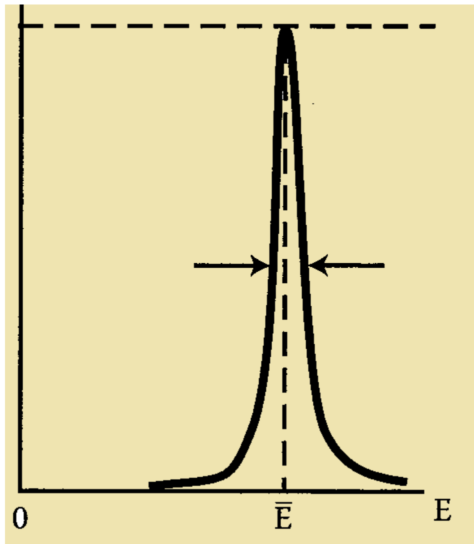
1.6 Παίγνια σε Μεγάλα Δίκτυα

Στο Κεφάλαιο 4 μελετώνται παίγνια στα οποία οι παίκτες αλληλεπιδρούν πάνω σε ένα μεγάλο δίκτυο. Υπάρχουν πολλά κίνητρα για τη μελέτη τέτοιων παιγνίων. Ένα χαρακτηριστικό παράδειγμα είναι η αλληλεπίδραση των διαφόρων παραγωγών, καταναλωτών και ευφυών μικρο-δικτύων πάνω στο δίκτυο μεταφοράς/διανομής ηλεκτρικής ενέργειας. Υπάρχουν, επίσης, πολλά παραδείγματα καταστάσεων παιγνίου σε κοινωνικά δίκτυα, όπως η αναζήτηση εργασίας, η επιλογή προϊόντων (π.χ. τηλεπικοινωνιακού παρόχου), η αποδοχή ή όχι ιδεών, η συμμετοχή σε διάφορες ομάδες, η εμπλοκή σε εγκληματική συμπεριφορά κ.λπ. στις οποίες η απόφαση του κάθε παίκτη εξαρτάται τόσο από τις προτιμήσεις του ιδίου όσο και από τις αποφάσεις των «φίλων» του. Υπάρχουν επίσης διάφορες οικονομικές καταστάσεις στις οποίες παρατηρείται τοπικός αλλά και ευρύτερος ανταγωνισμός ο οποίος μπορεί να παρασταθεί από αλληλεπιδράσεις πάνω σε ένα δίκτυο.

Υποθέτουμε ότι οι παίκτες δεν έχουν πλήρη γνώση του δικτύου των αλληλεπιδράσεων. Διαθέτουν όμως μια στατιστική περιγραφή του δικτύου π.χ. τυχαίο γράφημα τύπου Erdos–



(α′) Μια στατιστική συλλογή



(β′) Η συνολική ενέργεια στην κανονική συλλογή.

Σχήμα 1.4: Αναλογία με τη στατιστική Φυσική

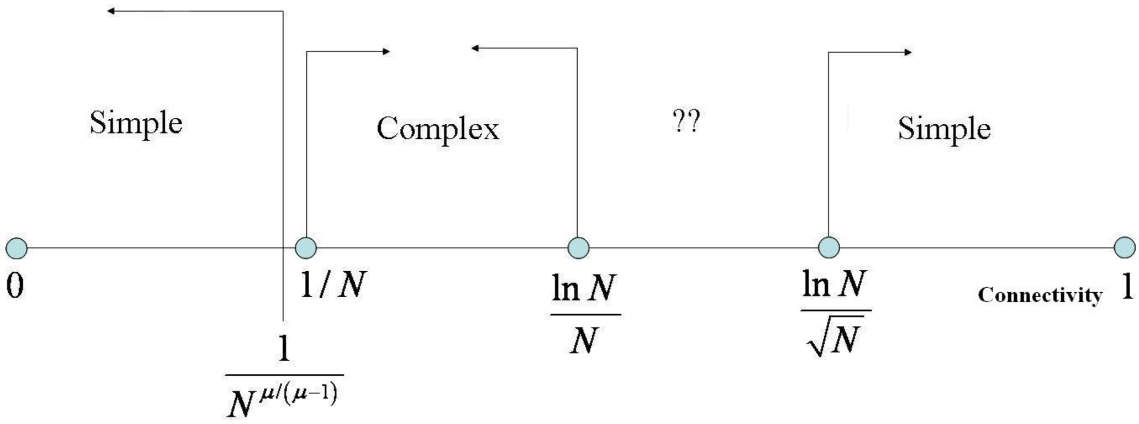
Renyi, small world network, πλέγμα, κ.λπ. Επιπροσθέτως, υποθέτουμε ότι διαθέτουν και κάποια τοπική πληροφορία, δηλαδή ότι γνωρίζουν τις συνδέσεις του δικτύου, καθώς και τις προτιμήσεις των παικτών σε μια γειτονιά τους. Η βασική ερώτηση που αναλύεται είναι υπό ποιες προϋποθέσεις είναι δυνατή η ύπαρξη ή/και ο υπολογισμός μιας προσεγγιστικής ισορροπίας σε τέτοια παιχνίδια.

Αρχικά ορίζεται μια έννοια προσεγγιστικής ισορροπίας εμπνευσμένη από έννοιες της στατιστικής φυσικής. Στη στατιστική φυσική, για την περιγραφή ενός μακροσκοπικού συστήματος, είναι αδύνατον να μετρηθούν οι αρχικές συνθήκες (π.χ. θέσεις και ταχύτητες) από ένα πολύ μεγάλο αριθμό (π.χ. $\sim 10^{23}$) σωματιδίων και να επιλυθούν οι δυναμικές εξισώσεις για την εξέλιξη τους. Το πρόβλημα αυτό αντιμετωπίζεται θωρώντας μια συλλογή από νοητικά αντίγραφα του συστήματος που έχουν διαφορετικές αρχικές συνθήκες, η οποία καλείται στατιστική συλλογή (statistical ensemble, π.χ. Εικόνα 1.4α′). Διάφορες μακροσκοπικές ιδιότητες έχουν τιμή πολύ κοντά σε μια ντετερμινιστική σταθερά για όλα τα συστήματα στη στατιστική συλλογή, εκτός πιθανόν από ένα σύνολο συστημάτων με πολύ μικρή πιθανότητα (π.χ. Εικόνα 1.4β′). Μια αναλογία με την κατάσταση που απεικονίζεται στην Εικόνα 1.4β′ χρησιμοποιείται για να οριστεί μια έννοια προσεγγιστικής ισορροπίας η οποία θα ονομάζεται κατά πιθανοτικά προσεγγιστική ισορροπία Nash (Probabilistic Approximate Nash (PAN) equilibrium).

Στη συνέχεια ασχολούμαστε με το παρακάτω ερώτημα:

Ερώτηση: Ποια είναι η ελάχιστη ποσότητα πληροφορίας που πρέπει να έχουν οι παίκτες έτσι ώστε να υπάρχει ένα σύνολο στρατηγικών σε προσεγγιστική ισορροπία.

Με βάση την απάντηση στην ερώτηση αυτή ορίζεται μια έννοια πληροφοριακής πολυπλοκότητας (necessary information complexity (NIC)), ως η ελάχιστη ποσότητα πληροφορίας που



Σχήμα 1.5: Η πολυπλοκότητα για διάφορα διαστήματα της πιθανότητας σύνδεσης.

είναι απαραίτητη για την ύπαρξη της προσεγγιστικής ισορροπίας.

Ακολούθως, αναλύονται κλάσεις παιχνιδιών στις οποίες η πολυπλοκότητα είναι χαμηλή. Είναι, δηλαδή, εφικτή εύρεση προσεγγιστικής ισορροπίας υποθέτοντας μόνο μια μικρή ποσότητα από τοπική και στατιστική πληροφορία. Επεκτείνεται, συνεπώς, η θεωρία των παιγνίων μέσου πεδίου (MFG) και σε περιπτώσεις που οι αλληλεπιδράσεις δεν είναι συμμετρικές.

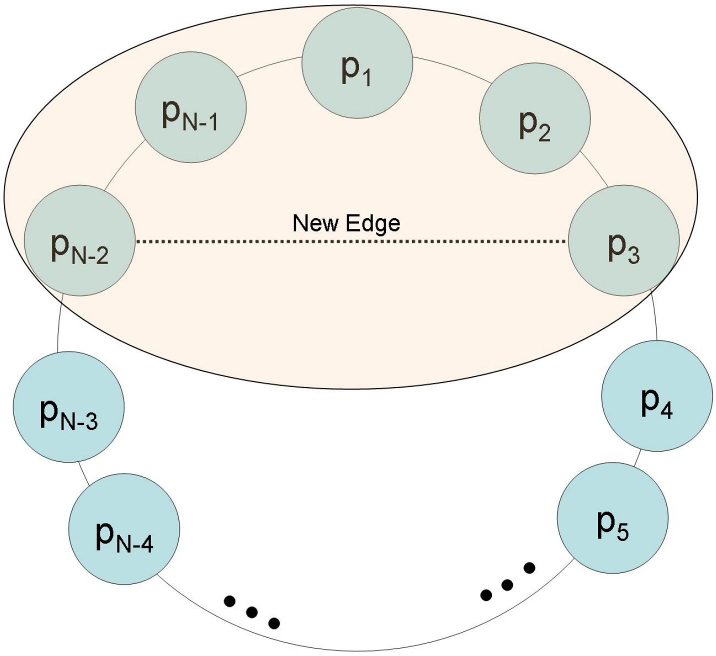
Παραδείγματα τέτοιων κλάσεων είναι τα στατικά και κάποια γραμμικά τετραγωνικά δυναμικά παιχνίδια σε τυχαία γραφήματα τύπου Erdos–Renyi. Τα παίγνια αυτά είναι απλά όταν η πιθανότητα σύνδεσης είναι του τυχαίου γραφήματος είναι μεγάλη. Για τα στατικά παιχνίδια, η πολυπλοκότητα παρουσιάζεται στην Εικόνα 1.5. Παρόμοια αποτελέσματα ισχύουν και στην περίπτωση αλληλεπίδρασης σε ένα Small World Network.

Μια δεύτερη ειδική περίπτωση που αναλύεται είναι τα τετραγωνικά παίγνια με παίκτες που αλληλεπιδρούν σε ένα πολυδιάστατο πλέγμα. Σε αυτή την περίπτωση μπορούν να κατασκευαστούν προσεγγιστικές ισορροπίες, χρησιμοποιώντας επανειλημμένα την απεικόνιση βέλτιστης απόκρισης. Έτσι, η πληροφοριακή πολυπλοκότητα είναι πολυωνυμική και η τάξη του πολυωνύμου ταυτίζεται με τη διάσταση του πλέγματος.

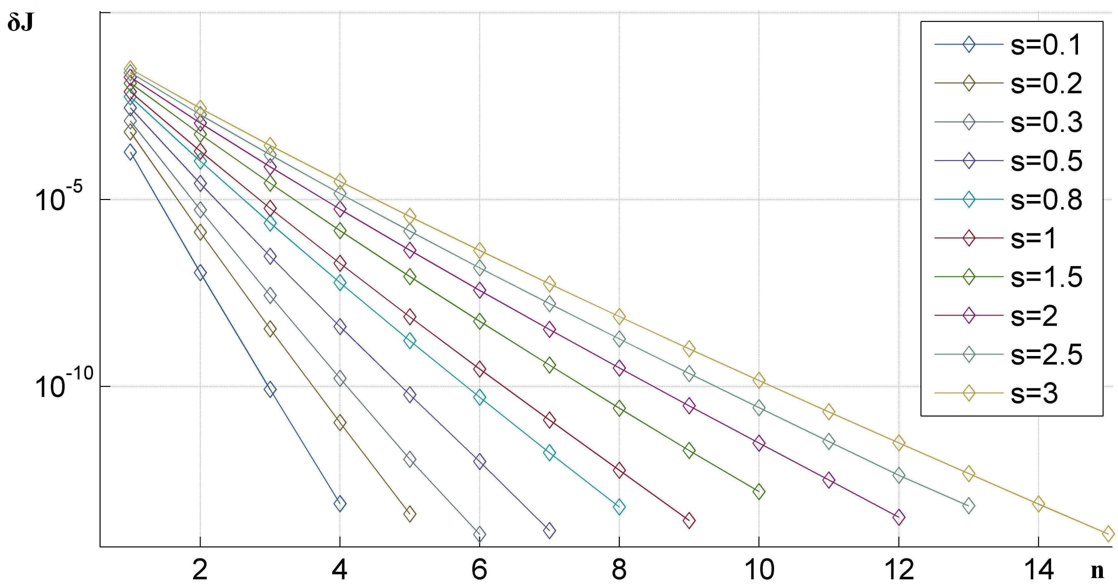
Στη συνέχεια αναλύεται ένα παράδειγμα στο οποίο οι παίκτες αλληλεπιδρούν πάνω σε ένα δακτύλιο και έχουν μη τετραγωνικά κόστη. Η βέλτιστη απόκριση στο παράδειγμα αυτό είναι χαοτική. Όμως, μπορούν να κατασκευαστούν στρατηγικές που αποτελούν προσεγγιστική ισορροπία αν οι παίκτες έχουν μια (έστω και μικρή) συνεργασία.

Τέλος αναλύεται αριθμητικά ένα παράδειγμα γραμμικών τετραγωνικών παιγνίων με παίκτες που αλληλεπιδρούν πάνω σε ένα δακτύλιο. Στρατηγικές που αποτελούν προσεγγιστική ισορροπία κατασκευάζονται με μια τεχνική αναγωγής. Αρχικά ο κάθε παίκτης θεωρεί ένα μικρότερο παιχνίδι, στη συνέχεια βρίσκει μια στρατηγική ισορροπίας στο παιχνίδι αυτό. Τέλος, εφαρμόζει τη στρατηγική αυτή στο αρχικό παιχνίδι. Η αναγωγή φαίνεται στην Εικόνα 1.6.

Η διαφορά μεταξύ της απόδοσης της προσεγγιστικής στρατηγικής και της βέλτιστης (υποθέτοντας ότι οι άλλοι παίκτες χρησιμοποιούν την προσεγγιστική), ως συνάρτηση της



Σχήμα 1.6: Ο παίκτης p_1 θεωρεί ότι συμμετέχει σε ένα παιχνίδι μόνο με τους παίκτες $p_2, p_3, p_{N-1}, p_{N-2}$ θεωρώντας μια νέα σύνδεση μεταξύ p_3 ανδ p_{N-2}.



Σχήμα 1.7: Η διαφορά $J^i(u^{i,n}, u^{-i,n}) - \min_u J^i(u, u^{-i,n})$ ως συνάρτηση του n

διάστασης του ανηγμένου παιχνιδιού, για διάφορες τιμές των παραμέτρων, φαίνεται στην Εικόνα 1.7.

Στα παραδείγματα που αναλύονται, οι βασικοί λόγοι που οδηγούν σε χαμηλή πολυπλοκότητα είναι κάποιοι νόμοι μεγάλων αριθμών, η συστολικότητα της απεικόνισης της βέλτιστης από- κρισης (**best response map**), η συνεργασία μεταξύ των παικτών ή μικρά κέρδη σε απομακρυσμέ- νους παίκτες.

1.7 Δυναμικοί Κανόνες και Εξαπάτηση

Το Κεφάλαιο 5 εξετάζει τη χρήση δυναμικών κανόνων σε καταστάσεις παιγνίου με ελλιπή δομική πληροφορία. Οι δυναμικοί κανόνες αντανακλούν την περιορισμένη ορθολογικότητα των παικτών και μπορεί να εκφράζουν προσαρμογή, εξέλιξη ή εκμάθηση. Ένα κίνητρο για τη μελέτη δυναμικών κανόνων είναι η μοντελοποίηση των επαναλαμβανόμενων αλληλεπιδράσεων ανάμεσα σε παραγωγούς ηλεκτρικής ενέργειας. Δομικά στοιχεία για τον κάθε παίκτη, όπως το κόστος παραγωγής, δεν είναι πλήρως γνωστά στους υπόλοιπους παίκτες. Συνεπώς η χρήση δυναμικών κανόνων είναι εύλογη.

Οι διάφοροι δυναμικοί κανόνες στην τυπική περίπτωση δεν βρίσκονται σε ισορροπία Nash. Συνεπώς αν κάποιος παίκτης γνωρίζει τον δυναμικό κανόνα των αντιπάλων του, έχει τη δυνατότητα να τους εξαπατήσει και να εκμεταλλευτεί τη γνώση αυτή. Οι βασικές ερωτήσεις που εξετάζει το Κεφάλαιο 5 είναι:

- Σε ποιες περιπτώσεις ένα σύνολο από δυναμικούς κανόνες είναι ικανοποιητική πρόβλεψη για μια κατάσταση παιγνίου.
- Ποια είναι η επίδραση της δυνατότητας εξαπάτησης στο αποτέλεσμα του παιγνίου.

Αναφορικά με την πρώτη ερώτηση διατυπώνονται δύο κριτήρια για την αξιολόγηση των δυναμικών κανόνων. Το κριτήριο της ευκαιρίας, εξετάζει κατά πόσο θα μειώνονταν το κόστος ενός παίκτη αν χρησιμοποιούσε την βέλτιστη απόκριση στη στρατηγική των αντιπάλων του. Το δεύτερο κριτήριο εξαρτάται από την αύξηση του κόστους ενός υπό εξέταση παίκτη όταν αντί για το δυναμικό κανόνα, ο αντίπαλος χρησιμοποιεί τη βέλτιστη απόκριση.

Στη συνέχεια εξετάζεται μια υποκατηγορία των στρατηγικών εξαπάτησης, την οποία ονομάζουμε στρατηγικές υποκρισίας. Σε αυτή την περίπτωση κάποιος παίκτης ή κάποιοι παίκτες χρησιμοποιούν τους δυναμικούς κανόνες τους σαν να είχαν κάποιο διαφορετικό τύπο (προτιμήσεις). Στην περίπτωση που υπάρχει αρκετή αβεβαιότητα και μόνο ένας παίκτης υποκρίνεται, προκύπτει ότι μπορεί να έχει το ίδιο κόστος με την περίπτωση που ήταν αρχηγός Stackelberg. Στη συνέχεια θεωρούμε την περίπτωση στην οποία όλοι οι παίκτες υποκρίνονται και κατασκευάζουμε ένα βοηθητικό παιχνίδι το οποίο καλούμε παιχνίδι των υποκριτών. Με τη βοήθεια αυτού του παιγνίου, διατυπώνουμε μια πρόβλεψη για το αρχικό παιχνίδι. Τέλος εξετάζεται ένα παράδειγμα που υπάρχουν πολλοί ισοδύναμοι παίκτες. Στην περίπτωση αυτή, το κίνητρο για υποκρισία είναι πολύ μικρό, και η βέλτιστη υποκρισία για κάθε παίκτη είναι πολύ κοντά στην πραγματικότητα.

Στη συνέχεια εξετάζονται κάποια μοντέλα για αγορές ηλεκτρικής ενέργειας. Αρχικά αναλύεται ένα μοντέλο ολιγοπωλίου Cournot, δηλαδή ενός ολιγοπωλίου στο οποίο οι παραγωγοί ενέργειας αποφασίζουν για την ενέργεια που θα προσφέρουν, ενώ η τιμή εξαρτάται από τη ζήτηση και τη συνολική προσφορά. Στην περίπτωση αυτή οι παραγωγοί έχουν συμφέρον να συμπεριφερθούν σαν να είχαν μικρότερο κόστος από το πραγματικό τους. Έτσι ο ανταγωνισμός μεταξύ τους ενισχύεται. Ένα παράδειγμα, στο οποίο υποθέτουμε γραμμική

Πίνακας 1.1: Συμμετρικό ολιγοπώλιο με 2 παίκτες

Προσποιείται	q_1	q_2	L_1	L_2	p
Κανείς	0.4	0.4	-0.16	-0.16	1.2
Παίκτης 1	0.6	0.3	-0.18	-0.09	1.1
Παίκτης 2	0.3	0.6	-0.09	-0.18	1.1
Και οι δύο	0.48	0.48	-0.1152	-0.1152	1.04

συνάρτηση ζήτησης, δίνεται στον πίνακα 1.1. Οι ποσότητες που παράγονται συμβολίζονται με q_1 και q_2, η τιμή με p και με L_1, L_2 συμβολίζονται τα κόστη των παικτών.

Σε ένα άλλο μοντέλο αγοράς, οι παίκτες υποβάλουν προφορές σαν γραμμικές συναρτήσεις ποσότητας-τιμής (Supply Function model). Στην περίπτωση αυτή η εξαπάτηση οδηγεί την τιμή υψηλότερα και το κόστος των παικτών χαμηλότερα. Υπό την έννοια αυτή η εξαπάτηση στη συγκεκριμένη περίπτωση είναι μια πράξη που ενισχύει τη συνεργασία ανάμεσα στους παραγωγούς.

Τέλος αναλύεται ένας μηχανισμός από τη βιβλιογραφία ([RT14]). Στην περίπτωση αυτή μια κατηγορία κανόνων εξαπάτησης κάνει τις τιμές να αυξάνονται απεριόριστα, κάνοντας το συνολικό σύστημα να μην δουλεύει.

Chapter 2

Stability and LQ Control of MJLS

This chapter studies the mean square stability and the LQ control of discrete time Markov Jump Linear Systems where the Markov chain has a general state space. The mean square stability is characterized by the spectral radius of an operator describing the evolution of the second moment of the state vector. Two equivalent tests for the mean square stability are obtained based on the existence of a positive definite solution to a Lyapunov equation and a uniformity result respectively. An algorithm for testing the mean square stability is also developed based on the uniformity result. The finite and infinite horizon LQ problems are considered and their solutions are characterized by appropriate Riccati integral equations. An application to Networked Control Systems (NCS) is finally presented and a simple example is studied via simulation.

2.1 Introduction

Markov Jump Linear Systems (MJLS) are linear dynamic models where the matrices describing the evolution of the state vector, depend on the state of a Markov chain. The existing results on the stability and LQ control of MJLS are dealing with Markov chains having a discrete state space, i.e. finite or countably infinite. However, in several applications, the natural choice for the state space of the Markov chain is not discrete. Some examples of Markov chains with continuous state space are given in chapters 1 and 2 of [MT93]. In this chapter, we extend the stability analysis and the LQ control of discrete time MJLS to a more general state space Markov chain case, including the continuous state space case.

An example of such applications is the study of systems with dependent random communication delays, such as Networked Control Systems [GC10],[Nil98]. The amount of time delay stands for the state variable of the Markov chain. A natural choice for the Markov chain state space is an uncountable subset of the real numbers, such as a closed interval. Another example could be a dynamic linear economic model, having coefficients depending on the price of some asset traded in a stock market. The price of a stock market is usually modeled as a geometric Brownian motion [BS73]. The state space of this Markov chain, is the positive real numbers. Examples of MJLS with continuous state space also arise in the study of gain scheduling control of nonlinear systems with a Markovian desired trajectory or a Markovian measurable disturbance. More generally, several systems have been modeled as linear uncertain systems, where the matrices describing the evolution of the state vector, belong to a compact polyhedron [BGFB94]. Assuming that there exists a Markovian model for the uncertainty, design problems involving MJLS with continuous state space Markov chain will be obtained and will lead to less conservative stability conditions. Another example, which also motivates the current work, comes from the optimal control problems, arising in the mean field approximation of LQ games involving a large number of randomly entering players, where the players are considered to belong to a continuum (see chapter 3).

The first work studying the mean square stability of MJLS with finite state space Markov chain was [Bha61]. Other contributions include [BP93] where the stochastic additive case was studied and [JC90],[CF93],[FL02] where necessary and sufficient conditions involving Lyapunov equations were derived. The mean square stability for MJLS with a countably infinite Markov chain was studied in [CF95], using an operator theoretic point of view. Some problems related to the mean square stability of MJLS with general state space were studied in [Mou98], [HKM02] under some ergodicity assumptions. The relation among several notions of stochastic stability for MJLS was studied in [FL02].

The first work studying LQ control problems related to MJLS was in a continuous time setting [Swo69]. A lot of work has been done on the Linear Quadratic control of the discrete time MJLS as well. The finite horizon LQ control problem for the finite state space Markov chain case was solved in [BS75] and its infinite horizon counterpart in [CWC86]. The existence of a solution was first studied using controllability notions in [JC88] and testable conditions were derived in [JC90]. A related work with infinite horizon ergodic criterion and safety constraints is [HAK10]. Filtering problems for MJLS are studied in [SBA99] and a review of several results is given in the books [CFM05] and [Bou05]. The LQ control problem, for a system involving a Markov chain with countably infinite state space was studied in [CF95].

The contribution of the work in this chapter is twofold. The first part, is the study of the mean square stability of MJLS when the Markov chain has a general state space. The mean square exponential stability notion is characterized by the spectral radius of a certain operator. Then, testable equivalent conditions are derived based on the operator theoretic result. The second part of the contribution of this work is the extension of the solution of finite and infinite horizon LQ problems to MJLS with general state space. The basic difference between the current work and the literature is that the techniques applied for the stability analysis of MJLS with discrete state space could not be directly extended to the continuous or general state space. In a comparison to older results, a more general class of models could be analyzed. Examples of models of Markov chains with general state spaces could be found in [MT93]

The following notation is used. The probability is denoted by $Pr(\cdot)$ and the expectation by $E[\cdot]$. The value of the Markov chain is denoted by y_k and its state space is the metric space

D. Denote by $\mathcal{B}(D)$ the σ-algebra of Borelian subsets of D. The evolution y_k is described by the notion of stochastic kernel, i.e. a function $\bar{K}(\cdot,\cdot) : D \times \mathcal{B}(D) \to [0,1]$ such that $Pr(y_{k+1} \in B | y_k = y) = \bar{K}(y, B)$. A matrix function $Q : D \to \mathbf{R}^{n\times n}$ will be called strictly positive definite, if there exists a positive constant c such that $Q(y) > cI$, for any $y \in D$. Finally, the spectral radius of an operator T is denoted by $r(T)$.

2.2 Problem Description

The system under consideration is the following:

$$x_{k+1} = A(y_k)x_k + B(y_k)u_k + w_k \tag{2.1}$$

where $x_k \in \mathbf{R}^n$ is the state vector, $u_k \in \mathbf{R}^m$ is the control input, $A(\cdot)$ and $B(\cdot)$ are Borel measurable, bounded matrix functions of appropriate dimensions and w_k are zero mean i.i.d. random variables with finite second moments.

Two types of problems are considered. The first type is the stability problem and it is stated as follows:

Stability problem: "Under which conditions the free system:

$$x_{k+1} = A(y_k)x_k \tag{2.2}$$

is stable". The notions of stochastic stability that we study in the current work are given in the following definition.

Definition 1. *The system given by* (2.2) *is:*

(i) Pointwise mean square stable, if $E[x_k^T x_k] \to 0$, *for any* $x_0 \in \mathbb{R}, y_0 \in D$.

(ii) Mean square stable, if $E[x_k^T x_k] \to 0$ *for any* x_0, y_0 *random variables, such that* $E[x_0^T x_0] < \infty$.

(iii) Mean square exponentially stable, if for any x_0, y_0 *random variables, such that* $E[x_0^T x_0] < \infty$, *there exist constants* $r \in (0,1)$ *and* $M > 0$ *such that it holds* $E[x_k^T x_k] < Mr^k$, *for any positive integer* k.

(iv) Stochastically mean square stable, if for any x_0, y_0 *random variables, such that* $E[x_0^T x_0] < \infty$, *it holds* $\sum_{k=0}^{\infty} x_k^T x_k < \infty$.

(v) Almost surely stable, if $Pr(x_k \to 0) = 1$, *for any* x_0, y_0 *random variables.*

The second type of problems considered is the LQ control problems. These problems are stated in a finite or an infinite horizon setting as follows:

Finite Horizon LQ Control Problem: "Find the control law $u_k = \gamma(x_k, y_k, k)$, that minimizes the LQ criterion:

$$E\left[x_{N+1}^T Q_{N+1}(y_{N+1})x_{N+1} + \right.$$
$$\left. + \sum_{k=0}^{N} \left(x_k^T Q_k(y_k)x_k + u_k^T R_k(y_k)u_k\right)\right] \text{''} \tag{2.3}$$

Infinite Horizon LQ Control Problem: "Find the control law $u_k = \gamma(x_k, y_k)$ (if any), that minimizes the LQ criterion:

$$E\left[\sum_{k=0}^{\infty} a^k \left(x_k^T Q(y_k)x_k + u_k^T R(y_k)u_k\right)\right] \text{''} \tag{2.4}$$

The standard assumptions are made on the matrices involved in the cost functions. More specifically, we assume that Q_k and Q are positive semidefinite, bounded matrix functions and R_k and R are strictly positive definite bounded matrix functions. For the discount factor it holds $a \in (0, 1)$.

2.3 Stability Analysis

In order to examine the mean square stability of a system in the form (2.2), let us introduce the following quantity:

$$P_k(C) = E[x_k x_k^T \chi_{y_k \in C}] \tag{2.5}$$

for C any Borelian subset of D. For any k, $P_k(\cdot)$ is a set function $P_k : \mathcal{B}(D) \to \mathbf{R}^{n\times n}$. It will be shown that $P_k(\cdot)$ is a symmetric matrix of signed measures.

The stability analysis is based on the evolution of the quantity $P_k(\cdot)$. Thus, we introduce the space in which $P_k(\cdot)$ belongs. Denote by X' the space of signed measures on $(D, \mathcal{B}(D))$ and by $|\cdot|$ the total variation norm. Then, the space of symmetric matrices of signed measures is defined as $X = \prod_{j=1}^{n(n+1)/2} X'$. Let us introduce on X, the norm $||\cdot||$, where $||[\mu_{ij}]|| = \sum_{i=1}^{n} \sum_{j=1}^{i} |\mu_{ij}|$, i.e. the sum of total variations. With this norm, X becomes a Banach space.

The evolution of $P_k(\cdot)$ is described using a linear operator $T = T_{A,\bar{K}} : X \to X$. It will be shown that T has the form:

$$(TP)(B) = \sum_{l=1}^{n} \sum_{m=1}^{n} \int_D \bar{A}_{l,m} \bar{K}(y, B) P^{l,m}(dy) \tag{2.6}$$

where $P^{l,m}$ is the l, m element of P and the formulae for the matrices $\bar{A}_{l,m}(y)$ are given in the proof of Theorem 1 (see Section 2.8).

The stability properties for a system in the form (2.2) depend on the spectral radius $r(T)$ ([Arv02]) of the operator T as, shown in the following theorem.

Theorem 1. *The quantity $P_k(\cdot)$ is a symmetric matrix of signed measures and its evolution is given by equation* (2.6). *Furthermore, the following hold:*

(i) The system is mean square exponentially stable if $r(T) < 1$.

(ii) If $r(T) > 1$ then the system is not mean square stable.

(iii) The system is mean square stochastically stable if and only if $r(T) < 1$.

(iv) The system is mean square exponentially stable then $r(T) < 1$.

Proof. See Section 2.8. □

The conditions of Theorem 1 involve the spectral radius of an infinite dimensional operator and thus, they are not easy to check. Based on Theorem 1, we are going to study further properties of exponentially stable systems that will allow us to obtain conditions that are easier to check.

The next proposition studies the uniformity of the exponential mean square stability on the initial conditions. Particularly, it is shown that in an exponential mean square stabile system, $E[x_k^T x_k]$ is going to be small in a finite number of steps irrespectively of the initial condition y_0. It is also shown that the converse is true.

Proposition 1. *The following are equivalent:*

(i) The system is exponentially mean square stable.

(ii) There exist a positive constant $M > 0$ and $a \in (0, 1)$ such that $E[x_k^T x_k] < Ma^k E[x_0^T x_0]$ for any x_0, y_0 random variables.

(iii) There exist an $a \in (0, 1)$ and a positive integer k_0 such that $E[x_{k_0}^T x_{k_0}] < a x_0^T x_0$, for any x_0, y_0 non-random initial conditions.

Proof. See Section 2.8. □

Condition (ii) of Proposition 1 has a stronger formulation than the definition of the mean square exponential stability, because the constants M and a are independent of the initial conditions, i.e. the exponential convergence to 0 is uniform on the initial conditions. Part (iii) of Proposition 1 shows that the mean square exponential stability is equivalent to the fact that the function $V(x) = x^T x$ is a “k_0-step Lyapunov function" for (2.2). Furthermore, Proposition 1, shows that a system which is not mean square exponentially stable could not be uniformly mean square stable.

This result also leads to a computational test for mean square exponential stability. The following Algorithm uses recursive computations to decide if $E[x_k^T x_k] < a x_0^T x_0$, for any x_0, y_0 non-random initial conditions.

Formula (2.7), computes recursively the matrix $E[A^T(y_0) \dots A^T(y_{k_0-1}) A(y_{k_0-1}) \dots A(y_0)]$ and thus Algorithm 1, is valid due to Proposition 1 (iii).

Algorithm 1 Stability Test

1: Set $L_{New}(y) = I$ and $Cnt = 1$.
2: Set $L(y) = L_{New}(y)$.
3: Compute:

$$L_{New}(y) = \int A^T(y')L(y')A(y')K(y, dy') \tag{2.7}$$

4: If $L_{New} < I$ for any $y \in D$ then return "The system is exponentially mean square stable" and halt.
5: Set $Cnt = Cnt + 1$
6: If $Cnt > MaxCnt$ then halt. Else go to Step 2.

The algorithm does not necessarily halt. Thus, a maximum number of steps $MaxCnt$ is introduced.

An alternative way to deal with the stability problem is to study the system using "one step quadratic Lyapunov functions" of the form $V(x, y) = x^T M(y)x$. In Proposition 2 the exponential mean square stability is proved to be equivalent to the existence of a positive definite solution to a Lyapunov equation.

Proposition 2. *Consider a strictly positive definite matrix function $Q(y)$. The following are equivalent:*

(i) The system is exponentially mean square stable

(ii) There exists a bounded, strictly positive definite matrix function $M(y)$ that satisfies the Lyapunov equation:

$$A^T(y)E[M(y_1)|y_0 = y]A(y) - M(y) = -Q(y) \tag{2.8}$$

Proof. See Section 2.8. □

Let us note that if the stochastic kernel is continuous, i.e. it could be described using densities, the Lyapunov equation (2.8) becomes a linear vector integral equation of Fredholm type. Thus, in several cases (2.8) could be solved numerically.

Remark 1. *The techniques applied to study the mean square stability of MJLS with discrete state space, could not be applied to a MJLS with general state space. More precisely, the quantity involved in the stability analysis of MJLS with discrete state space is, in several cases, identically zero when applied to MJLS with general state space. Thus, it is not appropriate for stability analysis.*

Another interesting notion is stabilizability, i.e. the existence of a stabilizing control law. It refers to a system under control in the form:

$$x_{k+1} = A(y_k)x_k + B(y_k)u_k \tag{2.9}$$

Let us define stabilizability:

Definition 2. *The system under control* (2.9) *is stabilizable, if there exists a bounded matrix function* $L(y)$ *such that the closed loop system given by:*

$$x_{k+1} = [A(y_k) + B(y_k)L(y_k)]x_k \tag{2.10}$$

is mean square exponentially stable. In this case, the pair $(A(\cdot), B(\cdot))$ *will be called stabilizable.*

2.4 Relations among the Stability Notions and Relations to other Control Problems

The relations among the stability notions of Definition 1 are then studied. Theorem 1 shows that $(iii) \Leftrightarrow (iv) \Rightarrow (ii) \Rightarrow (i)$. Next lemma shows that $(iii) \Rightarrow (v)$.

Lemma 1. *Exponential mean square stability implies almost sure stability.*

Proof. See Section 2.8. □

An example of a system which is almost sure stable but not mean square stable is given in [JC90]. Thus, $(v) \not\Rightarrow (i)$ and $(v) \not\Rightarrow (ii)$. In [CF95] an example is given, showing that $(ii) \not\Rightarrow (iii)$.

The following example shows that $(i) \not\Rightarrow (ii)$.

Example 1. *Let* $D = \mathbb{N}$ *be the state space of the Markov chain. The evolution of* y_k *is given by* $y_{k+1} = y_k + 1$ *and the function (sequence)* $A(\cdot)$ *is given by* $(3, 0, 3, 3, 0, 3, 3, 3, 0, \dots)$.

The system is pointwise mean square stable, since for any non-random initial condition, it holds $E[x_k^T x_k] \to 0$ *in finite steps. To see that the system is not mean square stable, take an initial condition with distribution* $Pr(y_0 = i_\nu) = 1/2^\nu$, *where* i_ν *is the first element after* ν *zeros. It is easy to see that* $E[x_k^T x_k] \to \infty$. *Hence, the system is not mean square stable.*

The following example shows that $(i) \not\Rightarrow (v)$.

Example 2. *Consider the infinite graph shown in Figure 2.1: Each node corresponds to a state of the state space of the Markov chain. The number inside the node is the value of the function* $A(\cdot)$ *on that member of the state space. The numbers on the arrows correspond to the transition probabilities. On any part of the tree that does not have any forks the product of the members is* $1/2$. *For any non-random initial condition it holds* $E[x_k^T x_k] \to 0$. *However, for* $x_0 = 1$ *and* y_0 *the root, it holds:*

$$Pr\left(\limsup_k x_k = 1\right) = 1.$$

Thus, $(i) \not\Rightarrow (v)$. *It remains open whether or not* $(ii) \not\Rightarrow (v)$.

The results are summarized in the following corollary.

Corollary 1. *The stability notions are related as follows:*

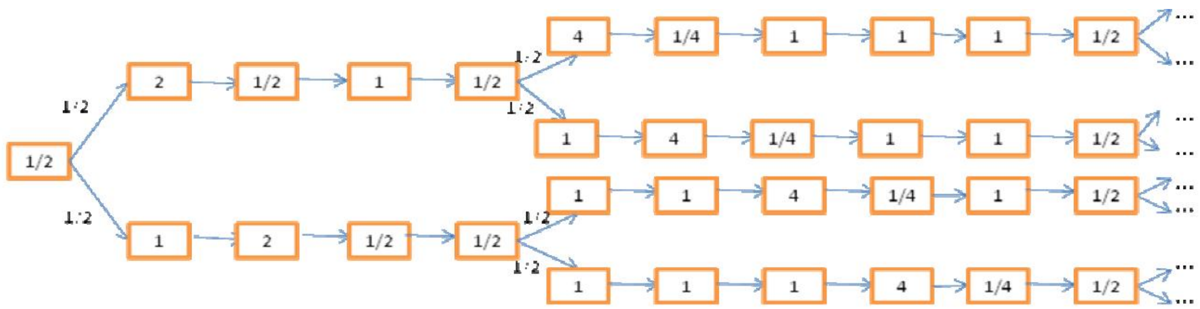


Figure 2.1: Evolution of the system in Example 2

(i) It holds: $(v) \Leftarrow (iii) \Leftrightarrow (iv) \Rightarrow (ii) \Rightarrow (i)$.

(ii) It holds: $(v) \not\Rightarrow (i)$, $(v) \not\Rightarrow (ii)$, $(ii) \not\Rightarrow (iii)$ *and* $(i) \not\Rightarrow (v)$.

In the following examples, some problems related to control theory are stated in terms of MJLS.

Example 3. *Uniform asymptotic stability of linear time varying systems.*
Consider a linear time varying system:

$$x_{k+1} = A_k x_k.$$

The uniform asymptotic stability [Kha02] for this system is equivalent to the mean square stability of the MJLS:

$$x_{k+1} = \tilde{A}(y_k)x_k,$$

where y_k *is a Markov chain with state space* $D = \mathbb{N}$ *and deterministic transition* $y_{k+1} = y_k + 1$ *and the matrix function* $\tilde{A}(\cdot)$ *satisfies* $\tilde{A}(y) = A_y$.

Example 4. *Uniform asymptotic stability of linear uncertain systems.*
Consider the linear uncertain system:

$$x_{k+1} = A(\theta)x_k,$$

where $\theta \in D \subset \mathbb{R}^l$. *The uniform asymptotic stability of the uncertain system is equivalent to the mean square stability of the MJLS:*

$$x_{k+1} = \tilde{A}(y_k)x_k,$$

where y_k *is a Markov chain with state space* D *and transition* $y_{k+1} = y_k$ *and the matrix function* $\tilde{A}(\cdot)$ *satisfies* $\tilde{A}(y) = A(y)$.

Example 5. *Convergence to zero of all periodic products of two matrices..*
This problem was stated in a connection with the finiteness conjecture [BTV03]. Consider two matrices A_1 *and* A_2. *Consider also the set* S *of finite sequences of* $1's$ *and* $2's$. *Consider finally*

the set $D = \{(s,m) : s \in S, m \in \mathbb{N}$ and $m \in dom(s)\}$ which will serve as the state space of the Markov chain. The transition of the Markov chain if $y_k = (s,m)$ is given by:

$$y_{k+1} = \begin{cases} (s, m+1) & \text{if } m \text{ is not the last element of the sequence} \\ (s, m') & \text{otherwise.} \end{cases},$$

where m' is the first element of the sequence. The value of $\tilde{A}(y_k)$ is given by:

$$\tilde{A}(y_k) = \begin{cases} A_1 & \text{if } s_m = 1 \\ A_2 & \text{if } s_m = 2 \end{cases},$$

The pointwise mean square stability of the MJLS is equivalent to the convergence to zero of all periodic products of the two matrices.

Example 6. *Uniform asymptotic stability of switched linear systems under arbitrary switching [LM99].*

Consider a set of matrices $A_1, \ldots, A_m$. Consider also the following switched system:

$$x_{k+1} = A_{\sigma(k)} x_k$$

where $\sigma(k)$ is the switching signal. The system is uniformly asymptotically stable under arbitrary switching iff there exist a positive integer N such that $\|x_N\| < \epsilon < 1$ for any x_0 such that $\|x_0\| = 1$ and any switching signal $\sigma(k) : \mathbf{N} \to \{1, \ldots, m\}$.

Consider also the logistic map $f : [0,1] \to [0,1]$ with $f(z) = 4z(1-z)$. Let $\bar{m}$ be a minimum positive integer that $2^{\bar{m}} \geq m$. The proof is based on the following fact:

Fact 1*: For any switching sequence, $\sigma_1, \ldots \sigma_N$, there exist a real number, $z_0 \in [0,1]$ such that:*

$$\begin{aligned} z_0 &\in \left[\frac{\sigma_1 - 1}{2^{\bar{m}}}, \frac{\sigma_1}{2^{\bar{m}}}\right), \\ f^{\bar{m}}(z_0) &\in \left[\frac{\sigma_2 - 1}{2^{\bar{m}}}, \frac{\sigma_2}{2^{\bar{m}}}\right), \\ &\vdots \\ f^{(N-1)\bar{m}}(z_0) &\in \left[\frac{\sigma_N - 1}{2^{\bar{m}}}, \frac{\sigma_N}{2^{\bar{m}}}\right) \end{aligned}$$

Let us denote by $S_i = \left[\frac{i-1}{2^{\bar{m}}}, \frac{i}{2^{\bar{m}}}\right)$. The Fact 1 could be proved inductively and uses that $f^{\bar{m}}[S_i] = [0,1]$ for any $i = 1, \ldots, 2^{\bar{m}}$.

Consider now the function $\bar{A} : [0,1] \to \mathbf{R}^{n \times n}$ with:

$$\bar{A}(y) = \begin{cases} A_i & \text{if } y \in S_i, i = 1, \ldots, m-1 \\ A_m & \text{otherwise} \end{cases}$$

The uniform asymptotic stability of the switched linear system, is equivalent with the exponential mean square stability of the following system:

$$x_{k+1} = \bar{A}(y_k)x_k$$
$$y_{k+1} = f^{\bar{m}}(y_k)$$

The transition of the variable y_k is deterministic and thus it is a Markov chain with state space $[0, 1]$. *Fact 1 and Proposition 1 complete the proof.*

2.5 Optimal Control Problems

At first, the finite horizon linear quadratic control problem is studied. The system under control is slightly more general than the system given by (2.1):

$$x_{k+1} = A_k x_k + B_k u_k + w_k \tag{2.11}$$

i.e. time varying matrices A and B are allowed. The problem under consideration is to find a control law $u_t = \gamma(x_t, y_t, t)$ that minimizes the cost function given by (2.3). The solution to the finite horizon linear quadratic control problem is given recursively by the following equations:

$$K_{N+1}(y_{N+1}) = Q_{N+1}(y_{N+1}) \tag{2.12}$$

$$\begin{aligned} \Lambda_{k+1}(y_k) =& E[K_{k+1}(y_{k+1}|y_k)] \\ =& \int_D K_{k+1}(y')\bar{K}(y_k, dy') \end{aligned} \tag{2.13}$$

$$K_k = Q_k + A_k^T[\Lambda_{k+1} - \Lambda_{k+1}B_k \cdot (R + B_k^T\Lambda_{k+1}B_k)^{-1}B_k^T\Lambda_{k+1}]A_k \tag{2.14}$$

$$L_k = -(R + B^T\Lambda_{k+1}B_k)^{-1}B_k^T\Lambda_{k+1}A_k \tag{2.15}$$

$$u_k = L_k(y_k)x_k \tag{2.16}$$

Proposition 3. *Consider the system given by* (2.11) *and the cost criterion* (2.3). *Then, the control law computed recursively using the equations* (2.12) - (2.16) *is optimal.*

Proof. Application of dynamic programming. □

Let us now study the infinite horizon linear quadratic control problem, i.e. minimize (2.4) subject to (2.1). The solution of this problem depends on the following Riccati integral equation:

$$\begin{aligned} K(y) = Q(y) + A^T(y)[a\Lambda(y) - a\Lambda(y)B(y)\cdot \\ \cdot (R(y)/a + B^T(y)\Lambda(y)B(y))^{-1}B(y)^T\Lambda(y)]A(y) \end{aligned} \tag{2.17}$$

where:

$$\Lambda(y) = E[K(y_{k+1})|y_k = y] = \int_D K(y')\bar{K}(y, dy') \tag{2.18}$$

The following Theorem 2 characterizes the optimal control policy in terms of the solution of the Riccati equation (2.17). Before stating Theorem 2, let us denote by $J_\mu(x, y)$ the value of the cost function (2.4) when $u_k = \mu(x_k, y_k, k)$ and $x_0 = x$, $y_0 = y$. Let us also denote by $J^*(x, y)$, the optimal value of the cost function (2.4).

Theorem 2. *Consider the system given by equation* (2.1) *and the cost function* (2.4)*. Then:*

(i) Assume that there exists a policy μ that makes the criterion (2.4)*, finite i.e. $J_\mu(x, y) < \infty$ for any x, y. Then, optimal cost has the form $J^*(x, y) = x^T K(y)x + c(y)$, where $K(y)$ satisfies the Riccati equation* (2.17)*. Furthermore, the optimal control is given by:*

$$u_k = L(y_k)x_k = -(B^T(y_k)\Lambda(y_k)B(y_k) + R(y_k)/a)^{-1} \cdot B^T(y_k)\Lambda(y_k)A(y_k)x_k \tag{2.19}$$

(ii) Conversely, assume that a bounded function $K(y)$ satisfies the Riccati equation (2.17)*. Assume that the undisturbed closed loop system given by:*

$$x_{k+1} = (A(y_k) + B(y_k)L(y_k))x_k \tag{2.20}$$

is mean square exponentially stable. Then the policy given by (2.19) *is optimal.*

Proof. The proof shares many ideas with [CFM05] or [CF95]. The basic difference is the proof of the finiteness of the cost when the controller given by (2.19) it is used. That proof uses essentially the results of Section 2.3. The differences are, however, of a technical character and thus the detailed proof is omitted. □

Theorem 2 characterizes the optimal control law when $a \in (0, 1)$. However, the equations (2.17) - (2.19) provide also the optimal controller when $a = 1$ and $w_k = 0$.

Remark 2. *(i) The existence of a policy μ that makes $J_\mu(x, y)$ finite for any x, y, is implied by the stabilizability of the pair $(\sqrt{a}A(\cdot), B(\cdot)))$. Furthermore, it is equivalent if Q is strictly positive definite.*

(ii) Equation (2.17)*, is a new form of Riccati equation. Specifically, it is a nonlinear vector integral equation. The solution of equation* (2.17) *could be approximated using the value iteration method (ex. [Ber07]).*

(iii) If the matrices $A(\cdot)$, $B(\cdot)$, $Q(\cdot)$, $R(\cdot)$ are continuous and the stochastic kernel is strongly Feller [MT93], then any solution of (2.17) *is continuous.*

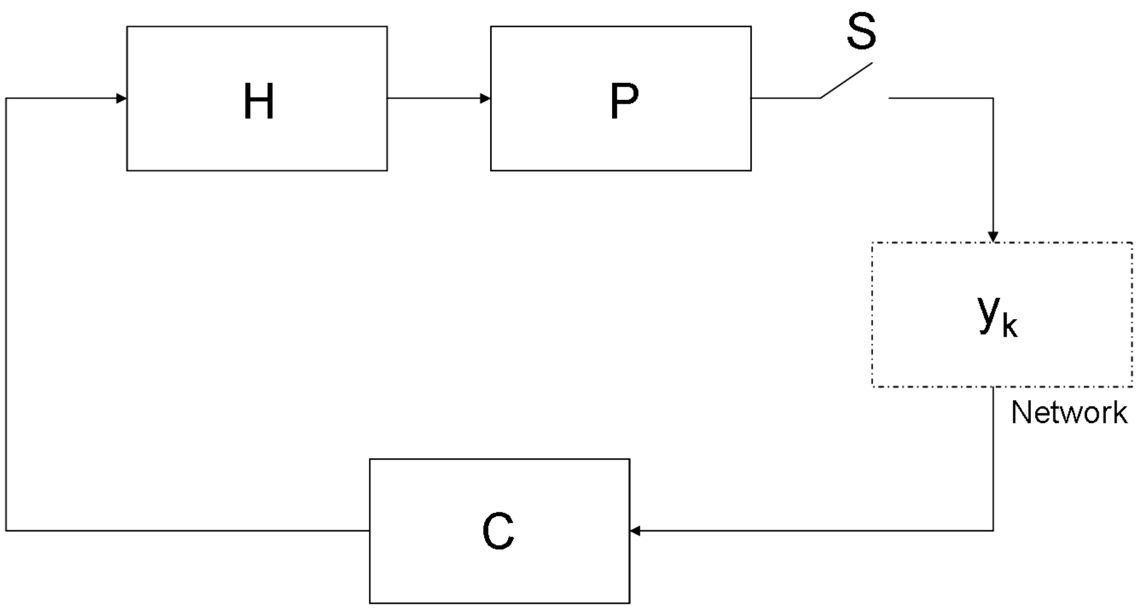


Figure 2.2: Networked Control System

2.6 An Application

In this section, we study a simple example of application of MJLS with general state space Markov chain on systems with random delays. Examples of such systems include Networked Control Systems (NCS) (ex. [GC10], [Nil98], [HBSW07] and [LHN09]) and distributed optimization algorithms (ex. [BP93]). Particularly, we study a simple model of NCS with dependent delays. Furthermore, a simple numerical example is given, illustrating the solvability of the equations derived in sections 2.3 and 2.5.

Consider a NCS as in Figure 2.2 ([LHN09] or [HBSW07]). The continuous time plant P is controlled by a controller C. The sampler S works at a constant rate and the time intervals among the sampling times have length T. The information is transmitted from the sampler to the controller through a communication channel, which introduces a random delay. Let us denote the delay of the transmission of the k-th measurement as y_k. In order to keep the model as simple as possible, we consider time delays only from the sampler to controller and we assume that the time delay introduced by the channel, is less than the sampling time, i.e. $y_k \in [0, T]$. A Markovian model for the time delay is introduced, i.e. there exists a stochastic kernel $\bar{K}(\cdot, \cdot)$ such that $\bar{K}(y, B) = Pr(y_k + 1 \in B | y_k = y)$. Finally, assume that the zero order hold H is event triggered, i.e. it holds the old value of the control until the new value comes.

The system under control P, is linear and its equation is given by:

$$\dot{x}_c = A_c x_c + B_c u_c \tag{2.21}$$

We will study the discretization of the system (2.21) on the time steps $t = kT$, for $k = 0, 1, \ldots$. Let us, thus, define $x_k = x_c(kT)$ and $u_k = u_c(kT + y_k)$, i.e. the control value obtained using the measurement of x_k. The system (2.21) has an input $u_c(t) = u_{k-1}$, on the time interval

$t \in [kT, kT + y_k)$ and $u_c(t) = u_k$ on the interval $t \in [kT + y_k, (k+1)T)$. Thus, in order to describe the evolution of x_k, we use the augmented state vector $\tilde{x}_k = [x_k^T \; u_k^T]^T$. The evolution of $\tilde{x}_k$ is given by:

$$\tilde{x}_{k+1} = A_d(y_k)\tilde{x}_k + B_d(y_k)u_k \tag{2.22}$$

where:

$$A_d(y_k) = \begin{bmatrix} e^{A_cT} & \int_0^{y_k} e^{A_c(T-\tau)}B_c d\tau \\ 0 & 0 \end{bmatrix},$$
$$B_d(y_k) = \begin{bmatrix} \int_0^{T-y_k} e^{A_c(T-y_k-\tau)}B_c d\tau \\ I \end{bmatrix} \tag{2.23}$$

Thus, the problem is reduced to the design of a controller for the MJLS (2.22) and the techniques of sections 2.3 and 2.5 could be applied. In the following example, a controller is designed for a simple system under control P.

Example 7. *Consider the plant P described by:*

$$\dot{x}_c = 2x_c + u_c$$

Assume that $T = 1$, the maximum delay is 0.5 and the stochastic kernel is described using the density function:

$$f_y(z) = \begin{cases} 4z/y & \text{if } 0 \leq z < y \leq 0.5 \\ 4 & \text{if } 0 \leq y = z \leq 0.5 \\ 4(1-2z)/(1-2y) & \text{if } 0 \leq y < z \leq 0.5 \end{cases}$$

i.e. $\bar{K}(y, B) = \int_B f_y(z)dz$. The matrices of the discretized system (2.22) are given by:

$$A_d = \begin{bmatrix} e^2 & e^2(1-e^{-2y})/2 \\ 0 & 0 \end{bmatrix}, B_d = \begin{bmatrix} (e^{2-2y}-1)/2 \\ 1 \end{bmatrix}$$

A LQ control law is designed. The matrices describing the quadratic criterion (2) are given by $Q(y_k) = diag(3, 0)$ and $R(y_k) = 1$. For the matrix functions $A(y), B(y), Q(y)$ and $R(y)$, the Riccati integral equation (2.17) is solved using the value iteration method [Ber07]. The components of the gain vector $L(y_k) = \begin{bmatrix} L_1(y_k) & L_2(y_k) \end{bmatrix}$ are plotted in Figure 2.3.

The closed loop system is simulated and several sample paths are presented in Figure 2.4.

Remark 3. *(i) Example 7 was considered, in order to illustrate that equations of section 2.5 could be used to design LQ control laws for NCS. with dependent time delays described by a Markov chain with continuous state space. It is worth noting that this model for the delays is more general than the models used in the literature.*

(ii) The model used is as simple as possible Thus it can be generalized in several directions. For example the hypothesis $y_k \leq T$ could be dropped, a time delay from the controller to the ZOH can be considered or packet losses could be studied.

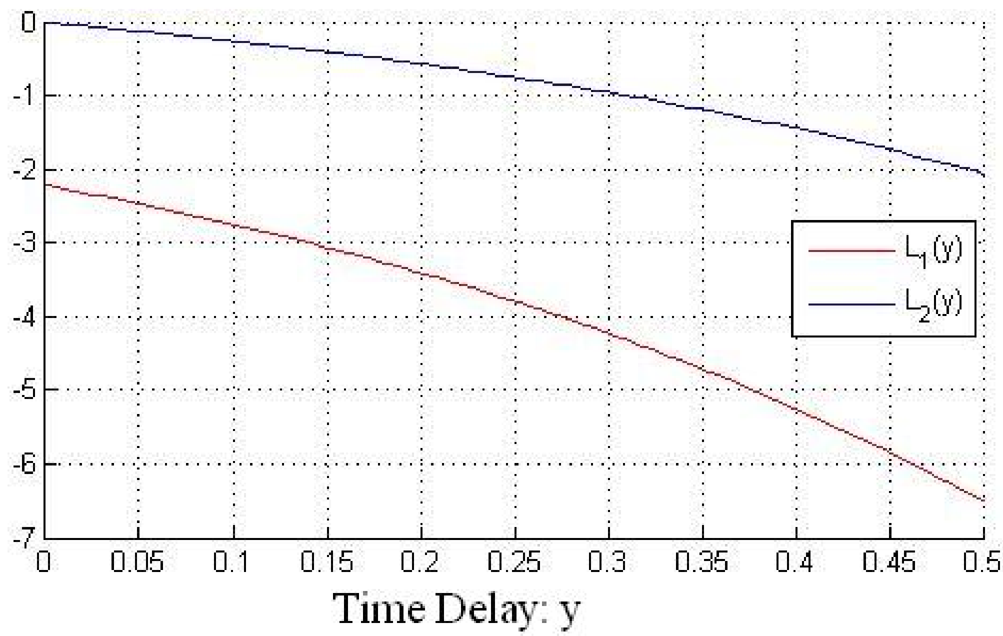


Figure 2.3: The gains $L_1(y)$ and $L_2(y)$

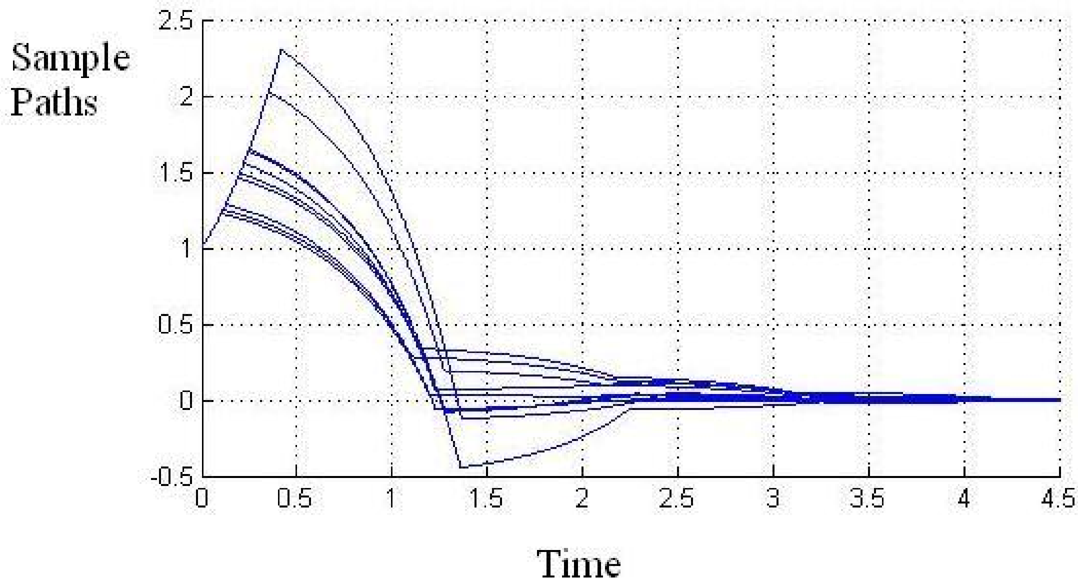


Figure 2.4: Several sample paths of the closed loop system

2.7 Conclusion

The class of MJLS with general state space was studied. The mean square exponential stability is characterized by the spectral radius of an infinite dimensional operator and proved to be uniform. An algorithm for testing stability was derived based on the uniformity result. The mean square exponential stability is also proved to be equivalent to the existence of a positive definite solution of a vector integral equation of Fredholdm type: the Lyapunov integral equation. The solution to the LQ control problem is characterized by a Riccati integral equation. The results derived, were used to design a controller to N.C.S. with dependent random delays. The model of the random delays is more general than those used in the literature.

2.8 Proofs

2.8.1 Proof of Theorem 1

Let us first show that P_k is a matrix of signed measures. It holds $P_k(\emptyset) = 0$. In order to show that each element of the matrix P_k is a signed measure, it suffices to show σ-additivity. For a sequence of disjoint sets $(A_m)_{m=1}^{\infty}$, σ-additivity could be shown using the functions $f_{i,j}^{k,m} : \Omega \to \mathbf{R}$ with $f_{i,j}^{k,m} = e_i^T x_k x_k^T e_j \chi_{y \in \bigcup_{p=1}^{m} A_p}$ and dominated convergence theorem ([Bil08]).

In order to derive the formula for T, we make the following computations:

$$\begin{aligned} P_{k+1}(B) &= E[A(y_k)\phi(y_k)A^T(y_k)\chi_{y_{k+1}\in B}] \\ &= \int_D A(y)\phi(y)A^T(y)\bar{K}(y,B)\mu_k(dy) \end{aligned}$$

where, $\phi(y_k)$ is a version of $E[x_k x_k^T | y_k]$ and μ_k is the distribution of y_k. Let $\phi_{l,m}$ and $P_k^{l,m}$ be the l, m, elements of ϕ and P_k respectively. Let also $\bar{A}_{l,m}(y)$ be functions such that $A(y)\phi(y)A^T(y) = \sum_{l=1}^{n}\sum_{l=1}^{n} \bar{A}_{l,m}(y)\phi_{l,m}(y)$. Thus, since $\phi_{l,m}$ is the Radon - Nikodym derivative $dP_k^{l,m}/d\mu_k$, it holds:

$$P_{k+1}(B) = \sum_{l=1}^{n}\sum_{m=1}^{n}\int_D A_{l,m}\bar{K}(y,B)P^{l,m}(dy)$$

which completes the proof of equation (2.6).

Assuming that $r(T) < 1$ spectral formula implies $\|P_k(\cdot)\| \to 0$ exponentially. Thus, the inequality $E[x_k^T x_k] \leq \|P_k(\cdot)\|$ completes the proof of (i).

To prove (ii) let us assume that $r(T) > 1$. Then there exists an initial value $\bar{P}$ such that $T^k\bar{P} \to \infty$ as $k \to \infty$. Let $\bar{P}^+$, $\bar{P}^-$ be the Hahn decomposition of $\bar{P}$. Without loss of generality $T^k\bar{P}^+ \to \infty$ as $k \to \infty$. It is not difficult to show that there exist x_0, y_0 random variables such that $\bar{P}^+(B) = E[x_0 x_0^T \chi_{y_0 \in B}]$. The proof of (ii) is completed using the following fact:

$$\|P_k(\cdot)\| \leq 2nE[x_k^T x_k]$$

To prove (iii) let us observe that the “only if" part of (iii) is a direct consequence of [Kub85] Lemma 1 and the “if part" is a consequence of (i).

The proof of (iv) follows directly from (iii).

2.8.2 Proof of Proposition 1

We first show that (i) implies (ii). Using Fact 1, it is not difficult to show that $E[x_k^T x_k] \leq 2nE[x_0^T x_0]\|T^k\|$. The spectral formula implies that there exist an integer k_0 and a positive constant ϵ such that $\|T^{k_0}\| < 1-\epsilon$. Thus using Euclidian division of k by k_0 we conclude to the desired result:

$$E[x_k^T x_k] < Ma^k E[x_0^T x_0]$$

with $M = 2n \max_{k=0,\ldots,k_0} \|T^k\|/(1-\epsilon)$ and $a = (1-\epsilon)^{1/k_0}$.

The fact that (ii) implies (iii), is obvious.

It remains to show that (iii) implies (i). Let us introduce the following quantities:

$$\Phi_i(y) = E[A^T(y_0)\ldots A^T(y_{i\cdot k_0})A(y_{i\cdot k_0})A^T(y_0)]$$

i,e, the expectation of the product of $2ik_0$ matrices. It could be shown inductively that $\Phi_i(y) < a^i I$. Let N be a positive integer such that $2na^N = \bar{a} < 1$. Then for every x_0, y_0 random variables it holds: $E[x_{k_0\cdot N}^T x_{k_0\cdot N}] \leq a^N E[x_0^T x_0]$ and using Fact 1 we obtain:

$$\|P_{k_0\cdot N}(\cdot)\| \leq 2nE[x_{k_0\cdot N}^T x_{k_0\cdot N}] \leq \bar{a}E[x_0^T x_0] \leq \bar{a}\|P_0(\cdot)\|$$

Thus, $r(T) < \bar{a}^{1/(iN)} < 1$ which completes the proof.

2.8.3 Proof of Proposition 2

We first show that (i) implies (ii). Let us consider a sequence of matrix functions $M_{N-k}(y)$ given by:

$$x_k^T M_{N-k} x_k = E\left[\sum_{t=k}^{N} x_t^T Q(y_t) x_t | x_t, y_t\right]$$

and their limit $M(y) = \lim_{N\to\infty} M_{N-k}(y)$. The matrix function Q is bounded, thus there exists a positive constant c such that $Q(y) < cI$. It holds:

$$x_0^T M_N x_0 \leq cE\left[\sum_{t=0}^{N} x_t^T x_t\right] \leq \frac{c\bar{M}}{1-a} x_0^T x_0$$

where $\bar{M}$ and a satisfy the Proposition 1 (ii). Thus $M(y)$ is bounded. Furthermore, it holds:

$$x_0^T M_N(y_0)x_0 - E[x_1 M_{N-1} x_1 | x_0, y_0] = x_0 Q(y_0) x_0$$

Taking limits, we conclude to Equation (2.8).

It remains to show that (ii) implies (i). Following the same steps as Theorem 2.1 of [JC90] we conclude that:

$$E[x_k^T x_k] \leq \frac{c_1}{c_2} a^k E[x_0^T x_0]$$

where c_1 and c_2 are positive constants such that $c_1 I < Q(y) < c_2 I$ and

$$a = 1 - \min_{y \in D} \{ \frac{\lambda_{min}(Q(y))}{\lambda_{max}(Q(y))} \} \in (0, 1)$$

Thus, Proposition 1 (iii) completes the proof.

2.8.4 Proof of Lemma 1

Consider the sets:

$$B_{k,l} = \{\omega \in \Omega : \|x_k\| > 1/l\}.$$

Markov inequality implies $\sum_{k=0}^{\infty} Pr(B_{k,l}) < \infty$. Thus, using Borel-Cantelli lemma we obtain:

$$Pr\left(\limsup_k B_{k,l}\right) = 0.$$

Hence:

$$Pr\left(\{\omega \in \Omega : x_k \to 0\}^c\right) = Pr\left(\cup_{l=1}^{\infty} \limsup_k B_{k,l}\right) = 0,$$

which completes the proof.

Chapter 3

LQ Nash Games with Random Entrance.

This chapter studies Dynamic Games with randomly entering players, staying in the game for different lengths of time. Particularly, a class of Discrete Time Linear Quadratic (LQ) Games, involving a major player who has an infinite time horizon and a random number of minor players is considered. The number of the new minor players, entering at some instant of time, is random and it is described by a Markov chain. The problem of the characterization of a Nash equilibrium, consisting of Linear Feedback Strategies, is reformulated as a set of coupled finite and infinite horizon LQ optimal control problems for Markov Jump Linear Systems (MJLS). Sufficient conditions characterizing Nash equilibrium are then derived. The problem of Games involving a large number of minor players is then addressed using a Mean Field (MF) approach and asymptotic ε - Nash equilibrium results are derived. Sufficient conditions for the existence of a MF Nash equilibrium are finally stated.

3.1 Introduction

For the most of the dynamic game models, the time interval during which the players are involved in the game, as well as the number of players that participate in the game at each instant of time are quite structured. For example, in finite or infinite horizon dynamic games (e.x. [SH69b], [BO99]) all the players participate in the game for identical time intervals. In overlapping generation games ([Sam58], [BS80], [JY99]), a known number of players of a new generation enters into the game at each time step and stays for a certain period of time. Several attempts to impose less structure on the players' time intervals or on the number of players that participate in the game have been made. For example, in games with population uncertainty or in Poisson games [Mye98] the number of players that participate in the game is not known a priori. Games with random horizon have been studied in a repeated game setting in [AU05] and in a differential game setting in [YP11]. In this class of games, the time intervals in which the players are involved in the game are identical; however, the duration of the game is random. In [JY05], a game with overlapping generations involving players which remain in the game for two time steps is considered. The number of players of each generation is however random.

The current work studies games with random entrance in a LQ setting and imposes less structure on the time intervals during which the players participate in the game, as well as on the number of the active players at each time step. In particular we, consider a player with infinite time horizon, called the major player and many players with finite time horizons, called minor players. The number of new players entering the game at any time step is a random variable that has a distribution which depends only on the number of active players at that moment. The random entrance is, thus, described by a Markov chain. The problem considered here is the characterization of the Perfect Nash equilibria. After that, we study the case where the number of minor players is very large. A Mean Field (MF) approximation is used to characterize strategies, which are asymptotically optimal as the number of new minor players in each step tends to infinity. The equations derived using MF approximation are often much easier.

The structure proposed for the participation of the minor players in the game is not unusual in practice. There are several examples of game situations where there is a long living agent or institution which, at each time step, interacts with a number of agents and the interaction with each agent is maintained for a certain, rather small amount of time. For instance, a bank that gives loans to households may be considered as a major player with an infinite horizon and each person that assumes a loan as a minor player with a finite pre-specified time horizon. Another example is a liberalized energy market in which there is a public power corporation with an infinite time horizon and many renewable energy producers that have a permission to enter the system for a certain amount of time [KP10]. A third example is University-Student Games [Pap12], where the students of each semester stand for the minor players and the university as a major player. Cases involving players with different time horizons were studied also in [JPAK12], [PAKJ13]. Other examples involve the study of repeated games with long-run and short run-players [FKM90], such as the chain store game and the study of reputation effects (ex. [KW82], [FT91b]).

In the current work, the problem of random entrance is reduced to the study of coupled finite and infinite horizon LQ problems for Markov Jump Linear Systems (MJLS). Thus, the Nash equilibria are characterized using appropriate coupled Riccati type equations. There are two types of coupling; the first corresponds to the Markov Jump character of the optimal control problems and the second to the LQ Game coupling. In the case of a large number of players, the Mean Field approach involves the statement of approximate optimal control problems assuming an infinity of players. In that case, ε - equilibrium results are proved. The method used to prove the ε - equilibrium results is based on some results connecting the stability and the LQ control of MJLS with the convergence of a sequence of Markov chains. These results are proved in the Section 3.8 and are also of independent interest.

The rest of the paper is organized as follows: In section 3.2, the dynamics and the cost functions of major and minor players are defined. In Section 3.3, the optimal control problems that the participants of the game face are reformulated as a set of coupled finite and infinite horizon LQ problems for MJLS. In Section 3.4, sufficient conditions on a set of linear feedback strategies to constitute a Nash equilibrium are derived. In Section 3.5, the problem with a large number of players is approximated using a Mean Field model. Then some ε - Nash equilibrium results are obtained. In Section 3.6, an algorithm for computing a Nash equilibrium is stated and it is shown to converge under certain conditions. Furthermore, some numerical examples are studied. In Section 3.7, we conclude. The proofs of some results in the text are relegated to Section 3.8.

Notation: The transpose of a matrix is denoted by $\cdot^T$. In all the text, except Section 3.6.2, $\|\cdot\|$ denotes the usual 2-norm. The underlying probability space is denoted by $(\Omega, \mathcal{F}, Pr)$ and the spectral radius of a matrix or an operator by $r(\cdot)$. The Borelian subsets of a set D are denoted by $\mathcal{B}(D)$. The notion of a stochastic kernel is also used to describe the evolution of a Markov chain. Particularly, for a Markov chain y_k with state space D, we denote by $\bar{K}(\cdot,\cdot) : D \times \mathcal{B}(D) \to \mathbb{R}$ the stochastic kernel, i.e. $\bar{K}(y, B) = Pr(y_{k+1} \in B | y_k = y)$, for $y \in D$ and $B \in \mathcal{B}(D)$. The fact that the random variable y has probability distribution F is denoted by $y \sim F$ and the weak convergence of probability measures is denoted by " $\Rightarrow$ ". The Kronecker delta δ_{ij} is also used, where $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ otherwise. A matrix function $A : D \to \mathbb{R}^{n \times n}$ is called strictly positive definite if there exists a positive constant c such that $A(y) > cI$ for any $y \in D$. Finally, the i, j element of a matrix A is denoted by $A^{i,j}$. The dependence of a function on the Markov chain state variable y_k will be omitted, in several points in the text.

3.2 Description of the Game

At first, the random entrance of the minor players is described. The minor players have time horizon T, i.e. each one of them stays in the game for T time steps. Consider a countably infinite set of minor players $\Delta = \{1, 2, \dots\}$. For any minor player $i \in \Delta$, let $t_i : \Omega \to \mathbb{N}$ be a stopping time describing the time step at which the player i enters the game. At the time step k, the number of the minor players that participate in the game may be described by the vector:

$$y_k = (N_k^0, \dots, N_k^{T-1})/s_c, \tag{3.1}$$

where $N_k^l = \# I_k^l$ and I_k^l is the set of players with entrance time $k - l$ and and s_c, which will be called the 'scale variable', is the maximum possible number of active players, i.e. $s_c = \max\{N_k^0 + \cdots + N_k^{T-1}\}$. Let us finally denote by I_k, the set of active players at time step k.

The number of new minor players that enter the game at the time step $k + 1$ is a random variable with a distribution depending on y_k. Thus, the random entrance is modeled by the Markov chain y_k having a finite state space. Let $1, 2, \dots, M$ be an enumeration of the state space and $\Pi = [p_{ij}]$ the transition matrix of the Markov chain, where p_{ij} is the ij entry of the matrix Π. We shall use the vector form (3.1) and the enumeration interchangeably.

Each player participating in the game has its own dynamic equation. The evolution of the state vector of each player depends on the state vectors of the currently active players in a symmetric manner. The dynamic equation of the major player is given by:

$$x^M(k+1) = A^M x^M(k) + \frac{1}{s_c}\sum_{i\in I_k} F^M x^i(k) + B^M u^M(k) + w^M(k), \tag{3.2}$$

where x^M and x^i are the state vectors of the major player and minor player i respectively. The stochastic disturbances $w^M(k)$ are zero mean, finite variance, i.i.d. random variables, independent of the state vectors. The initial condition for the major player is given by $x^M(0) = w^M(-1)$.

The dynamics of the minor player i is described by:

$$x^i(k+1) = Ax^i(k) + \sum_{j\in I_k} Fx^j(k)/s_c + Gx^M(k) + Bu^i(k) + w^i(k), \tag{3.3}$$

where the stochastic disturbances $w^i(k)$ are zero mean finite variance random variables, independent of the state vectors $x^M(k)$ and $x^i(k)$, $i \in I_k$. The initial values of the state vectors of the minor players are given by $x^i(t_i) = w^i(t_i - 1)$. The dependence of $w^i(k)$ with $w^j(k)$, $i \neq j$ is not disallowed.

In order to define the cost functions of the players, let us introduce the mean field quantities $z^l(k) = \sum_{i\in I_k^l} x^i(k)/s_c$ and the vector of the mean field quantities:

$$\tilde{z}(k) = \left[z^0, \ldots, z^{T-1}\right].$$

The cost function of the major player is given by:

$$\begin{aligned} J^M = E\Bigg\{\sum_{k=0}^{\infty} a^k \bigg[\Big[\left(x^M(k)\right)^T \tilde{z}^T(k)\Big] Q^M(y_k)\cdot \\ \cdot\Big[\left(x^M(k)\right)^T \tilde{z}^T(k)\Big]^T + (u^M(k))^T R^M u^M(k)\bigg]\Bigg\}, \end{aligned} \tag{3.4}$$

where $Q^M(y), y \in \{1,\ldots,M\}$ and R^M are positive semidefinite and positive definite matrices of appropriate dimensions respectively and $a \in (0,1)$ a discount factor.

For the minor player i, the cost function is given by:

$$\begin{aligned} J^i = E\Bigg\{ &(\tilde{x}^i(t_i+T))^T Q_f(y_{t_i+T})\tilde{x}^i(t_i+T) + \\ &\sum_{k=t_i}^{t_i+T-1} (\tilde{x}^i(k))^T Q(y_k)\tilde{x}^i(k) + (u^i(k))^T R u^i(k)\Bigg\}, \end{aligned} \tag{3.5}$$

where $\tilde{x}^i = [(x^M)^T\ \tilde{z}^T\ (x^i)^T]^T$, $Q_f(y)$ and $Q(y)$ positive semidefinite matrices of appropriate dimensions for any $y \in \{1,\ldots,M\}$ and R positive definite matrix of appropriate dimensions. Let us note that (3.5) indicates that the minor players have symmetric cost functions.

The problem considered here is the characterization of a Nash equilibrium that satisfies the Dynamic Programming (Perfect equilibrium [Sel75]). We shall focus on Linear Feedback Strategies (i.e. strategies with no memory; see [BO99], Def 5.2). Furthermore, due to the symmetry of the dynamic equations and cost functions, we shall further concentrate to strategies in the following form:

$$u^i = L^{1M}(k - t_i, y_k)x^M + \sum_{l=0}^{T_j - 1} L(l, k - t_i, y_k)z^l + \bar{L}(k - t_i, y_k)x^i(k), \tag{3.6}$$

and:

$$u^M = L^{MM}(y_k)x^M + \sum_{l=0}^{T-1} L^M(l, y_k)z^l. \tag{3.7}$$

The equations (3.6) and (3.7) serve only as a general form of the feedback strategies. Equations characterizing the gains $L^{1M}, L, \bar{L}, L^{MM}$ and L^M are determined in the next sections.

For the compactness of the presentation, the following notation will be used:

$$\tilde{L}^M(y) = [L^{MM}(y)L^M(0, y) \dots L^M(T - 1, y)] \tag{3.8}$$

$$\hat{L}_k(y) = [L^{1M}(k, y)L(0, k, y) \dots L(T - 1, k, y)\bar{L}(k, y)]$$

$$\tilde{L}(y) = [\hat{L}_0(y), \dots, \hat{L}_{T-1}(y)] \tag{3.9}$$

Remark 4. *A set of strategies in the form* (3.6)*,* (3.7) *has two types of symmetries. At first, the feedback gains are the same for all the minor players. Consider a strategy in that form. Then the control values depend on the mean field quantities z^l. Thus, the feedback gains corresponding to the players of the same entrance time are the same, which is a second form of symmetry. These symmetry assumptions are justified by the structure of the dynamics and the cost functions.* □

Remark 5. *Although we know that for Linear Quadratic games, closed loop Nash equilibria in nonlinear strategies may also exist, it is only the linear ones that survive if we introduce noise in the state equation or the measurements [Bas75]. This is the reason due to which we restrict our attention to Linear Feedback Nash equilibria.* □

Remark 6. *An interesting extension is to study games involving minor players with different time horizons*[1]*. This does not make the problem more difficult and all the results in this work can be immediately generalized to that case.*

Remark 7. *The major player may be viewed as a coordinator helping the stabilization of the overall system. An interesting alternative is to see the major player as a common adversary of*

[1] In fact, the first version of this work involved several types of minor players having different time horizons. However, for simplicity and clarity of the presentation reasons, after reviewers' recommendation, we restrict ourselves to the case where all the minor players have the same time horizon.

all the minor players in the spirit of [BTB⁺12][2]. The techniques used in the current work should, however, be adapted in order to study this alternative.

Remark 8. *An interesting extension would involve the continuous time analog of the current formulation. This can be done either by taking the limit of the discrete problems as the discretization time tends to zero in the spirit of [TH11] or by stating the corresponding problem in continuous time directly.*

Remark 9. *The study of Stackelberg equilibria with the major player as a leader is a related interesting problem. The same techniques can be used in order to characterize feedback Stackelberg equilibria.*

Remark 10. *The only necessary measurements for a player to implement a strategy in the form of* (3.6) *or* (3.7) *are the value of the state vector of the major player x^M, the mean field quantities $\tilde{z}(k)$, the value of its own state vector and the value of the Markov chain state variable y_k. Thus, we shall make the following assumption:*

Assumption 1. *All the players have access to the current values of x^M, $\tilde{z}$ and y_k. Furthermore, each player can measure its own state vector.* □

3.3 Optimal control problems

The problem of the Nash equilibrium characterization for LQ games with random entrance is converted to the problem of finding a solution to a set of coupled LQ control problems for MJLS. Particularly, the optimal control problems are stated in spaces of smaller dimensions and the random entrance problem is transformed to a random coefficients problem of a linear dynamic equation, depending on the Markov chain given by (3.1). This reduction is possible, due to the symmetric form of the dynamic equations, the cost functions and the control strategies. We shall assume that the players follow strategies in the general form (3.6), (3.7).

3.3.1 Optimal Control Problem for the Major Player

The evolution of the state vector of x^M and the cost function J^M depend only on x^M, $\tilde{z}$ and u^M. Assuming that the minor players use the strategies in the general form (3.6), the evolution of the components z_k^l of $\tilde{z}$ depend only on and x^M and $\tilde{z}$, as well. Hence, symmetry implies that the the evolution of the quantities in the cost function (3.4) can be described by a state vector of smaller dimension: $[(x^M)^T, \tilde{z}^T]^T$. The dynamics, after straightforward manipulations, is given by:

$$\begin{bmatrix} x^M(k+1) \\ \tilde{z}(k+1) \end{bmatrix} = \tilde{A}^M(y_k) \begin{bmatrix} x^M(k) \\ \tilde{z}(k) \end{bmatrix} + \begin{bmatrix} B^M \\ 0 \end{bmatrix} u^M(k) + W^M(k), \tag{3.10}$$

[2]This alternative was proposed by an anonymous reviewer.

where:

$$\tilde{A}^M(y_k) = \begin{bmatrix} \tilde{A}^M_{x^M x^M} & \tilde{A}^M_{x^M z^{1,0}} & \dots & \tilde{A}^M_{x^M z^{p,\bar{T}}} \\ \tilde{A}^M_{z^0 x^M} & \tilde{A}^M_{z^0 z^0} & \dots & \tilde{A}^M_{z^0 z^{\bar{T}}} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{A}^M_{z^{\bar{T}} x^M} & \tilde{A}^M_{z^{\bar{T}} z^0} & \dots & \tilde{A}^M_{z^{\bar{T}} z^{\bar{T}}} \end{bmatrix},$$

$\bar{T} = T - 1$. The entries of the first row of the matrix are given by: $\tilde{A}^M_{x^M x^M} = A^M$ and $\tilde{A}^M_{x^M z^l} = F^M$. The second row consists of zeros. For the rest of the entries it holds:

$$\tilde{A}^M_{z^{l+1} x^M} = \frac{N^l_k}{s_c}(G + BL^{1M}(l, y_k)), \text{ and}$$
$$\tilde{A}^M_{z^{l+1} z^{l'}} = \frac{N^l_k}{s_c}(F + BL(l', l, y_k)) + \delta_{l,l'}(A + B\bar{L}(l, y_k)).$$

The matrix $\tilde{A}^M$ depends on y_k through the terms N^l_k/s_c. Thus, the study of the optimal control problem of the major player, under the random entrance of the minor players is reduced to the study of the following infinite horizon LQ control problem for a MJLS:

OC Problem 1: "Minimize the cost function (3.4) subject to the dynamics (3.10) and (3.1)". □

3.3.2 Optimal Control Problem for the Minor Players

For the minor players, a similar reasoning applies. Consider a minor player i_0 with entrance time t_{i_0}. Assume that the other players use the feedback strategies in the general form (3.6) and (3.7). The evolution of the state vector and the cost of the player i_0 depend only on the quantities: x^M, x^{i_0}, $\tilde{z}$ and y. Thus, the cost function of player i_0 can be described using the state vector $\tilde{x}^{i_0} = [(x^M)^T \ \tilde{z}^T \ (x^{i_0})^T]^T$, which evolves according to:

$$\begin{aligned} \tilde{x}^{i_0}(k+1) = \tilde{A}(k - t_{i_0}, y_k)\tilde{x}^{i_0}(k) + \\ + \tilde{B}(k - t_{i_0}, y_k)u^{i_0} + W^{i_0}(k), \end{aligned} \tag{3.11}$$

where:

$$\tilde{A}(k - t_{i_0}, y_k) = \begin{bmatrix} \tilde{A}_{x^M x^M} & \tilde{A}_{x^M z^0} & \dots & \tilde{A}_{x^M z^{\bar{T}}} & \tilde{A}_{x^M x^{i_0}} \\ \tilde{A}_{z^0 x^M} & \tilde{A}_{z^0 z^0} & \dots & \tilde{A}_{z^0 z^{\bar{T}}} & A_{z^0 x^{i_0}} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \tilde{A}_{z^{\bar{T}} x^M} & \tilde{A}_{z^{\bar{T}} z^0} & \dots & \tilde{A}_{z^{\bar{T}} z^{\bar{T}}} & A_{z^{\bar{T}} x^{i_0}} \\ \tilde{A}_{x^{i_0} x^M} & \tilde{A}_{x^{i_0} z^0} & \dots & \tilde{A}_{x^{i_0} z^{\bar{T}}} & A_{x^{i_0} x^{i_0}} \end{bmatrix},$$

$\bar{T} = T - 1$. The entries of the matrix are computed by simple but lengthy calculations. For the first row, it holds:

$$\begin{aligned} \tilde{A}_{x^M x^M} &= A^M + B^M L^{MM}(y_k), \\ \tilde{A}_{x^M z^l} &= F^M + B^M L^M(l, y_k). \end{aligned}$$

The entries of the second row are zero. For the rest of the rows except the last one, the entries are given by:

$$\tilde{A}_{z^{l+1}x^M} = \frac{N_k^l}{s_c}(G + BL^{1M}(l, y_k)) - \frac{\delta_{l,k-t_{i_0}}}{s_c} BL^{1M}(l, y_k),$$

$$\tilde{A}_{z^{l+1}z^{l'}} = \delta_{l,l'}\left(A + \frac{N_k^l}{s_c}B\bar{L}(l, y_k)\right) +$$

$$+ \frac{N_k^l}{s_c}(F + BL(l', l, y_k)) - BL(l', l, y_k)\delta_{l,k-t_{i_0}}/s_c,$$

$$\tilde{A}_{z^{l+1}x^{i_0}} = -\frac{\delta_{l,k-t_{i_0}}}{s_c}B\bar{L}(l, y_k).$$

The entries of the last row are determined by (3.3). Thus, $\tilde{A}_{x^{i_0}x^M} = G$, $\tilde{A}_{x^{i_0}z^l} = F$ and $\tilde{A}_{x^{i_0}x^{i_0}} = A$.

The $\tilde{B}$ matrix is given by:

$$\tilde{B}(k - t_{i_0}, y_k) = [0\ \tilde{B}_{z^0}^T\ \dots\ \tilde{B}_{z^{T-1}}^T\ B^T]^T,$$

where $\tilde{B}_{z^{l+1}} = \delta_{l,k-t_{i_0}}B/s_c$.

The matrix $\tilde{A}$ is time varying and depends on the Markov chain through the terms N_k^l/s_c. Hence, the random entrance problem is transformed to a MJLS problem for the minor players, as well. The optimal control problem that a minor player faces is the following finite LQ control problem for a MJLS:

OC Problem 2: "Minimize the cost function (3.5) subject to the dynamics (3.11) and (3.1)".

□

Remark 11. *The state vectors of the dynamics of the major and minor players* (3.10) *and* (3.11) *have a much smaller dimention than the state vector consisting of all the active players. Furthermore, the dimensions of the state vectors do not depend on the number of players.* □

Remark 12. *The Markov jump character of the OC problems has two origins. The first is the random entrance and affects several terms such as $A^M_{z^{l+1}x^M}$, through the factor N_k^l/s_c. The second is the dependence of the Q matrices on y_k. Hence, the dependence of the Q matrices on y_k, does not make the problem more difficult.* □

3.4 Optimality Conditions and Nash Equilibrium

The optimality conditions for the OC Problems 1 and 2 are derived and then used to characterize a perfect Nash equilibrium. The general form of the solutions of the finite and infinite horizon discounted LQ control problems for MJLS can be found in [KP14a].

For the infinite horizon OC Problem 1, the optimality conditions are given in terms of a set of coupled Riccati type equations. Particularly, let us consider a set of matrices $K(y)$, $\Lambda(y)$, for

$y = 1, \dots, M$ such that:

$$K^M(y) = Q^M + (\tilde{A}^M)^T \Big[a\Lambda^M - a\Lambda^M \tilde{B}^M \cdot$$
$$\cdot \Big(R^M/a + (\tilde{B}^M)^T \Lambda^M \tilde{B}^M \Big)^{-1} (\tilde{B}^M)^T \Lambda^M \Big] \tilde{A}^M \tag{3.12}$$
$$\Lambda^M(y) = E[K^M(y_{k+1}) | y_k = y] = \sum_{j=1}^{M} p_{yj} K^M(j). \tag{3.13}$$

Let us also consider the control law given by:

$$u^M(k) = \tilde{L}^{M,\star}(y_k)[(x^M(k))^T \;\; \tilde{z}^T(k)]^T, \tag{3.14}$$

where:

$$\tilde{L}^{M,\star}(y) = -\Big((\tilde{B}^M)^T \Lambda^M \tilde{B}^M + R^M/a \Big)^{-1} \Lambda^M \tilde{A}^M. \tag{3.15}$$

The control law given by (3.14) is optimal, provided that it makes the cost function (3.4) finite. The criterion for the finiteness of the cost function (3.4) is given in terms of the closed loop matrix:

$$\tilde{A}^M_{cl} = \tilde{A}^M - \tilde{B}^M \tilde{L}^{M,\star}. \tag{3.16}$$

The finiteness criterion is based on the existence of a positive definite solution to a set of coupled Lyapunov equations (see equation (2.8)).

Finiteness Criterion 1: There exist positive definite matrices $S(i)$, $i = 1, \dots, M$ that solve the following set of coupled Lyapunov equations:

$$S(i) - \sum_{j=1}^{M} a p_{ij} (\tilde{A}^M_{cl}(i))^T S(j) \tilde{A}^M_{cl}(i) = I, \tag{3.17}$$

for $i = 1, \dots, M$. □

The optimality conditions for a minor player i_0 are given in terms of the solution of the following set of coupled Riccati type difference equations:

$$K_T(y) = Q_f(y), \tag{3.18}$$
$$\Lambda_{k+1}(y) = E[K_{k+1}(y_{k+1+t_{i_0}}) | y_{k+t_{i_0}}]$$
$$= \sum_{j=0}^{M} p_{yj} K_{k+1}(j), \tag{3.19}$$
$$K_k(y) = Q + (\tilde{A}(k))^T [\Lambda_{k+1} - \Lambda_{k+1} \tilde{B} (R + \tilde{B}^T \Lambda_{k+1} \tilde{B})^{-1} \tilde{B}^T \Lambda_{k+1}] \tilde{A}, \tag{3.20}$$

for $k = T-1, \dots, 0$. The optimal control law is then given by:

$$u^{i_0}(k + t_{i_0}) = \hat{L}^\star_k \tilde{x}^{i_0}(k + t_{i_0}), \tag{3.21}$$

where:

$$\hat{L}_k^\star(y) = -(R + \tilde{B}^T \Lambda_{k+1} \tilde{B})^{-1} \Lambda_{k+1} \tilde{A} \tag{3.22}$$

We are interested in Nash equilibria, i.e. a set of strategies, each of which is optimal given the others. Therefore, we define the notion of a consistent set of strategies.

Definition 3. *Consider a set of strategies in the general form* (3.6), (3.7)*, with gains $\tilde{L}$ and $\tilde{L}^M$ and the set of matrices $\tilde{A}^M(y)$ and $\tilde{A}(k, y)$ for $k = 0, \dots, T$ and $y = 1, \dots, M$. Then, the set of strategies is called consistent if:*

(i) There exists a set of matrices $K^M(y)$, $\Lambda^M(y)$ for $y = 1, \dots, M$ that satisfies (3.12), (3.13). *Moreover, it holds:*

$$\tilde{L}^M(y) = L^{M,\star}(y), \tag{3.23}$$

for any $y = 1, \dots, M$.

(ii) The Finiteness Criterion 1 is satisfyied.

(iii) It holds:

$$\tilde{L}(y) = [\hat{L}_0^\star(y) \dots \hat{L}_{T-1}^\star(y)], \tag{3.24}$$

for any $y = 1, \dots, M$. □

It is not difficult to show that a consistent set of strategies constitutes a perfect Nash equilibrium.

Proposition 4. *Consider a consistent set of strategies in the general form* (3.6) *and* (3.7)*. Then it constitutes a Perfect Nash equilibrium.*

Proof: The strategies of the minor players are optimal. Using equation (2.8), with $\sqrt{a}A_{cl}^M$ in the place of the A matrix, we conclude that equation (3.17) implies the finiteness of the total cost of the Major player and thus the optimality of the control law (3.14). Thus, the strategy of each player is optimal and the consistent set of strategies constitutes a Nash equilibrium. Furthermore, the strategies of the players satisfy the Dynamic Programming and thus the Nash equilibrium is perfect. □

Sufficient conditions for the existence of a Nash equilibrium are given in Section 3.6, for a special case. They are expressed as sufficient conditions for the convergence of an algorithm approximating the Nash equilibrium.

Remark 13. *The optimality conditions given by the equations* (3.12)-(3.15) *and* (3.18)-(3.22) *are Riccati type equations with two kinds of coupling. The first is due to the involvement of the gains in the $\tilde{A}$ matrices and has the same nature as the coupled Riccati equations of the LQ games [SH69b], [PMC79]. The second kind of coupling is through the Λ matrices and it has the same nature as the interconnected Riccati equations in the study of LQ control of MJLS [AKFJ95].* □

3.5 Large Number of Players Case

In this section we use a Mean Field approximation in order to study games with a very large number of players. This approach assumes a continuum of players. A set of optimal control problems that correspond to the limit of those in section 3.3, as the scale variable tends to infinity, is then stated. The Markov chain with a large number of states is approximated by a Markov chain with a continuum of states and thus a notion of convergence of Markov chains is first recalled in the subsection 3.5.1. Then the solution of the approximate optimal control problems for the major and minor players is characterized by appropriate Riccati integral equations and consistency conditions analogous of those of Section 3.4 are stated. Finally, it is proved that a set of feedback strategies satisfying those consistency conditions constitutes an ε - Nash equilibrium, for a game with a very large number of players.

Another motivation for the use of the continuous approximation is computational. The state space of the Markov chain that describes the random entrance grows fast as the maximum number of players increases. For example if the minor players have a time horizon 5 and the new minor players in each step belong to the set 1,2,...,N then the state space of the Markov chain describing the entrance has N^5 points. Thus, the equations characterizing a Nash equilibrium depend on many parameters and therfore, they are very complicated. On the other hand, in several cases the situation is much simplified using the continuous approximation.

3.5.1 Convergence of Markov Chains

In this section the vector representation of y_k will be used. The state space of the Markov chain is contained in the set:

$$D = \left\{ (y^0, \dots, y^{T-1}) \in \mathbf{R}^T : \sum_{i=0}^{T-1} y^i \leq 1, y^i \geq 0, \right\}. \tag{3.25}$$

The continuous approximation will be defined on the set D.

A Markov chain could be described using the notion of the stochastic kernel.

Definition 4. *Let $D' = \{d_1, \dots, d_M\} \subset D$ and $P = [p_{ij}]$ be an $M \times M$ stochastic matrix. The stochastic kernel that corresponds to the Markov chain with state space D' and transition matrix P is defined as:*

$$\bar{K}(y, B) = Pr(y_{k+1} \in B | y_k = y) = \sum_{j: d_j \in B} p_{ij}, \tag{3.26}$$

where $i = \min\{\arg\min_l\{\|y - d\| : d \in D'\}\}$, $y \in D$ and B a Borel subset of D. □

Let us then recall a notion of convergence of stochastic kernels from [Kar75] and a notion of continuity from [MT93].

Definition 5. *(i) We shall say that a sequence of stochastic kernels $\bar{K}_\nu$ converges weakly to a stochastic kernel $\bar{K}$ if for any sequence y_ν of elements of D converging to an element y of D and any (bounded) continuous function g, it holds:*

$$\int_D g(y')\bar{K}_\nu(y_\nu, y') \to \int_D g(y')\bar{K}(y, y') \tag{3.27}$$

(ii) A stochastic kernel $\bar{K}$ is called Feller continuous if $\bar{K}(y_\nu, \cdot) \Rightarrow \bar{K}(y, \cdot)$ when $y_k \to y$. □

Let us turn back to games described by the relationships (3.2) - (3.5) and a large number of minor players. To do so, we consider a sequence g^ν of games with increasing number of minor players; i.e. for the scale variable we assume $s_c \to \infty$ as $\nu \to \infty$. The state of the Markov chain describing the entrance is denoted by y_k^ν, the number of states of the Markov chain by M^ν and the corresponding stochastic kernel is denoted by $\bar{K}^\nu$.

Conclusions about the final part of this sequence of games are obtained under the assumption that the sequence of stochastic kernels $\bar{K}^\nu$ converges weakly to a to a Feller continuous stochastic kernel $\bar{K}$. The stochastic kernel $\bar{K}$, hence, approximates the final part of the sequence of Markov chains. We finally assume that the matrix functions $Q^M(\cdot), Q_f(\cdot), Q(\cdot)$ are continuous.

The following example shows that the continuum approximation often simplifies a lot the description of the random entrance.

Example 8. *Consider games involving only one type of minor players of time horizon 2. At each time step each one of ν players enters the game with probability p. Thus, the number of new minor players at each step follows a binomial distribution. The entrance dynamics is thus described by the Markov chain $y_k^\nu = [N_k^{0,\nu} \ N_k^{1,\nu}]/s_c^\nu$, $s_c = 2\nu$ and*

$$Pr(N_{k+1}^{\nu,0} = i) = \binom{\nu}{i} p^i (1-p)^{\nu - i}.$$

The random variable $N_k^{\nu,0}/s_c$ converges weakly to the deterministic constant $p/2$. Thus, the Markov chain with large ν may be approximated by a Markov chain with continuous state space and a stochastic kernel given by $\bar{K}((y_1, y_2), \cdot) = \delta(p/2, y_1)$, where δ denotes the Dirac measure. Thus, for a large number of players, the approximate description of the Markov chain is much simpler than the original. □

3.5.2 Approximate Optimal Control Problems

The approximate optimal control problems are then stated. These problems correspond to the limits of the OC Problems 1 and 2 of section 3.3 as the scale variable tends to infinity.

The reduced order dynamics for the major player, given by (3.10), remains unchanged under the limiting procedure. Thus, the limit optimal control problem for the major player is stated as follows:

OC Problem 3: "Minimize the cost function (3.4) subject to the dynamics (3.10) and $y_{k+1} \sim \bar{K}(y_k, \cdot)$". $\square$

The solution of the optimal control Problem 3 depends on the solution of a Riccati integral equation given by:

$$K^M(y) = Q^M + (\tilde{A}^M)^T \Big[a\Lambda^M - a\Lambda^M \tilde{B}^M \cdot \cdot \left(R^M/a + (\tilde{B}^M)^T \Lambda^M \tilde{B}^M\right)^{-1} (\tilde{B}^M)^T \Lambda^M \Big] \tilde{A}^M, \tag{3.28}$$

$$\Lambda^M(y) = E[K^M(y_{k+1})|y_k = y] = \int_D K^M(y')\bar{K}(y, dy'). \tag{3.29}$$

Consider the matrix functions $K(\cdot), \Lambda(\cdot)$ satisfying the Riccati integral equation (3.28), (3.29). Then the control law given by:

$$u^M(k) = \tilde{L}^{M,\star}(y_k)[(x^M(k))^T \;\; \tilde{z}^T(k)]^T, \tag{3.30}$$

where:

$$\tilde{L}^{M,\star}(y) = -\left((\tilde{B}^M)^T \Lambda \tilde{B}^M + R^M/a\right)^{-1} \Lambda \tilde{A}^M, \tag{3.31}$$

solves the optimal control Problem 3, under the following finiteness criterion:

Finiteness Criterion 2: Consider the closed loop matrix given by:

$$\tilde{A}^M_{cl} = \tilde{A}^M - \tilde{B}^M \tilde{L}^{M,\star}.$$

There exists a strictly positive definite matrix function $S(\cdot)$ satisfying:

$$\int_{y' \in D} a(\tilde{A}^M_{cl}(y))^T S(y') \tilde{A}^M_{cl}(y) \bar{K}(y, dy') - S(y) = -I, \tag{3.32}$$

for any $y \in D$. $\square$

The reduced order dynamics for a minor player is simplified under the limiting procedure. Specifically, consider a minor player i_0 with entrance time t_{i_0}. The limit dynamics is given by:

$$\tilde{x}^{i_0}(k+1) = \tilde{A}(k - t_{i_0}, y_k)\tilde{x}^{i_0}(k) + + \tilde{B}(k - t_{i_0}, y_k)u^{i_0} + W^{i_0}(k). \tag{3.33}$$

The entries of the matrix are computed by simple but lengthy calculations. The entries of the first row are given by:

$$\begin{aligned} \tilde{A}_{x^M x^M} &= A^M + B^M L^{MM}(y_k), \\ \tilde{A}_{x^M z^l} &= F + B^M L^M(l, y_k). \end{aligned}$$

The entries of the second row are zero. For the rest of the rows except the last one, the entries are given by:

$$\begin{aligned} &\tilde{A}_{z^{l+1} x^M} = y_k^l(G + BL^{1M}(l, y_k)), \\ &\tilde{A}_{z^{l+1} z^{l'}} = \delta_{l,l'}\left(A + y_k^l B\bar{L}(l, y_k)\right) + (F + BL(l', l, y_k))y_k^l, \\ &\tilde{A}_{z^{l+1} x^{i_0}} = 0. \end{aligned}$$

The entries of the last row remain the same, i.e. $\tilde{A}_{x^{i_0}x^M} = G$, $\tilde{A}_{x^{i_0}z^l} = F$ and $\tilde{A}_{x^{i_0}x^{i_0}} = A$.

The limit optimal control problem for a minor player i_0 is, thus, the following:

OC Problem 4: "Minimize the cost function (3.5) subject to the dynamics (3.33) and $y_{k+1} \sim \bar{K}(y_k, \cdot)$". □

The solution of the limit optimal control Problem 4 depends on the following Riccati type difference integral equaitons.

$$K_T(y) = Q_f(y), \tag{3.34}$$

$$\begin{aligned}\Lambda_{k+1}(y) &= E[K_{k+1}(y_{k+1+t_{i_0}})|y_{k+t_{i_0}}] \\ &= \int_D K_{k+1}(y')\bar{K}(y, dy'),\end{aligned} \tag{3.35}$$

$$\begin{aligned}K_k(y) = Q + (\tilde{A}(k))^T[\Lambda_{k+1} - \Lambda_{k+1}\tilde{B}\cdot \\ \cdot (R + \tilde{B}^T\Lambda_{k+1}\tilde{B})^{-1}\tilde{B}^T\Lambda_{k+1}]\tilde{A}.\end{aligned} \tag{3.36}$$

The optimal control law is then given by:

$$u^{i_0}(k + t_{i_0}) = \hat{L}^\star_k \tilde{x}^{i_0}(k + t_{i_0}), \tag{3.37}$$

where:

$$\hat{L}^\star_k(y) = -(R + \tilde{B}^T\Lambda_{k+1}\tilde{B})^{-1}\Lambda_{k+1}\tilde{A} \tag{3.38}$$

3.5.3 Consistency Conditions and ε - Nash Equilibrium

The consistency conditions for the solutions of the OC Problems 3 and 4 are stated and then used to characterize approximate Nash equilibrium.

Definition 6. *Consider a set of strategies that belong in the general form* (3.6), (3.7) *with gains* $\tilde{L}$ *and* $\tilde{L}^M$. *Assume that they depend continuously on* $y \in D$. *Compute the matrix functions* $\tilde{A}^M(y)$ *and* $\tilde{A}(y)$. *The set of strategies will be called consistent if:*

(i) There exist continuous matrix functions $K^M(y), \Lambda^M(y)$ *satisfying* (3.28), (3.29). *Moreover it holds:*

$$\tilde{L}^M = \tilde{L}^{M,\star}, \tag{3.39}$$

where $\tilde{L}^{M,\star}(y)$ *is given by* (3.31).

(ii) The closed loop matrix $\tilde{A}^M_{cl}(y)$ *satisfies the Finiteness Criterion 2.*

(iii) The matrix functions $\hat{L}^\star_k(y)$ *computed by* (3.34) - (3.38) *satisfy:*

$$\tilde{L}(y) = [\hat{L}^\star_0(y), \ldots, \hat{L}^\star_{T-1}(y)], \tag{3.40}$$

for any $y \in D$. □

Let us consider a game with a large number of players g^ν, where the participants use a set of approximately consistent strategies, characterized by gains $\tilde{L}$ and $\tilde{L}^M$. Under certain conditions, this set of strategies is shown to constitute an ε - Nash equilibrium, i.e. the cost of any player is at most ε - far from the optimal cost. This property is illustrated by the following Theorem 3 and its Corollary 2. Before stating the Theorem 3, let us introduce some notation:

Notation: For the game g^ν, let us denote by $J^{M,\nu}(\pi_{\tilde{L},\tilde{L}^M})$ and $J^{i,\nu}(\pi_{\tilde{L},\tilde{L}^M})$ be the values of the cost functions, (3.4) and (3.5), when all the players use the policies given by (3.6) and (3.7) with gains $\tilde{L}, \tilde{L}^M$.

Let also $(J^{M,\nu}(\pi_{\tilde{L},-M}))^\star$ be the minimum value of the cost function of the major player, assuming that the other players use the policies given by (3.6) with gains $\tilde{L}$. Finally, denote by $(J^{i,\nu}(\pi_{\tilde{L},\tilde{L}^M,-i}))^\star$ the minimum value of the cost function of the player i, assuming that the other players use the strategies given by (3.6) and (3.7) with gains $\tilde{L}, \tilde{L}^M$. □

Theorem 3. *Consider an approximately consistent set of strategies given by* (3.6) *and* (3.7) *with gains $\tilde{L}, \tilde{L}^M$. Then for any positive constant ε, there exists a positive integer ν_0 such that:*

$$J^{M,\nu}(\pi_{\tilde{L},\tilde{L}^M}) \leq (J^{M,\nu}(\pi_{\tilde{L},-M}))^\star + \varepsilon \tag{3.41}$$

$$J^{i,\nu}(\pi_{\tilde{L},\tilde{L}^M}) \leq (J^{i,\nu}(\pi_{\tilde{L},\tilde{L}^M}, -i))^\star + + \varepsilon(1 + E[(\tilde{x}^i(t_i))^2]), \tag{3.42}$$

for all the minor players $i \in \Delta$ and any $\nu \geq \nu_0$.

Structure of the Proof: The proof of the second inequality is based on the fact that the optimal policies for a minor player involve continuous functions of the state vector and the Markov chain and some properties of the weak convergence.

A basic step in the proof of the first inequality is given in Section 3.8.1, where it is shown that some stability properties of the MJLS are preserved under weak convergence. It is then shown that the final part of the series involved in the cost function is small in some sense, uniformly in the initial conditions, and thus it suffices to compare finite series. The result for finite series is similar to the proof of the second inequality. The detailed proof is relegated to Section 3.8. More general results are first shown in section 3.8.2 and particularly in Propositions 7 and 8. Theorem 1 is then proved as a consequence of the Propositions of section 3.8.2. □

Corollary 2. *Consider a set of strategies in the form* (3.6), (3.7). *In addition to the assumptions of Theorem 3, assume that the closed loop system is mean square exponentially stable, i.e. the following Lyapunov equation,*

$$\int_{y' \in D} (\tilde{A}^M_{cl}(y))^T S(y') \tilde{A}^M_{cl}(y) \bar{K}(y, dy') - S(y) = -I,$$

admits a strictly positive definite solution $S(\cdot)$. Then for any positive constant ε, there exists a positive integer ν_0 such that (3.6) *and* (3.7) *constitute an ε - Nash equilibrium for any $\nu \geq \nu_0$,* *i.e. it holds:*

$$J^{M,\nu}(\pi_{\tilde{L},\tilde{L}^M}) \leq (J^{M,\nu}(\pi_{\tilde{L},-M}))^\star + \varepsilon \tag{3.43}$$

$$J^{i,\nu}(\pi_{\tilde{L},\tilde{L}^M}) \leq (J^{i,\nu}(\pi_{\tilde{L},\tilde{L}^M,-i}))^\star + \varepsilon. \tag{3.44}$$

Proof: The inequality (3.44) is a consequence of the inequality (3.42) and the mean square stability. □

Remark 14. *The approximate consistency conditions (Def. 6) involve nonlinear matrix integral equations and in general are not simpler than the consistency conditions of Section 3.4. However, in several cases the situation is extremely simplified as illustrated in the Example 3 of the next section.* □

3.6 Computing the Nash Equilibria

An algorithm for solving the consistency conditions derived in Section 3.4, is described in subsection A. Conditions under which the algorithm converges are stated and thus sufficient conditions for the existence of a Nash equilibrium are then derived. In the third subsection, some numerical examples are given.

3.6.1 Algorithm

The algorithm initially guesses a value for the feedback gains. With the assumed values, it computes the matrices for the optimal control problems. Then, the optimal control problems are solved and new feedback gains are computed. The new feedback gains are then used to compute the system matrices and solve the optimal control problems and so forth. This algorithm is presented in Algorithm 2 table.

Algorithm 2

1: Guess $\tilde{L}$, $\tilde{L}^M$.
2: Compute the matrices $\tilde{A}(\cdot,\cdot)$ using (3.11).
3: Compute the new values for the gains $\hat{L}_1^\star, \dots, \hat{L}_{T-1}^\star$, using equations (3.18) - (3.22).
4: Set $\tilde{L} = [\hat{L}_1^\star, \dots, \hat{L}_{T-1}^\star]$
5: Compute the matrix $\tilde{A}^M$ using (3.10).
6: Set $\tilde{L}_{old}^M = \tilde{L}^M$.
7: Compute matrices $K^M(\cdot), \Lambda^M(\cdot)$ to satisfy (3.12), (3.13).
8: Use (3.15) to update the values of $\tilde{L}^M$, i.e. set $\tilde{L}^M = \tilde{L}^{M,\star}$.
9: If the difference $\max_y \|\tilde{L}^M - \tilde{L}_{old}^M\|$ is small enough then halt. Else go to Step 2.

An analogous algorithm can be used for the Mean field case, as well.

Remark 15. *Step 7 of the algorithm may be implemented in several ways. Probably the simpler one is to use the value iteration algorithm. A variant of the algorithm would involve the use of a single step of the value iteration method, instead of the steps 7 and 8 of the Algorithm 2.* □

3.6.2 Convergence of the Algorithm

The convergence of the Algorithm 2 depends on the existence of a Nash solution, as well as on some stability properties. Such problems are hard to solve and remain open even in simpler settings (ex. [FJAK96]).

In this subsection, we study the convergence of the Mean Field variant of Algorithm 2. Particularly, sufficient conditions for convergence of the algorithm and hence, the existence of a Mean Field Nash solution are stated, based on contraction mapping ideas. A special class of games is analyzed. Specifically, it is assumed that there are only minor players coupled only though costs having state vectors of dimension one. The time horizon of the players is three.

In what follows, for vectors $\|\cdot\|$ denotes the 1-norm, for a matrix A, $\|A\|$ denotes the induced 1-norm and for a matrix function $A(\cdot)$, $\|A\|$ denotes the essential supremum of the induced 1-norm. Assuming that $A = \alpha$, $B = 1$ and $r = 1$, the matrices $\tilde{A}$ and $\tilde{B}$ take the form:

$$\tilde{A} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ \alpha + (L(0,0) + \bar{L}(0))y_k^0 & L(1,0)y_k^0 & L(2,0)y_k^0 & 0 \\ L(0,1)y_k^1 & a + (L(1,1) + \bar{L}(1))y_k^1 & L(2,1)y_k^1 & 0 \\ 0 & 0 & 0 & \alpha \end{bmatrix},$$

$\tilde{B} = [0\ 0\ 0\ 1]^T$. Let us denote by:

$$\begin{aligned} \tilde{L}^{\tilde{A}} &= [\tilde{L}^{\tilde{A}}(0), \tilde{L}^{\tilde{A}}(1)] = \\ &= [[L(0,0) \quad L(1,0) \quad L(2,0) \quad \bar{L}(0)], \\ &\qquad [L(0,1) \quad L(1,1) \quad L(2,1) \quad \bar{L}(1)]], \end{aligned}$$

i.e. all the entries of $\tilde{L}$ that affect $\tilde{A}$.

We consider the following mappings:

$$\tilde{L}^{\tilde{A}} \xrightarrow{T_1} \left(\tilde{A}; \begin{bmatrix} K_1 \\ K_2 \end{bmatrix}\right) \xrightarrow{T_2} \left(\tilde{A}; \begin{bmatrix} \Lambda_1 \\ \Lambda_2 \end{bmatrix}\right) \xrightarrow{T_3} \left(\tilde{L}^{\tilde{A}}\right)^{\star}. \tag{3.45}$$

This mappings compute the best response of a player if the other players use strategies described by $L^{\tilde{A}}$. Sufficient conditions for the contractivity of the mapping $T = T_1 \circ T_2 \circ T_3$ are then derived.

The mappings T_1 and T_2 are computed according to the following equations:

$$\begin{aligned} K_3(y) &= Q_f(y), \\ K_2(y) &= Q + \tilde{A}^T[\Lambda_3 - \Lambda_3 \tilde{B}\tilde{B}^T \Lambda_3/(1 + \Lambda_3^{4,4})]\tilde{A}, \\ K_1(y) &= Q + \tilde{A}^T[\Lambda_2 - \Lambda_2 \tilde{B}\tilde{B}^T \Lambda_2/(1 + \Lambda_2^{4,4})]\tilde{A}, \\ \Lambda_i(y) &= \int_{y' \in D} K_i(y')\bar{K}(y, dy') \end{aligned}$$

The mapping T_3 is given by:

$$(\tilde{L}^{\tilde{A}}(i))^{\star} = -\frac{1}{1 + \Lambda_{i+1}^{4,4}} \Lambda_{i+1}\tilde{A}, \;\; i = 0, 1 \tag{3.46}$$

Lemma 2. *It holds:*

$$\left\| \int_{y' \in D} (K(y') - K'(y')) \bar{K}(y, dy') \right\| \leq \|K(y) - K'(y)\|$$

Proof: Immediate □

The mapping T_2 is, thus, non-expansive (weakly contractive). Sufficient conditions for the contractivity of T are found using the following technical result.

Lemma 3. *If* $\|A\|, \|A'\| < d_1$, $\|K\|, \|K'\| < d_2$ *and* $\|A - A'\| < c_1$, $\|K - K'\| < c_2$*, then it holds:*

$$\|f(A, K) - f(A', K')\| < 2(d_1 d_2 + d_1 d_2^2)c_1 + d_1^2(1 + d_2^2 + 2d_2 + c_2)c_2,$$

where: $f(A, K) = Q + A^T[K - K\tilde{B}\tilde{B}^T K/(1 + K^{4,4})]A$.

Proof: The proof is long but straightforward. It uses repeatedly the matrix identity $XY - X'Y' = (X - X')Y + X'(Y - Y')$ and the sub-multiplicative property of the matrix norm. □

For a constant $0 < \rho < 1$, a region, $\bar{R}_\rho$, containing 0 that the mapping T is ρ-contractive, will be determined. Assuming that the algorithm starts with zero gains, $\tilde{L}^{\tilde{A}} = 0$, after the application of T we have:

$$\begin{aligned} &\|(\tilde{L}^{\tilde{A}})^\star\| \leq \beta = \\ &= 2q\alpha + \alpha^3(q + (1 + \alpha^2)(q_f + q_f^2) + (q + \alpha^2(q_f + q_f^2))^2), \end{aligned}$$

where $q = \|Q\|$ and $q_f = \|Q_f\|$. This inequality can be derived using the sub-multiplicative property of the matrix norm. If after a number of iterations of T, $\tilde{L}^{\tilde{A}}$ remains in $\bar{R}_\rho$, then it holds:

$$\begin{aligned} &\|\tilde{L}^{\tilde{A}}\| < \beta/(1 - \rho), \\ &\|\tilde{A}\| < d_1 = a + \beta/(1 - \rho) \ \text{ and} \\ &\left\| \begin{bmatrix} K_1 \\ K_2 \end{bmatrix} \right\| < d_2, \end{aligned}$$

where:

$$d_2 = 2q + d_1^2(q + (1 + d_1^2)(q_f + q_f^2) + (q + d_1^2(q_f + q_f^2))^2).$$

The last inequality is also shown using the sub-multiplicative property of the matrix norm.

Proposition 5. *Assume that the parameters* α, q, q_f *are such that:*

$$\begin{aligned} 4(d_1 d_2 + d_1 d_2^2) + 2d_1^2[1 + d_2^2 + 2d_2 + \\ + 2(d_1 d_2 + d_1 d_2^2)\beta](d_1 d_2 + d_1 d_2^2) < \rho_1. \end{aligned} \tag{3.47}$$

Furthermore, assume that $\rho_2 = d_1 + d_2 + 2d_1 d_2$ *is such that:*

$$\rho = (1 + \rho_1)\rho_2 < 1. \tag{3.48}$$

Then T *is* ρ*-contractive, the Algorithm 2 converges and there exists a MF Nash solution for the game.*

Proof: We first determine a Lipschitz constant for T_1. Assuming that $\|\tilde{L}^{\tilde{A}} - \tilde{L}'^{\tilde{A}}\| < c$, Lemma 3 implies:

$$\begin{aligned}
&\|\Lambda_2 - \Lambda_2'\| \leq \|K_2 - K_2'\| < 2(d_1 d_2 + d_1 d_2^2)c, \\
&\|K_1 - K_1'\| < 2(d_1 d_2 + d_1 d_2^2)c + \\
&\quad + 2d_1^2(1 + d_2^2 + 2d_2 + 2(d_1 d_2 + d_1^2 d_2)c)(d_1 d_2 + d_1 d_2^2)c.
\end{aligned}$$

Hence, inequality (3.47) implies that:

$$\|K_1 - K_1'\| + \|K_2 - K_2'\| < \rho_1 c.$$

Thus, T_1 has Lipschitz constant less than $1 + \rho_1$. Using (3.46) and the sub-multiplicative property, it is straightforward to show that T_3 is ρ_2 - contractive. These, in combination with the non-expansive property of T_2, complete the proof. □

Remark 16. *The generalization to the many steps case does not add any further difficulties. The existence of a dynamic coupling except the cost coupling makes the matrix A, time varying. The generalization to the multi-dimensional case is also immediate. However, only this special case is analyzed in order to keep the results as simple as possible.*

Example 9. *In this example major and minor players have scalar state equations. The minor players have time horizon 2 and the maximum possible number of minor players participating in the game at some time step is 4. The number of new minor players that enter the game at each instant of time is either 1 or 2. Thus, the entrance dynamics is described by a Markov chain with state space: $(\frac{1}{4}, \frac{1}{4})$, $(\frac{1}{4}, \frac{2}{4})$, $(\frac{2}{4}, \frac{1}{4})$ and $(\frac{2}{4}, \frac{2}{4})$.For the states of the Markov chain, we shall use the enumeration $1, 2, 3, 4$ respectively. The entrance dynamics is described by the following transition matrix:*

$$\Pi = \begin{bmatrix} 0.9 & 0 & 0.1 & 0 \\ 0.2 & 0 & 0.8 & 0 \\ 0 & 0.3 & 0 & 0.7 \\ 0 & 0.8 & 0 & 0.2 \end{bmatrix}.$$

The dynamic equation of the major player is given by:

$$x^M(k+1) = x^M(k) + \frac{c_1}{4} \sum_{i \in I_k} x^i(k) + u^M(k) + w^M(k),$$

and for a minor player by:

$$x^i(k+1) = x^i(k) + \frac{c_1}{4} \sum_{j \in I_k} x^j(k) + c_1 x^M(k) u^i(k) + w^i(k),$$

where by c_1, we denote all the coupling coefficients of the dynamic equations. Thus, the parameters of the state equations are given by $A^M = B^M = A = B = 1$ and $F^M = G = F = c_1$.

Table 3.1: Gain matrices for the major player, $c_1 = c_2 = 1$

y_k	$L^{MM}(y_k)$	$L^M(0, y_k)$	$L^M(0, y_k)$
1	-0.6411	-0.6618	-0.6326
2	-0.6560	-0.6787	-0.6467
3	-0.7338	-0.7211	-0.7240
4	-0.6825	-0.6682	-0.6715

The cost function matrices Q for the major player are given by $Q^M(1) = Q^M(2) = Q^M(3) = I_3$ *and* $Q^M(4) = (1 + c_2)I_3$*. The cost matrices* Q^1 *for the minor players are all units, i.e.* $Q(y) = Q_f(y) = I_4$ *and* $R = R^M = 1$*.*

For example, if $c_1 = c_2 = 1$*, after* 30 *steps of the Algorithm 2, the gain matrices change less than* 10^{-15}*. The feedback gains for the major player are gain in Table I.*

In what follows, we study the dependence of the feedback gains on the coupling parameters c_1 *and* c_2*. Let us first consider the case where* $c_2 = 0$ *and* $c_1 \neq 0$*. Then, each player will interact dynamically with an unknown number of minor players. The distribution of the number of the minor players in the next step depends on* y_k*. Thus, the dependence of the feedback gains on* y *is larger, for larger values of* $|c_1|$*. The dependence of* $L^{MM}(y)$ *for* $y = 1, \ldots, 4$ *on* c_1 *is illustrated in Fig. 1.*

We next assume no dynamic coupling, i.e. $c_1 = 0$ *and* $c_2 \neq 0$*. The dependence of* $L^{MM}(y)$ *for* $y = 1, \ldots, 4$ *on* c_2 *is illustrated in Fig. 2. Again, the dependence on* y_k *is larger, for larger values of* $|c_2|$*.*

The Algorithm 2 does not always converge. For example, for $c_1 = c_2 = 10$*, the Algorithm 2 does not converge.*

Finally we fix a dynamics and cost function for the participants of the game and present the sample paths of a run. The dynamics of the players are as before, with $c_1 = 1$ *and the cost functions slightly different. Particularly,* $Q^M(y) = I_3$*,* $R^M = 1$*,* $R = 0.75$ *and:*

$$Q(y) = Q_f(y) = \begin{bmatrix} 10 & 0 & 0 & -10 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -10 & 0 & 0 & 10 \end{bmatrix},$$

The sample paths are shown in Figure 3.3. □

The following example studies a very simple game with a large number of minor players. It illustrates that, in certain cases, the mean field approximation simplifies very much the analysis of the game. For simplicity reasons, we assume that there is no major player.

Example 10. *There is only one type of minor players with time horizon two. At each time step, each one of* ν *minor players tosses a fair coin and with probability* $1/2$ *enters the game.*

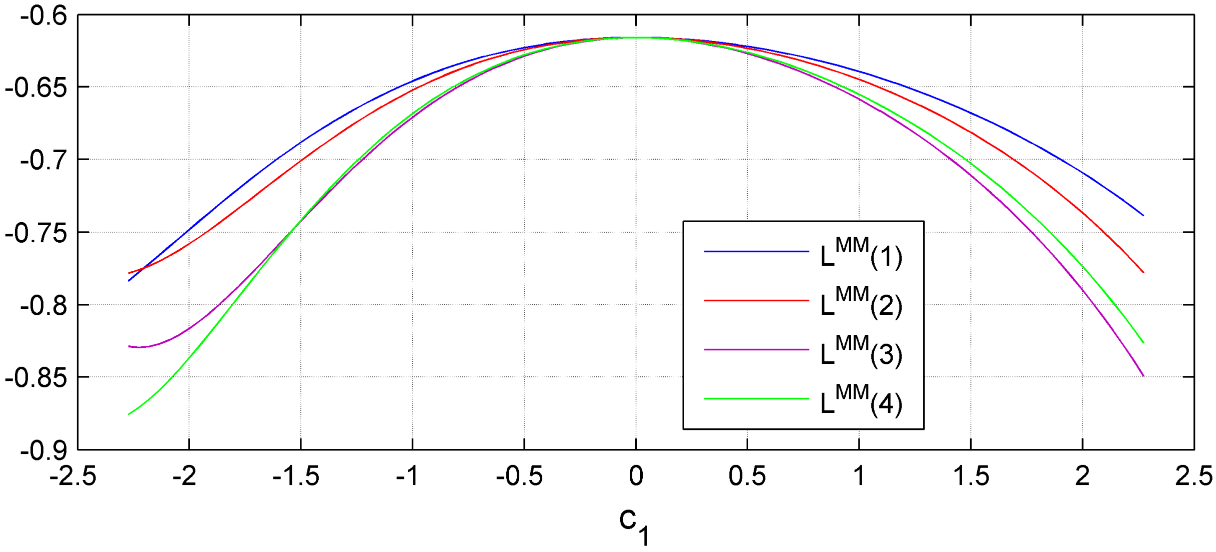


Figure 3.1: The dependence of the feedback gains $L^{MM}(y)$ on c_1.

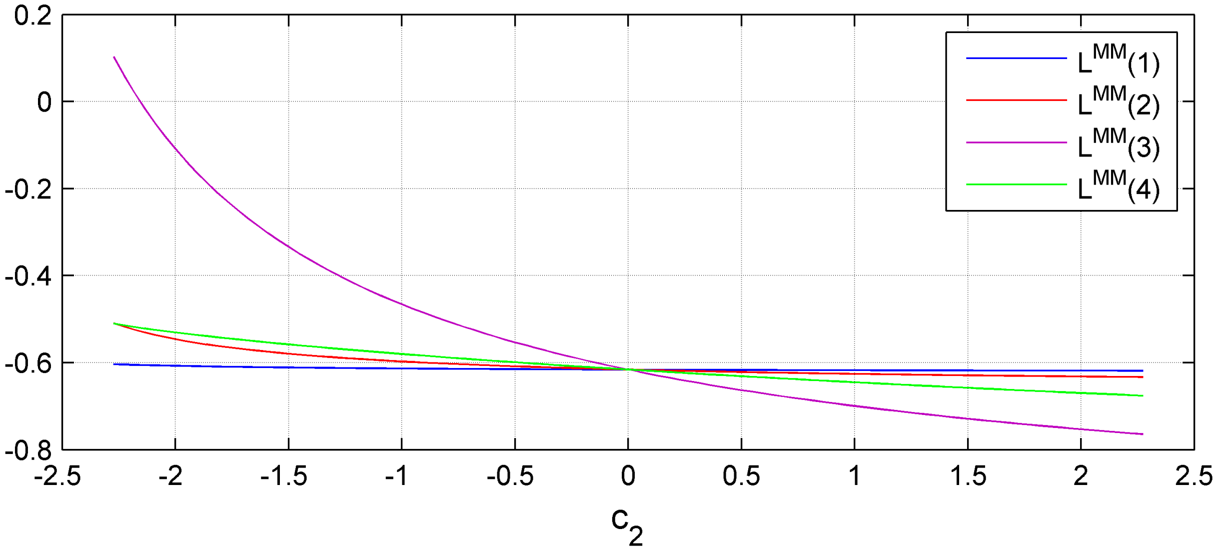


Figure 3.2: The dependence of the feedback gains $L^{MM}(y)$ on c_2.

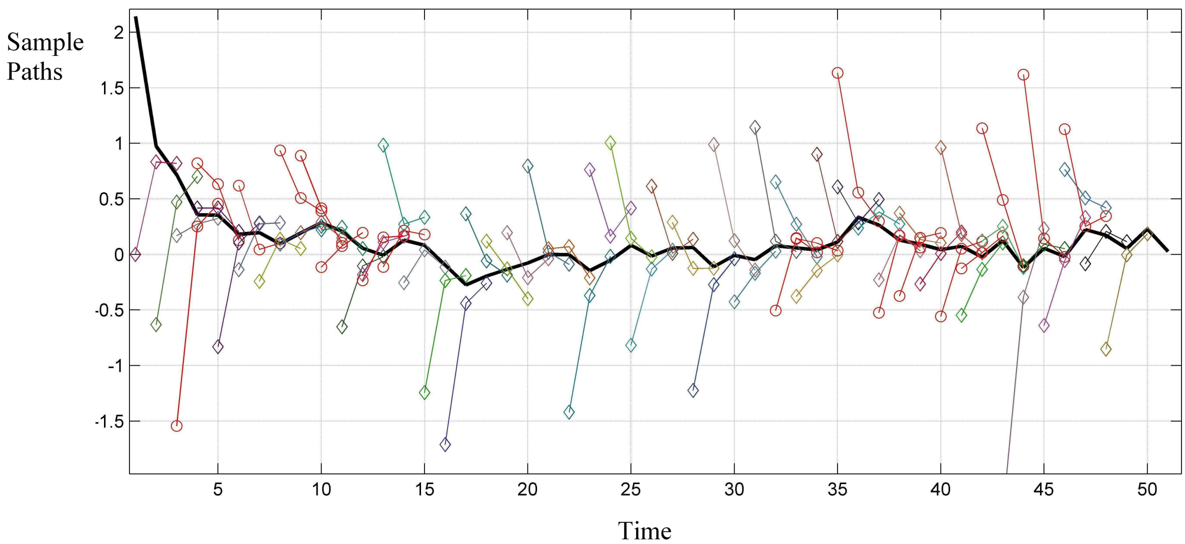


Figure 3.3: The sample paths of the major and minor players. The black line corresponds to the major player. The red lines correspond to cases where there is a second player entering at some instant of time.

Table 3.2: Gain matrices computation

y	$L_1(y)$	$L_2(y)$	$L_3(y)$	$L_4(y)$	$L_5(y)$	$L_6(y)$
$\tilde{y}$	-0.813	-0.68	-0.5	-0.5	-0.667	-0.5
y'	-0.813	-0.68	-0.444	-0.444	-0.667	-0.444
y	-0.754	-0.66	-0.444	-0.444	-0.626	-0.444

The dynamic equation of the minor players is given by:

$$x^i(k+1) = x^i(k) + \sum_{j\in I_k} x^j(k)/s_c + u^i(k) + w^i(k).$$

The cost function matrices are given by $Q_f = 4y^1 I_3$, $Q = 12yy^T I_3$ *and* $R = 1$ *(where the notation* Q_f, Q, R, y^1, y^0 *is used instead of* $Q_f^1, Q^1, R^1, y^{1,1}, y^{1,0}$*).*

The scale variable has a value $s_c = 2\nu$ *and the approximate description of the Markov chain when* $\nu \to \infty$ *is given by the stochastic kernel:*

$$\bar{K}((y_1, y_2), \cdot) = \delta_{(1/4, y_1)}(\cdot),$$

where δ *is the Dirac measure.*

Due to the absence of a major player the approximate consistency conditions involve only (3.33) - (3.36). *The unknown quantities can be expressed in terms of the functions* $L_1(y) = L(0,0,y), L_2(y) = L(1,0,y)$, $L_3(y) = L(0,1,y), L_4(y) = L(1,1,y), L_5(y) = \bar{L}(0,y)$ *and* $L_6(y) = \bar{L}(1,y)$.

Due to the special form of the stochastic kernel, the integral equation (3.35) *becomes:*

$$\Lambda_{k+1}((y_1, y_2)) = K_{k+1}(\bar{y}, y_1),$$

where $\bar{y} = 1/4$*. Hence, the form of the stochastic kernel implies a decoupling in the consistency conditions. Particularly, for* $\tilde{y} = (\bar{y}, \bar{y})$*, the consistency conditions do not depend on the other values of y. Writing the consistency conditions for some* $y' = (\bar{y}, y_1)$*, the equations depend only on* $L_1(y'), \ldots, L_6(y')$ *and* $L_1(\tilde{y}), \ldots, L_6(\tilde{y})$*. Furthermore, for some* $y = (y_1, y_2)$*, the consistency conditions depend only on* $L_1(y), \ldots, L_6(y)$ *and* $L_1(y'), \ldots, L_6(y')$*.*

This structure of the consistency conditions suggests the following procedure: Compute the values of $L_1, \ldots L_6$ *on* $\tilde{y}$*, solving six equations with six unknowns. Then, for each* $y' = (\bar{y}, y_1)$*, compute the values of* $L_1, \ldots L_6$ *on* y'*. Finally, for any* $y = (y_1, y_2)$*, compute the values of* $L_1, \ldots L_6$ *on* y*, again, by solving six equations with six unknowns involving* $L_1, \ldots L_6$ *on* y *and* y'*.*

As an example, we compute the values of the feedback gains on $y = (0.2, 0.6)$*. The computations are shown in Table II.*

It remains to show the stability of the limit system. For $k \geq 3$*, it holds* $y_k = \tilde{y}$ *a.s. Furthermore, it holds:*

$$A_{cl}^M(\tilde{y}) = \begin{bmatrix} 0 & 0 \\ 0.88 & 0.08 \end{bmatrix}.$$

Thus, Corollary 2 applies, i.e. for any $\varepsilon > 0$ the strategies computed constitute an ε - Nash equilibrium for large ν. □

Remark 17. *Example 10 shows that when the random entrance is independent, the approximate consistency conditions are decoupled. The solutions to the approximate optimal control problems could, thus, be obtained using this special form.* □

3.7 Conclusion

Games with a major player and Randomly Entering minor players were considered. The problem of the characterization of Symmetric Linear Feedback Strategies that constitute a Nash equilibrium, was converted to a set of coupled finite and infinite horizon LQ control problems for MJLS. Appropriate coupled Riccati type equations were derived to characterize Nash equilibrium. The case where there exists a very large number of minor players was addressed using a Mean Field approach. Particularly, the evolution of the number of players is approximately described using a Markov chain having a continuum of states. Some limit optimal control problems were then stated. A set of Symmetric Linear Feedback Strategies that solves the limit optimal control problems was proved to constitute an ε - Nash equilibrium, when the scale is sufficiently large. A sufficient condition for the existence of a Mean Field Nash equilibrium was derived using contraction mapping ideas. Numerical examples were also presented. It occurs that, in several cases, the Mean Field approximation simplifies considerably the analysis.

3.8 Proofs

3.8.1 Weak Convergence and Mean Square Stability

Consider a sequence of systems:

$$x_{k+1}^{\nu} = A(y_k)x_k^{\nu}, \quad y_{k+1}^{\nu} \sim \bar{K}^{\nu}(y_k^{\nu}, \cdot), \tag{3.49}$$

and a limit system:

$$x_{k+1} = A(y_k)x_k, \quad y_{k+1} \sim \bar{K}(y_k, \cdot). \tag{3.50}$$

Assume that $\bar{K}^{\nu} \to \bar{K}$ weakly, $\bar{K}$ is Feller continuous and that $A(\cdot)$ is a continuous matrix function. Assume also that the limit system, given by (3.50), is exponentially mean square stable. The basic topic of this section is to show that the system given by (3.49) is exponentially mean square stable, for large ν.

For any $a \in (0, 1)$, there exists an integer k such that:

$$E[x_k^T x_k] < aE[x_0^T x_0], \tag{3.51}$$

for any x_0, y_0 initial conditions. Choosing x_0 to be any nonrandom initial condition, the last inequality becomes:

$$aI - E[A^T(y_0)\dots A^T(y_{k-1})A(y_{k-1})\dots A(y_0)] > 0. \tag{3.52}$$

The positive definiteness of this matrix is equivalent, due to the Sylvester criterion, to a set of inequalities in the form:

$$f_j\left(E\begin{bmatrix}\bar{f}_1(y_0,\dots,y_{k-1})\\ \vdots \\ f_{n^2}(y_0,\dots,y_{k-1})\end{bmatrix}\right) > 0, \tag{3.53}$$

for $j = 1,\dots,n$, where $\bar{f}_i$, $i = 1,\dots,n^2$ correspond to the elements of the matrix in (3.52) and are continuous and f_j are the multinomials derived using the Sylvester criterion.

The inverse procedure shows that the conditions in (3.53) imply the mean square exponential stability of the limit system (3.50).

Let us then state the basic result of this section.

Proposition 6. *Under the assumptions stated above, there exists a positive integer ν_0 such that $r(T_{A,\bar{K}^\nu}) < 1$, for any $\nu \geq \nu_0$.*

Before proving the proposition, a lemma will be stated. This lemma illustrates a uniformity property of the weak convergence. The uniformity is expressed in terms of the Bounded Lipschitz metric ([Pan08] section 17) which is defined by:

$$\beta(P_1,P_2) = \sup\left\{\left|\int f dP_1 - \int f dP_2\right| : \|f\|_{BL} \leq 1\right\},$$

where P_1, P_2 are probability measures on D and $\|f\|_{BL} = \sup_{y\in D}\{f(y)\} + \inf\{L : f$ is L - Lipschitz$\}$. The metric $\beta(\cdot,\cdot)$ metrizes the weak convergence, due to the fact that D is seperable ([Pan08] section 17).

In order to state the lemma, let us consider the functions:

$$\Xi^\nu, \Xi : (D, \|\cdot\|) \to (\Pi(D^k), \beta),$$

where $\Pi(D^k)$ is the space of probability measures on D^k and for any $C \in \mathcal{B}(D^k)$ the functions Ξ^ν have the form $\Xi^\nu(y)(C) = Pr((z_0,\dots,z_{k-1}) \in C)$, where z_0 has a distribution concentrated on y and $z_{i+1} \sim \bar{K}^\nu(z_i,\cdot)$. In the same way the values of Ξ are defined. Thus, Ξ maps the initial condition y_0 to the distribution of $(y_0,\dots,y_{k-1})$.

Due to the Feller continuity of $\bar{K}$, it is not difficult to show that the function Ξ is continuous ([Kar75]). The following lemma illustrates a uniformity property of the convergence of Ξ^ν to Ξ. Particularly, it is shown that for the same initial condition $y_0^\nu = y_0$ and large ν, the distribution of $(y_0^\nu,\dots,y_{k-1}^\nu)$ is β - close to $(y_0,\dots,y_{k-1})$, uniformly in y_0.

Lemma 4. *Under the assumptions stated above, for any $\varepsilon > 0$, there exists a positive integer ν_0 such that $\beta(\Xi^\nu(y), \Xi(y)) < \varepsilon$ for any $\nu \geq \nu_0$ and any $y \in D$.*

Proof: To contradict, assume that there is a positive constant ε such that for any $\nu_0 \in \mathbb{N}$, there exists a $\nu \geq \nu_0$ and a $y \in D$ with $\beta(\Xi^\nu(y), \Xi(y)) > \varepsilon$. Then there exist sequences m_ν, y_{m_ν} such that $m_\nu \geq \nu$, $m_\nu > m_{\nu-1}$ and $\beta(\Xi^{m_\nu}(y_{m_\nu}), \Xi(y_{m_\nu})) > \varepsilon$. The compactness of D implies the existence of a further subsequence $y_{m_{\nu_l}}$ that converges to a value $\bar{y}$. Theorem 1 of [Kar75] implies that $\beta(\Xi^{m_{\nu_l}}(y_{m_{\nu_l}}), \Xi(\bar{y})) \to 0$.

However, the triangle inequality implies:

$$\begin{aligned}\beta(\Xi^{m_{\nu_l}}(y_{m_{\nu_l}}), \Xi(\bar{y})) \geq{}& \beta(\Xi^{m_{\nu_l}}(y_{m_{\nu_l}}), \Xi(y_{m_{\nu_l}})) \\ &- \beta(\Xi(y_{m_{\nu_l}}), \Xi(\bar{y})) \\ >{}& \varepsilon - \beta(\Xi(y_{m_{\nu_l}}), \Xi(\bar{y})).\end{aligned}$$

The continuity of Ξ implies that $\beta(\Xi^{m_{\nu_l}}(y_{m_{\nu_l}}), \Xi(\bar{y})) > \varepsilon/2$, which contradicts $\beta(\Xi^{m_{\nu_l}}(y_{m_{\nu_l}}), \Xi(\bar{y})) \to 0$. $\square$

Remark 18. *If the functions Ξ^ν are continuous for large ν, the proof of Lemma 4 becomes trivial.* $\square$

Let us then turn back to the proof of Proposition 6.

Proof of Proposition 6: The quantities:

$$g_j(y_0) = f_j\left(E\begin{bmatrix}\bar{f}_1(y_0, \dots, y_{k-1}) \\ \vdots \\ \bar{f}_{n^2}(y_0, \dots, y_{k-1})\end{bmatrix}\right), \tag{3.54}$$

are continuous functions of y_0. Thus, due to the compactness of D, there is a constant $\varepsilon_1 > 0$, with $g_j(y) > \varepsilon_1$ for any $y \in D$ and any $j = 1, \dots, n$. The functions $f_j(\cdot)$ are uniformly continuous in D and thus there exists a positive constant δ_1 such that $f_j(v_1) > \varepsilon_1$ implies $f_j(v_2) > 0$, for any $v_2 \in D^k$ with $\|v_1 - v_2\| < \delta_1$ and any $j = 1, \dots, n$.

Choose $y_0 = y_0^\nu$. The entries of the functions f_j, i.e. $E[\bar{f}_i(y_0^\nu, \dots, y_{k-1}^\nu)]$ and $E[\bar{f}_i(y_0, \dots, y_{k-1})]$, $i = 1, \dots, n^2$ can be written in the form:

$$\int \bar{f}_i(w)(\Xi(y_0))(dw) \quad \text{and} \quad \int \bar{f}_i(w)(\Xi^\nu(y_0))(dw).$$

We claim that:

Claim 1: For large ν, it holds:

$$\left|\int \bar{f}_i(w)(\Xi(y_0))(dw) - \int \bar{f}_i(w)(\Xi^\nu(y_0))(dw)\right| < \delta_1/n^2,$$

for any $y_0 \in D$ and any $i = 1, \dots, n^2$.

In order to prove the Claim 1, recall that any uniformly continuous function may be approximated by a Lipschitz one. Let $\bar{f}_i' : D \to \mathbb{R}$, $i = 1, \dots, n^2$ be Lipschitz functions such that $\|\bar{f}_i' - \bar{f}_i\| < \delta_1/(4n^2)$. Denote by $\bar{L}$ the maximum bounded Lipschitz norm of the functions $\bar{f}_i'$, i.e. $\bar{L} = \max_{i=1,\dots,n^2} \|\bar{f}_i'\|_{BL}$. Hence,

$$\begin{aligned} \left| \int \bar{f}_i(w)(\Xi(y_0))(dw) - \int \bar{f}_i(w)(\Xi^\nu(y_0))(dw) \right| \leq \\ \leq \int \left| \bar{f}_i(w) - \bar{f}_i'(w) \right| (\Xi(y_0) + \Xi^\nu(y_0))(dw) + \\ + \left| \int \bar{f}_i'(w)(\Xi(y_0))(dw) - \int \bar{f}_i'(w)(\Xi^\nu(y_0))(dw) \right|. \end{aligned}$$

The first term is bounded by $\delta_1/(2n^2)$, due to the fact that $\Xi(y_0)(\cdot)$ and $\Xi^\nu(y_0)(\cdot)$ are probability measures. To bound the second term, let us observe that Lemma 4 implies the existence of a positive integer ν_0 such that $\beta(\Xi^\nu(y_0), \Xi(y_0)) < \delta_1/(2n^2\bar{L})$, for any $\nu \geq \nu_0$. This completes the proof of Claim 1.

Therefore, for $\nu \geq \nu_0$ it holds:

$$f_j \left(E \begin{bmatrix} \bar{f}_1(y_0^\nu, \dots, y_{k-1}^\nu) \\ \vdots \\ f_{n^2}(y_0^\nu, \dots, y_{k-1}^\nu) \end{bmatrix} \right) > 0,$$

for any $y_0^\nu \in D$ and $j = 1, \dots, n$. Hence,

$$E[(x_k^\nu)^T x_k^\nu | x_0^\nu, y_0^\nu] < a(x_0^\nu)^T x_0^\nu.$$

Integrating over the distribution of x_0^ν, y_0^ν, we conclude:

$$E[(x_k^\nu)^T x_k^\nu] < aE[(x_0^\nu)^T x_0^\nu], \tag{3.55}$$

for any initial condition and any $\nu \geq \nu_0$, which completes the proof of the proposition. □

Let us then state a corollary of the Proposition 6, dealing with systems having also additive stochastic disturbance. Particularly, consider the systems given by:

$$x_{k+1}^\nu = A(y_k^\nu)x_k^\nu + w_k^\nu, \quad y_{k+1}^\nu \sim \bar{K}^\nu(y_k^\nu, \cdot), \tag{3.56}$$

and:

$$x_{k+1} = A(y_k)x_k + w_k, \quad y_{k+1} \sim \bar{K}(y_k, \cdot), \tag{3.57}$$

where w_k^ν and w_k are zero mean i.i.d. random variables with finite variances.

Corollary 3. *Consider the systems described by* (3.56) *and* (3.57). *Assume that* $\bar{K}^\nu \to \bar{K}$ *weakly and that* $\bar{K}$ *is Feller continuous. Let* $a < 1$ *and assume that* $T = T_{A,\bar{K}}$ *has spectral* *radius less than $1/a$. Then for any $\varepsilon > 0$, there exist positive integers k_0, ν_0 such that:*

$$E\left[\sum_{k=k_0}^{\infty} a^k (x_k^\nu)^T x_k^\nu\right] \leq \varepsilon\left(1 + E\left[(x_0^\nu)^T x_0^\nu\right]\right) \tag{3.58}$$

$$E\left[\sum_{k=k_0}^{\infty} a^k x_k^T x_k\right] \leq \varepsilon\left(1 + E\left[x_0^T x_0\right]\right), \tag{3.59}$$

for any $\nu \geq \nu_0$.

Proof: The proof is straightforward and uses equation (2.5), as well as the bound (3.55) and techniques form Chapter 2. □

3.8.2 Weak Convergence and ε - Optimality

In what follows, we assume that $\bar{K}^\nu \to \bar{K}$ weakly, $\bar{K}$ is Feller continuous and the functions $A^\nu(y,k), A(y,k), A(y)$ are continuous on their y argument. Let us then introduce some notation, needed in order to state the basic results.

Notation: Consider the system given by:

$$x_{k+1} = A'(y_k, k)x_k + B(y_k)u_k + w_k, \quad y_{k+1} \sim \bar{K}'(y_k, \cdot),$$

and the feedback control law $u_k = L_k(y_k)x_k$. Then, we denote by:

$$J_{\bar{K}', k_0, A', u_k = L_k(y_k)x_k}(x_0, y_0) = E\left[x_{k_0+1}^T Q_{k_0+1} x_{k_0+1} + \right.$$
$$\left. + \sum_{k=0}^{k_0} a^k x_k^T [L_k^T(y_k) R(y_k) L_k(y_k) + Q(y_k)] x_k\right],$$

and $J^\star_{\bar{K}', k_0, A'}(x_0, y_0)$ the optimal value, where K' can take the values $\bar{K}$ or $\bar{K}^\nu$ and the time horizon is allowed to take the infinity value. We use the notation $J_{\bar{K}', k_0, u_k = L_k(y_k)x_k}$ and $J^\star_{\bar{K}', k_0}(x_0, y_0)$ for $A'(y_k, k) = A(y_k)$. □

The basic topic of this section is the proof of the following two propositions about the ε - optimality in the finite and infinite horizon LQ control problems respectively.

Proposition 7. *Assume that $A^\nu(y_k, k) \to A(y_k, k)$ as $\nu \to \infty$. Let us denote by $u_k = L_k(y_k)x_k$ the optimal control law that attains the minimum $J^\star_{\bar{K}, A(y_k,k), k_0}$. Then for any $\varepsilon > 0$ there exists a positive integer ν_0 such that:*

$$J_{\bar{K}^\nu, k_0, A^\nu(y_k,k), u_k = L_k(y_k)x_k}(x_0, y_0) \leq J^\star_{\bar{K}^\nu, k_0, A(y_k,k)} +$$
$$+ \varepsilon(1 + x_0^T x_0), \tag{3.60}$$

for any $\nu \geq \nu_0$. □

Proposition 8. *Let us denote by $u_k = L(y_k)x_k$ the feedback strategy that attains the minimum $J^\star_{\bar{K},\infty}$. Then, for any $\varepsilon > 0$ there exists a positive integer ν_0 such that:*

$$J_{\bar{K}^\nu,\infty,u_k=L(y_k)x_k}(x_0, y_0) \leq J^\star_{\bar{K}^\nu,\infty} + \varepsilon(1 + x_0^T x_0), \tag{3.61}$$

for any $\nu \geq \nu_0$. □

The proof of the Propositions 7 and 8 depends on the following lemmas.

Lemma 5. *Consider a feedback control law $u_k = L_k(y_k)$ which is continuous in y_k. Then for any $\varepsilon > 0$, there exists a positive integer ν_0 such that:*

$$\begin{aligned} \left| J_{\bar{K},k_0,u_k=L_k(y_k)x_k}(x_0, y_0) - J_{\bar{K}^\nu,k_0,u_k=L_k(y_k)x_k}(x_0, y_0) \right| < \\ < \varepsilon(1 + x_0^T x_0) \end{aligned}$$

for any $\nu \geq \nu_0$.

Proof: The proof is a direct consequence of the properties of the weak convergence [Kar75]. □

Lemma 6. *Let $f_\nu : D \to \mathbb{R}$ be a sequence of continuous functions and f their pointwise limit. Then, it holds:*

$$\int_D f_\nu(y')\bar{K}^\nu(y, dy') \to \int_D f(y')\bar{K}(y, dy')$$

as $\nu \to \infty$.

Proof: The proof is a direct consequence of the compactness of D and the properties of the weak convergence. □

Lemma 7. *For any $\varepsilon > 0$, there exists a positive integer ν_0 such that:*

$$\left| J^\star_{\bar{K},k_0}(x_0, y_0) - J^\star_{\bar{K}^\nu,k_0,}(x_0, y_0) \right| < \varepsilon(1 + x_0^T x_0)$$

for any $\nu \geq \nu_0$.

Proof: The proof proceeds backwards in time from the step k_0 to 0. At each step the Dynamic programming equations, as well as the Lemma 6 are used. □

We then proceed to the proof of the basic results of the current section.

Proof of Proposition 7: Lemma 5 implies the existence of a positive integer ν_{01} such that:

$$\begin{aligned} & J_{\bar{K}^n u,k_0,u_k=L_k(y_k)x_k}(x_0, y_0) < \\ & \quad < J_{\bar{K},k_0,u_k=L_k(y_k)x_k}(x_0, y_0) + \varepsilon(1 + x_0^T x_0) \\ & \quad = J^\star_{\bar{K},k_0}(x_0, y_0) + \varepsilon(1 + x_0^T x_0), \end{aligned} \tag{3.62}$$

for any $\nu \geq \nu_{01}$. On the other hand, Lemma 7 implies the existence of a positive integer ν_{02} such that:

$$J^{\star}_{\bar{K},k_0}(x_0, y_0) < J^{\star}_{\bar{K}^{\nu},k_0}(x_0, y_0) + \varepsilon(1 + x_0^T x_0) \tag{3.63}$$

for any $\nu \geq \nu_{02}$. Inequalities (3.62) and (3.62) imply the desired result for $\nu_0 = max\{\nu_{01}, \nu_{02}\}$. □

Proof of Proposition 8: In order to complete the proof, the following series of comparisons is made:

$J_{\bar{K}^{\nu},\infty,u_k=L(y_k)x_k}$, $J_{\bar{K}^{\nu},k_0,u_k=L(y_k)x_k}$, $J_{\bar{K},k_0,u_k=L(y_k)x_k}$, $J^{\star}_{\bar{K},\infty}$, $J^{\star}_{\bar{K},k_0'}$, $J^{\star}_{\bar{K}^{\nu},k_0'}$, $J^{\star}_{\bar{K}^{\nu},\infty}$.

Particularly, each of these quantities is compared to the next one. It is shown that each of these quantities is at most slightly larger than the next.

At first let us compare $J_{\bar{K}^{\nu},\infty,u_k=L(y_k)x_k}$ with $J_{\bar{K}^{\nu},k_0,u_k=L(y_k)x_k}$. Corollary 3 implies that for any $\varepsilon > 0$ there exist integers k_0 and ν_{01} such that:

$$J_{\bar{K}^{\nu},\infty,u_k=L(y_k)x_k}(x_0, y_0) \leq\leq J_{\bar{K}^{\nu},k_0,u_k=L(y_k)x_k}(x_0, y_0) + \varepsilon(1 + x_0^T x_0)/4,$$

for any $\nu \geq \nu_{01}$.

Lemma 5 implies the existence of a positive integer ν_{02} such that:

$$J_{\bar{K}^{\nu},k_0,u_k=L(y_k)x_k}(x_0, y_0) \leq J_{\bar{K},k_0,u_k=L(y_k)x_k}(x_0, y_0) + \varepsilon(1 + x_0^T x_0)/4,$$

for any $\nu \geq \nu_{02}$, which serves as the second comparison.

Comparison three holds as an inequality, i.e.:

$$J_{\bar{K},k_0,u_k=L(y_k)x_k}(x_0, y_0) \leq J^{\star}_{\bar{K},\infty}(x_0, y_0).$$

To derive an inequality for the fourth comparison, let us observe that $K_k \to K$ and $c_k \to c$ uniformly as $k \to \infty$, where K_k, K, c_k and c are as in Proposition 3 and Theorem 2 of [KP14a]. It also holds $J^{\star}_{\bar{K},k} = x_0^T K_k(y_0)x_0 + c_k$ and $J^{\star}_{\bar{K},\infty} = x_0^T K(y_0)x_0 + c$. Thus, for any $\varepsilon > 0$, there exists a positive integer k_0' such that:

$$J_{\bar{K},\infty}(x_0, y_0) \leq J^{\star}_{\bar{K},k_0'}(x_0, y_0) + \varepsilon(1 + x_0^T x_0)/4.$$

Lemma 7 implies the existence of a positive integer ν_{03} such that:

$$J_{\bar{K},k_0'}(x_0, y_0) \leq J^{\star}_{\bar{K}^{\nu},k_0'}(x_0, y_0) + \varepsilon(1 + x_0^T x_0)/4,$$

for any $\nu \geq \nu_{03}$. This inequality shows the fifth comparison.

The last comparison holds as an inequality, i.e.:

$$J^{\star}_{\bar{K}^{\nu},k_0'}(x_0, y_0) \leq J^{\star}_{\bar{K}^{\nu},\infty}(x_0, y_0).$$

Hence, choosing $\nu_0 = \max\{\nu_{01}, \nu_{02}, \nu_{03}\}$ the desired result is shown. □

3.8.3 Proof of Theorem 1

The proof is a direct consequence of the Propositions 7 and 8 and the continuity of the functions involved.

Chapter 4

Games on Large Networks: Information and Complexity.

This chapter studies Static and Dynamic Games on large Networks of interacting agents assuming that the players have some statistical description of the interaction graph, as well as some local information. Inspired by Statistical Physics, we consider statistical ensembles of games and define a Probabilistic Approximate Equilibrium notion for ensembles of games. A Necessary Information Complexity notion is also introduced, quantifying the minimum amount of information needed in order to exist a Probabilistic Approximate equilibrium. Some classes of games are then analyzed and upper and lower bounds for the complexity are found. Particularly, static and dynamic games on random graphs are considered and it is shown to have high complexity for low connectivity and low complexity for high connectivity. In the latter case, Probabilistic Approximate equilibrium strategies are computed. Static games on lattices are then considered and upper and lower bounds for the complexity were derived using contraction mapping ideas. A LQ game on a large ring is also studied numerically and an upper bound for the complexity is found to be approximately linear.

4.1 Introduction

In the last decade, a lot attention has been devoted to the study of games with large number of players under the name of Mean Field Games (MFGs) [GLL10]. These models study situations where each individual interacts with the mass of the other players (mean field interaction). Asymptotic Nash equilibrium results are usually obtained using the assumption that each player can measure only her state variable and also knows the statistical distribution of the types and state variables of the rest of the players. This work aims to study more general interaction structures, in which it is possible to have an approximate equilibrium assuming only local and statistical information.

In several game situations involving many agents, the strategic interactions depend on a large interaction structure such as a network. An example is an electricity transmission grid where several entities, such as producers, consumers or smart micro-grids are connected in different places of the network and local, as well as global cooperation and/or competition arises (ex. [SHPB12]). Several examples involve interactions over social networks [Jac08], [DE10], such as the selection of a telecommunication company, the opinion about an idea or a product, the selection of fashion group and the engagement in criminal behavior. In these examples, the choice of each agent depends on her preferences, as well as the choices of her friends. There are also several examples of non-social interaction structures, such as the interaction among the owners of stores for renting and the gas station prices, where there exists a local, as well as a global competition. Several features characterizing Systems of Systems [KH11], such as operational and managerial independence, geographical distribution, heterogeneity of systems and networks of systems can be also captured by Dynamic Games on Networks models.

Two kinds of approach have been mainly used to predict the behavior of the participants in large games. The first approach is based on equilibrium concepts and the dominant notion in this approach is the Nash equilibrium. The knowledge of a large amount of information is often needed in order to be possible to determine equilibrium or approximate equilibrium. The second kind of approach assumes limited (bounded) rationality for the participants [Sim72], [Sel01] and it is based on dynamic formulations. In particular, some deterministic or stochastic rules describing the future actions of the agents as a function of their current actions are postulated and then evaluated theoretically or experimentally. Examples of dynamic rules include evolutionary dynamics, learning, adaptive control laws and best response.

This second kind of approach does not require a complete knowledge of the game. However, the dynamic rules used are not universal, in the sense that there is no reason to believe that all the players will follow some specified rule to determine their future actions. Furthermore, there is also the cheating problem [KP14b], i.e. that the knowledge of the dynamic/adaptation rule of a player may be exploited by the others, leading probably to different outcomes from the predicted.

Which kind of approach should be used in order to describe/predict the behavior of the players in a large game? The full rationality assumption for the players should depend on the difficulty of the problem they have to solve, as well as the informational requirements. In this work, we define an informational complexity concept, for a certain class of games, as the minimum amount of information needed in order to be achievable a certain form of approximate equilibrium. This complexity notion is studied asymptotically for large games. If the complexity is small, then the participants of a large game can use strategies that are in approximate equilibrium without assuming a lot of information. In a large game with very high complexity, it is reasonable to assume that the players would use a dynamic rule. In this sense, our approach aims at the one hand to identify classes of large games which admit approximate equilibrium solutions assuming only a small portion of the information and on the other hand to identify some informational limitations of the applicability of equilibrium concepts in static or dynamic games with a large number of players.

4.1.1 Contribution

We study Dynamic Games on Large Networks, assuming some stochastic description of the Graph such as Random Graph, Grid, Small World Network etc. The stochasticity is divided into two parts. The first part describes the structural stochasticity and corresponds to the lack of knowledge of the players the and the second corresponds to the lack of predictability of the future.

Instead of studying a single game, inspired by ideas arisen in the Statistical Physics domain, an ensemble (collection) of games is considered, assuming that a statistical model on that ensemble is available to the players. Furthermore, the information contained in a neighborhood of a certain order of each player is assumed to be available to that player. As the order of neighborhoods increases the information varies from only statistical to perfect. An analogy with a situation common in the Statistical Physics literature is used in order to define a Probabilistic Approximate Nash (PAN) equilibrium concept for ensembles of games. A complexity function is then defined for an ensemble of games as the minimum amount of information needed in order to be achievable a PAN equilibrium. This complexity function is studied asymptotically as the number of the participants of the game becomes large.

Special cases of games where upper or lower bounds for the complexity can be obtained are then analyzed. A class of static games on Random Graphs is shown to be simple under some connectivity assumptions and complex under other connectivity assumptions. A LQ opinion game on a Random Graph is also analyzed and approximate equilibrium strategies are computed, generalizing older work about consensus [NCMH13].

Quadratic Games on ordered interaction Structures (Grids or Rings) are then analyzed. They are found to have polynomial complexity functions where the order of the polynomial depends on the dimension of the interaction structure. A class of non-quadratic games on rings having chaotic best response maps are analyzed and shown to have a relatively low complexity. The approximate equilibrium is possible due to some cooperation among the players. Finally, a LQ game on a ring is studied numerically. Approximate equilibrium strategies are computed, using a reduced game and an upper bound for the complexity function is found to be approximately linear.

In the special cases analyzed, we established low complexity results using one of the following properties: a law of large numbers, contractivity of the best response maps, cooperation among the players or low gains assigned to distant players.

4.1.2 Related Topics

The interest for the games with large number of players is not new. Probably the first works dealing with games involving a continuum of players are [Aum64] and [MS78]. [Aum64] analyzes a market with a continuum of players and [MS78] studies games with a continuum of players, called Oceanic Games and introduces and a value for such games. The Mean Field

Games [GLL10] have been recently introduced to study static and dynamic games with large number of players. The closely related methodology of Nash Certainty Equivalence was also developed, in order to obtain asymptotic approximate Nash equilibrium results, as the number of players tends to infinity [HMC05],[HMC06]. These works study games, where each player interacts with the mass of the other players, which is approximated by a continuum. Large games involving a coordinator (major player) were studied in [Hua10]. Several extensions of the Mean Field game theory on models describing more general interactions are presented in [Tem11b].

Another related topic is Games with Local Interactions, in which each player interacts with some players important to her on some organized structure. In [BHO06], equilibria for complete and incomplete information Local Interaction Games were found, based on contraction mapping ideas. The dynamic game counterpart is presented in [BO11]. Models with discrete choice were introduced in [BD01].

Games where players move on a graph were analyzed in [OP88], [BB90] and finite games on graphs, where each node corresponds to a participant of the game, were studied in [KLS01]. Repeated games with random matching of the opponents were introduced in [Kan92], in the context of sustainability of cooperation and social norms. The probability of existence of a Nash equilibrium for games on random graphs is studied in [DDM11]. A quadratic game on networks is studied in [BCAZ06], using centrality notions. Games on networks with incomplete information are studied in [GGJ$^+$10]. Dynamic games on evolving state dependent graphs were studied in [BB10] and stochastic games in [NAB09]. A review of network games is given in [JZng].

Dynamic rules for updating the actions of the agents on lattices were studied in the context of Interacting Particle Systems [Lig85] and in [Blu93]. Several dynamic rules for games on graphs were introduced and studied analytically and computationally in [SF07]. Several sociological applications of evolutionary games on graphs were studied in [Sky03].

The impact of the quality of information that the agents receive on their costs is studied in [TP85] in the LQG Game framework. It was shown that, as the number of players or the time horizon becomes large, better information becomes beneficial for the all the participants of the game. The notion of the price of uncertainty was introduced in [BBM09] and [GJC10], in order to describe the difference in the costs of the players under different information structure using dynamic and equilibrium formulations respectively. The price of information was introduced in [BZ11] to describe the difference of the cost that the players have in deterministic dynamic games under different information patterns, i.e. feedback and open loop.

4.1.3 Notation

The underlying probability space is denoted by $(\Omega, \mathcal{F}, P)$. For a random variable X, denote by $\sigma(X)$ the σ algebra generated by X, i.e. the coarsest σ algebra such that X is $\sigma(X)$-measureable.

A directed or undirected graph will be denoted by $G = (V, E)$, where V is the set of vertices and E the set of edges (ex. [Wes01], [ME10]). For a vertex $v \in V$, the neighborhood of v is defined as $\mathcal{N}_v(G) = \{v' \in V : (v', v) \in E\}$ and the closed neighborhood of v as $\bar{\mathcal{N}}_v(G) = \{v\} \cup \{v' \in V : (v', v)$ or $(v, v') \in E\}$. The closed neighborhood of order n of v is defined as $\bar{\mathcal{N}}_v^n(G) = \cup_{j \in \bar{\mathcal{N}}_v(G)} \bar{\mathcal{N}}_j^{n-1}(G)$. For a subset A of V, denote by $G_A = (A, \{(v', v) \in E : v, v' \in A\})$, the largest subgraph of G with set of vertices A. With $|G'|$ we denote the number of nodes of a graph G'.

The indeterminacy $\frac{0}{0}$ is resolved as 0. An ordered tuple $(\gamma^1, \dots, \gamma^N)$ is denoted by $(\gamma^i)_i$ and the ordered tuple $(\gamma^1, \dots, \gamma^{i-1}, \gamma^{i+1}, \dots, \gamma^N)$ by γ^{-i}. By $[\cdot]$ we denote the integer part. The asymptotic notation will be also used. For real functions f and g, we write $f(x) \in O(g(x))$, if there exists a constant $c > 0$, such that $0 \leq f(x) \leq cg(x)$ for large x and $f(x) \in o(g(x))$, if for any $c > 0$ it holds $0 \leq f(x) \leq cg(x)$, for large x. We write $f(x) \in \omega(g(x))$ if for any $c > 0$ it holds $0 \leq cg(x) \leq f(x)$ for large x and $f(x) \in \Theta(g(x))$ if there exist constants c_1 and c_2 such that $0 \leq c_1 g(x) \leq f(x) \leq c_2 g(x)$ for large x. Finally, $f(x) \in \Omega(g(x))$ if for some constant c, it holds $f(x) \geq cg(x)$, for large x.

In what follows, we assume that all the functions involved are measurable within appropriate measurable spaces.

4.1.4 Organization

The rest of the current work is organized as follows. Section 4.2 describes the notion of an ensemble of static or dynamic games. In Section 4.3, the Probabilistic Approximate Nash equilibrium is introduced and compared with the Bayesian Nash equilibrium. Furthermore, the Necessary Information Complexity and the Asymptotic Necessary Information Complexity functions are defined. Several special cases are then analyzed. Particularly, static games on large Erdos-Renyi random graphs are studied in section 4.4.1. In section 4.4.3 the LQ dynamic game counterpart is considered. Section 4.5 studies static games on organized structures. Particularly, Subsection 4.5.1 studies static quadratic games on lattices and Subsection 4.5.2 studies games with chaotic best response maps on rings. LQ games on rings are considered and studied numerically in Section 4.6.2. Finally, Section 4.7 concludes.

4.2 Description

We first describe the general form of the structure of interactions among the players. For any given interaction structure S, a game g^S is defined. Then, an ensemble of interaction structures (or equivalently an ensemble of games) $\mathcal{E}$ is defined assuming that the players have a common probabilistic description for the interaction structures S in $\mathcal{E}$.

Let us first describe the general form of an interaction structure. There is a set of players $p_1, \dots, p_N$. Each player p_i has a type θ_i belonging to a set of possible types Θ. The interaction structure depends on a graph $G = (V, E)$, directed or not. Each vertex of the graph $v \in V$

corresponds to a player p_v and each edge $(v', v) \in E$ to the influence of the player $p_{v'}$ to the player p_v. The interaction structure is described in compact form by:

$$S = (\Pi, G), \tag{4.1}$$

where $\Pi = ((p_1, \theta_1), \dots, (p_N, \theta_N))$. Let us denote by $\mathcal{F}^S = \sigma(S)$, the structural information.

For a given interaction structure S, a discrete time dynamic game g^S among the players is described. Each player p_i has a state variable x^i the evolution of which is affected by the actions and the state variables of the other players. The influence of the neighboring players is different than the other players. The dynamics for the player p_i is given by:

$$\begin{aligned} x^i_{k+1} = f^{\theta_i}\Big(x^i_k, u^i_k, w^i_k, &\sum_{j\in\mathcal{N}_i(G)} f_1^{\theta_i,\theta_j}(x^i_k, x^j_k, u^i_k, u^j_k), \\ &\sum_{j\in V} f_2^{\theta_i,\theta_j}(x^i_k, x^j_k, u^i_k, u^j_k)\Big), \end{aligned} \tag{4.2}$$

where $u^i_k \in U$ is the action of player p_i at time step k, w^i_k are random variables with known distributions, $\mathcal{N}_i(G)$ is the neighbourhood of player p_i and V the set of all players. The initial conditions are random variables, possibly dependent on the interaction structure.

The cost functions, given the interaction structure, are given by:

$$\begin{aligned} J_i = E\Bigg\{\sum_{k=0}^{T} \rho^k g^{\theta_i}\Big(x^i_k, u^i_k, &\sum_{j\in\mathcal{N}_i(G)} g_1^{\theta_i,\theta_j}(x^i_k, x^j_k, u^i_k, u^j_k), \\ &\sum_{j\in V} g_2^{\theta_i,\theta_j}(x^i_k, x^j_k, u^i_k, u^j_k)\Big) \Bigg| \mathcal{F}^S\Bigg\}, \end{aligned} \tag{4.3}$$

where the time horizon T can be finite or infinite, $g \geq 0$ and $\rho \in (0, 1]$ is the discount factor. The cost function of the player p_i, also, depends differently on the players with whom she has a direct connection, than the rest of the players.

The players do not know the interaction structure characterizing the game that they are involved. Instead, they consider a statistical ensemble $\mathcal{E}$ of possible interaction structures, i.e. a collection of mental copies of the game having different interaction structures. With a slight abuse of notation, the corresponding ensemble of games is also denoted by $\mathcal{E}$. We assume that all the players consider the same ensemble and that they assume the same probability structure on that ensemble. That is, the players assume the same distribution of the random variable S in $\mathcal{E}$. Let us denote by $Q(\cdot)$ that distribution.

Apart form the statistical model of the interaction structure $(\mathcal{E}, Q)$, the players possess some local information. The local information of a player p_i consists of the structural facts and the state variables of the players contained in a neighborhood of order n of that player. Particularly, we assume that each player knows her type and can measure her own state variable. Furthermore, each player knows the interaction structure of a neighborhood of order n around her, that is she knows the types of the players and the subgraph and she can also measure the state variables of the players in that neighborhood. Thus, the local information available to the player p_i at time step k is:

$$I_k^{i,n} = (G_{\bar{\mathcal{N}}_i^n(G)}, (\theta_j)_{j\in\bar{\mathcal{N}}_i^n(G)}, (x_t^j)_{j\in\bar{\mathcal{N}}_n^i(G)}^{t=0,\dots k}), \tag{4.4}$$

where $\bar{\mathcal{N}}_i^n(G)$ was defined in Section 4.1.3 and denotes the neighborhood of order n of player p_i and $G_{\bar{\mathcal{N}}_i^n(G)}$ is the corresponding subgraph. The total information available a player p_i at time step k is $(\mathcal{E}, Q, I_k^{i,n})$.

Due to the fact that the players do not know which is the actual interaction structure S, they use strategies that can be applied in any member of the ensemble $\mathcal{E}$. The strategy of each agent can, however, depend on the local information available to her. We consider symmetric sets of strategies, where players with the same information (and hence type) behave in the same way. Furthermore, we focus on feedback strategies (strategies without memory). The strategies under consideration have, thus, the form:

$$u_k^i = \gamma_k(\bar{I}_k^{i,n}), \tag{4.5}$$

where:

$$\bar{I}_k^{i,n} = (G_{\bar{\mathcal{N}}_i^n(G)}, (\theta_j)_{j\in\bar{\mathcal{N}}_i^n(G)}, (x_k^j)_{j\in\bar{\mathcal{N}}_n^i(G)}). \tag{4.6}$$

The following classes of strategies will be useful in the next section.

Strategy Classes:

(i) The class of Feedback Local Information Strategies for player i is given by:

$$\Gamma_i^{FLI} = \{\gamma^i = (\gamma_1^i, \gamma_2^i, \dots) : \gamma_k^i : \bar{I}_k^{i,n} \to U, k = 1, 2, ..\}.$$

If necessary, $\Gamma_{i,n}^{FLI}$ will be used in the place of Γ_i^{FLI} to indicate the order n of the information neighborhoods.

(ii) The class of Closed Loop Perfect Information Strategies for player i is given by:

$$\Gamma^{CLPI} = \{\gamma^i = (\gamma_1^i, \gamma_2^i, \dots) : \gamma_k^i : I_k^{CLPI} \to U, \\ k = 1, 2, \dots\},$$

where $I_k^{CLPI} = (G, (\theta_j)_{j\in V}, (x_t^j)_{j\in V}^{t\leq k})$, G is the interaction graph and V is the set of all the players. □

Remark 19. *(i) The dynamics and the cost functions given by equations (4.2) and* (4.3) *describe two types of interactions. The first sum corresponds to local interactions and the second term to mean field interactions.*

(ii) There are two types of stochasticity presented in the model. The first is due to the lack of predictability and it is described by the random variables w_k^i. The second is the uncertainty due to the lack of information (knowledge) and it is described the random variable S which contains the structural features of the game. There some other works which divide

the uncertainty to lack of knowledge and lack of predictability. For example, in [AB06] incomplete information is treated using robust optimization and the stochasticity due to the randomization of the players using the expectation.

(iii) The members of the ensemble do not need to have the same number of players and the ensemble do not need to be finite.

(iv) Models involving graphs with information on their edges could also be studied, as well as structures more general than the graphs relating more than two agents (ex.[GZCN09]). For simplicity reasons, we study only the model defined in this section. □

The model described borrows some ideas from the Statistical Physics and Network Science. In Statistical Physics the inability to know the initial conditions of the system under consideration precisely, as well as the fact that it is very difficult to solve a huge number of equations describing individual particles, is coped by considering a collection of mental copies of the system having different initial conditions, called a statistical ensemble [Hua01]. Several macroscopic properties are then shown to have values close to a deterministic constant for all the systems in the ensemble, except possibly of a set of systems with very low probability. Similar results have been also obtained in the Network Science, such as results about Percolation, connectivity of large Random Graphs, etc [Jac08], [DE10].

In the following section, we study sets of strategies which constitute an approximate Nash equilibrium for the games corresponding to all the interaction structures of the ensemble, except possibly of a set of interaction structures with very low probability. Following some ideas from Statistical Physics and Network science, in order to study games with large number of players, we consider a sequence of ensembles of games $\mathcal{E}_\nu$ with increasing number of players and we study the tail of the sequence.

4.3 Approximate Equilibrium and Complexity

Consider a large game in which the actions of the players depend only on local and statistical information. Due to the fact that the agents do not know in which game they participate in, it is reasonable to expect that a set of strategies in the form (4.5) could not typically constitute a Nash equilibrium. A Probabilistic Approximate Nash (PAN) equilibrium concept is thus defined, based on the concept of ε - Nash equilibrium. We first recall the definition of the ε - Nash equilibrium for a single game.

Definition 7. *Consider a game g^S with $S \in \mathcal{E}$ and the set of dynamics and cost functions given by (4.2) and* (4.3)*. Then a set of strategies $(\gamma^i)_i$ with $\gamma^i \in \Gamma_i^{FLI}$ constitutes an ε - Nash equilibrium, if for every player p_i it holds:*

$$J_i(\gamma^i, \gamma^{-i}) - \min_{\gamma \in \Gamma^{CLPI}} \{J_i(\gamma, \gamma^{-i})\} < \varepsilon, \tag{4.7}$$

where the minimum is considered within the class of full information closed loop strategies. □

An approximate equilibrium concept is then defined for the ensemble of games. We are interested to characterize a set of strategies constituting an ε - Nash equilibrium for the games g^S that correspond to the most of the interaction structures in $S \in \mathcal{E}$.

Definition 8. *Consider an ensemble of interaction structures $\mathcal{E}$ and the set of dynamics and cost functions as before. Then a set of strategies $(\gamma^i)_i$ with $\gamma^i \in \Gamma_i^{FLI}$ is an ε-Probabilistic Approximate Nash equilibrium (ε-PAN equilibrium) for that ensemble if:*

$$P(\{S \in \mathcal{E} : (\gamma^i)_i \textit{ is } \varepsilon \textit{ - Nash equilibrium of } g^S\}) > 1 - \varepsilon, \tag{4.8}$$

i.e. $(\gamma^i)_i$ constitutes an ε - Nash equilibrium with high probability. The probability of the event in (4.8) *can be computed using the distribution Q.* □

The reason for studying sets of strategies constituting an ε-PAN equilibrium with small ε is that, with a very high probability, no player has non-negligible benefit from changing her strategy, even if she had access to all the available information at any time step.

Remark 20. *An alternative way to express inequality* (4.8) *of the manuscript is to use the Ky Fan metric among random variables (ex. [Dud02]) defined as follows. The distance d between random variables X_1 and X_2 is defined as:*

$$d(X_1, X_2) = \inf \{\varepsilon > 0 : P(|X_1 - X_2| > \varepsilon) \leq \varepsilon\} \tag{4.9}$$

Thus, inequality (4.8) *can be stated equivalently as:*

$$d(J^i(\gamma^{-i}, \gamma^i), \inf_{\gamma \in \Gamma^{CLPI}} J^i(\gamma^{-i}, \gamma)) < \varepsilon \tag{4.10}$$

in terms of the Ky Fan metric. □

Lemma 8. *The ε-PAN equilibrium has the following properties:*

(i) Consider a set of strategies $(\gamma^i)_i$ constituting an ε-PAN equilibrium. If the players receive more information, i.e. the information neighborhoods have order $n' > n$ and $I^{i,n'} \supset I^{i,n}$, then the set of strategies $(\gamma^i)_i$ remains an ε-PAN equilibrium. That is, the ε-PAN equilibrium is insensitive to new information.

(ii) An ε - PAN equilibrium of a static game remains an ε - PAN for the corresponding repeated game.

Proof: (i) It holds $\gamma^i \in \Gamma_{i,n}^{FLI} \subset \Gamma_{i,n'}^{FLI}$. Furthermore, (4.8) holds and thus $(\gamma^i)_i$ is an ε-PAN equilibrium for the ensemble of games where the players have information $I^{i,n'}$.

(ii) Consider an ensemble of static games with cost functions J_i and a set of strategies $(\gamma^i)_i$ constituting an $\varepsilon-$PAN equilibrium. For a discount factor $\rho \in (0,1)$, an ensemble of repeated games can be considered having cost functions:

$$\tilde{J}_i = E\left[(1-\rho)\sum_{k=0}^{\infty}\rho^k J_i(\gamma^i,\gamma^{-i})\middle|\mathcal{F}^S\right]. \tag{4.11}$$

The set of strategies $(\tilde{\gamma}^i)_i$ with $\tilde{\gamma}^i = (\tilde{\gamma}^i_1, \tilde{\gamma}^i_2, \dots)$ and $\tilde{\gamma}^i_k = \gamma^i$ constitutes an $\varepsilon-$PAN equilibrium for the ensemble of repeated games. □

Let us now compare the ε-PAN equilibrium with the notion of Bayesian Nash equilibrium (see ex. [FT91b]).

Remark 21. *Some differences among the two approaches involve:*

(i) It is very difficult to compute a Bayesian Nash equilibrium even for simple dynamic games. In the case of LQ stochastic Dynamic Games with imperfect state feedback information, Nash equilibria have been computed only for special information patterns [Pap82]. [GNLB14]. In the case where the structural information is also incomplete, the optimization problems are vey difficult even for single person games (optimal control problems), due to the fact that the dual control problem arises [Wit02].

(ii) If there is only structural uncertainty, in contrast to Bayesian Nash equilibrium, an ε - PAN set of strategies is insensitive to the risk profile of the players.

(iii) In contrast to Bayesian Nash equilibrium, an ε - PAN set of strategies satisfies the properties (i),(ii) of Lemma 8.

(iv) In some examples, we may have PAN equilibrium, even if the players have different prior probabilities on the ensemble. In these cases, the common prior assumption can be weakened (ex. Sec. IV). □

Consider a set of strategies constituting an ε-PAN equilibrium. Each player is interested and responds to a different set of players and in this sense, each player is involved in (perceives) a different game. For example, consider the game of in Figure 4.1. Each player is affected by her neighbors in the graph through the terms f_1, g_1 and by the rest of the players through f_2, g_2. Assume also that there is an ε-PAN equilibrium assuming information neighbourhoods of order 1. Then, player p_1 acts as if he is involved in a game only with the players p_2, p_3 and p_5, the player p_5 in a game with p_1, p_2 and p_7 and so on.

If the order of the information neighborhoods of the players is small then it is probably not possible to have an ε-PAN equilibrium. Thus, the following question is quite interesting:

Question 1. *"Given a positive constant ε, what is the minimum amount of information that the agents need to have in order to achieve an ε-PAN set of strategies?"* □

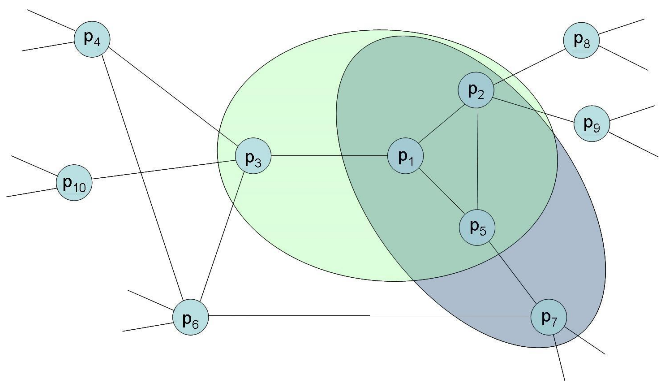


Figure 4.1: The information neighborhoods of players 1 and 5 with $n = 1$.

Based on the answer to this question, the Necessary Information Complexity (NIC) function with respect to the PAN equilibrium is defined for the ensemble of games.

Definition 9. *(i) Consider an ensemble of games as described above. Let us define the following function:*

$$\bar{n}(m) = \inf\{n \in \mathbb{N} : \exists \textit{ set of strategies } (\gamma^i)_i \textit{ with } \gamma^i \in \Gamma_{i,n}^{FLI}, \textit{ in the form (4.5) which is } 2^{-m}\textit{-PAN}\}. \tag{4.12}$$

The Necessary Information Complexity (NIC) function with respect to the PAN equilibrium is defined as:

$$C(m) = \max\{|\bar{\mathcal{N}}_i^{\bar{n}(m)}|\}, \tag{4.13}$$

where the maximum is considered over the several players and over the games of the ensemble where $(\gamma^i)_i$ is a 2^{-m} - Nash equilibrium.

(ii) Consider a sequence of ensembles $\mathcal{E}_\nu$ with cost functions J_i^ν and dynamics described by $f_1^{\cdot,\cdot,\nu}$, $f_2^{\cdot,\cdot,\nu}$. Denote by $C^\nu(\cdot)$ the NIC function of the ν-th ensemble. The Asymptotic Necessary Information Complexity (ANIC) function with respect to the PAN equilibrium is given by:

$$C_a(m) = \limsup_{\nu\to\infty} C^\nu(m). \tag{4.14}$$

The sequence of ensembles will be called asymptotically simple if the function $C_a(m)$ is bounded and asymptotically complex if for some $m \in \mathbb{N}$ it holds $C_a(m) = \infty$. □

In Definition 9, the informational complexity with respect to PAN equilibrium is introduced. The function $\bar{n}(m)$ quantifies the minimum order of the information neighborhood that the players need to have for the existence of a $2^{-m} - PAN$ equilibrium. The NIC function $C(m)$ quantifies the maximum number of players in a neighborhood of order $\bar{n}(m)$ or equivalently the maximum number of players that is required to be observed by a single player, in order for a $2^{-m} - PAN$ equilibrium to exist.

Remark 22. *Several static and dynamic games, that have only mean field interactions, have been studied in the literature [GLL10], [HMC05], [HMC06]. In these cases, under some conditions, each player interacts with the mass of the other players which behaves asymptotically deterministically, as the number of players increases. Thus, each player needs to know only her type and state variable, in order to behave nearly optimally. Thus, the Mean Field Games is a first example of asymptotically simple games.* □

Classes of games, where it is possible to find upper and lower bounds for the NIC and ANIC functions, are analyzed in the following sections.

4.4 Games on Random Graphs

In this section, ensembles of games where the players interact on large random graphs are studied. It is shown that the complexity of the ensemble varies, depending on the connection probability of the random graph. In the high connectivity regime the effects of the neighbors of any player can be approximated by a mean value. Using this approximation, appropriate consistency conditions are derived and then used to characterize a set of strategies which is shown to be ε-PAN for the large number of players case.

The first subsection considers static games and the second Linear Quadratic games.

4.4.1 Static Games on Random Graphs

Let us first describe the game for a given interaction structure. Each player has a type $\theta_i \in [0, L]$ and the cost functions are given by:

$$J_i = g(\theta_i - u^i) + g\left(u^i - \frac{1}{|\mathcal{N}_i(G)|}\sum_{j=1}^{N} d_{ij}u^j\right), \tag{4.15}$$

where g is a smooth, strictly convex function with $g(0) = g'(0) = 0$ and $d_{ij} = 1$, if there is an edge between vertexes i and j and zero otherwise.

In order to describe the ensemble of games, a stochastic structure on θ_i and the graph is described. We assume that θ_i are i.i.d. random variables uniformly distributed in $[0, L]$. The graph is an Erdos-Renyi random graph with connection probability c_N, i.e. each edge appears independently of the other edges with probability c_N. We further assume that the random variables d_{ij} and θ_i are mutually independent.

We first focus on strategies depending only on statistical information, assuming no knowledge about the neighbors of each player. The strategies under consideration have the form $u^i = \gamma(\theta_i)$.

A technique to derive strategies in this form is to approximate the terms in the cost function by their mean values. Specifically, we shall use the approximation:

$$\bar{u} \simeq \frac{1}{|\mathcal{N}_i(G)|}\sum_{j=1}^{N} d_{ij}u^j, \tag{4.16}$$

for all $i = 1, \ldots, N$.

With this approximation, the cost functions depend only on statistical information. The strategies that minimize the approximate cost functions have the form:

$$u^i = h(\theta_i, \bar{u}) = \arg\min_u \{g(u - \theta_i) + g(u - \bar{u})\}. \tag{4.17}$$

The function h will be shown to be well defined.

With the strategies given by equation (4.17), the mean value of the actions should satisfy the following compatibility condition:

$$\bar{u} = \frac{1}{L} \int_0^L h(\sigma, \bar{u}) d\sigma. \tag{4.18}$$

The following proposition shows that if connectivity is high, then the strategies described by (4.17), (4.18) constitute an ε-PAN equilibrium, for large N. The asymptotic notation used in the following proposition was defined in section 4.1.3.

Proposition 9. *Under the specified assumptions it holds:*

(i) Equation (4.18) *has a unique solution.*

(ii) If $c_N \in \omega\left(\frac{\ln N}{\sqrt{N}}\right)$*, then the strategies given by* (4.17) *constitute an* ε*-PAN equilibrium of the ensemble of games, for large* N*.*

(ii) If $c_N \in \omega\left(\frac{\ln N}{\sqrt{N}}\right)$*, then* $C_a(m) = 1$ *and the ensemble of games is asymptotically simple.*

Proof: (i) The strict convexity and the lower boundedness of g imply that the function h is well defined. The function $h(\theta, \bar{u})$ can be expressed as the solution to the following equation:

$$f_\theta(u, \bar{u}) = g'(u - \bar{u}) + g'(u - \theta) = 0,$$

with respect to u. Thus it holds:

$$\min\{\theta, \bar{u}\} \leq h(\theta, \bar{u}) \leq \max\{\theta, \bar{u}\}. \tag{4.19}$$

Consider the mapping:

$$\bar{u} \to T\bar{u} = \frac{1}{L} \int_0^L h(\sigma, \bar{u}) d\sigma.$$

Inequalities in (4.19) imply that $TL \leq L$ and $T0 \geq 0$. Thus, due to the intermediate value theorem, there is a $\bar{u}^\star$ such that $\bar{u}^\star = T\bar{u}^\star$.

The derivative of h with respect to $\bar{u}$ can be expressed, using the implicit function theorem as:

$$\begin{aligned} \frac{\partial h}{\partial \bar{u}} &= -\left(\frac{\partial f_\theta}{\partial u}\right)^{-1} \frac{\partial f_\theta}{\partial \bar{u}}\bigg|_{h(\theta,\bar{u}),\bar{u}} = \\ &= \frac{g''(h(\theta, \bar{u}) - \bar{u})}{g''(h(\theta, \bar{u}) - \bar{u}) + g''(h(\theta, \bar{u}) - \theta)} < 1. \end{aligned}$$

Thus, the solution $\bar{u}^\star$ is unique. In what follows, the unique solution of (4.18) will be denoted by $\bar{u}$.

(ii) The functions g and h are continuous. Using the strategies given by (4.17), the arguments of the functions belong to compact intervals. In those intervals, g and h are uniformly continuous. Thus, in order to show that the set of strategies given by (4.17) constitute an ε - PAN equilibrium, for large N, it suffices to show that for any $\varepsilon, \delta > 0$ it holds:

$$P\left(\exists i : |\bar{u} - \frac{1}{|\mathcal{N}_i(G)|}\sum_{j=1}^{N} d_{ij}u^j| > \delta\right) < \varepsilon,$$

for large N.

It holds:

$$\begin{aligned}\left|\bar{u} - \frac{1}{\mathcal{N}_i(G)}\sum_{j=1}^{N} d_{ij}u^j\right| \leq \left|\bar{u} - \frac{1}{N}\sum_{j=1}^{N} h(\theta_j, \bar{u})\right| + \\ + \left|\frac{1}{N}\sum_{j=1}^{N} h(\theta_j, \bar{u}) - \frac{1}{Nc_N}\sum_{j=1}^{N} d_{ij}u^j\right| + \\ + \left|\frac{1}{Nc_N}\sum_{j=1}^{N} d_{ij}u^j - \frac{1}{|\mathcal{N}_i(G)|}\sum_{j=1}^{N} d_{ij}u^j\right| \end{aligned} \tag{4.20}$$

It holds $\bar{u} = \int_0^L h(\sigma, \bar{u})/L d\sigma$. Due to the Glivenko-Cantelli theorem [Bil08], the empirical distribution, $\sum_{i=1}^{N} \delta_{x_i}/N$ converges a.s. to the uniform distribution as $N \to \infty$. Thus, there exists an integer N_{01}, such that:

$$\left|\bar{u} - \sum_{j=1}^{N} h(\theta_j, \bar{u})/N\right| < \delta/3,$$

with probability larger than $1 - \varepsilon/3$ for any $N \geq N_{01}$.

For the second term of the right hand side of (4.20) using (4.17) and Lemma 10 of the Appendix 4.8.1 it holds:

$$\left|\frac{1}{Nc_N}\sum_{j=1}^{N}(c_N - d_{ij})h(x_j, \bar{u})\right| < \delta/3,$$

with probability larger than $1 - \varepsilon/3$, for any $N \geq N_{02}$.

Lemma 11 of the Appendix 4.8.1 implies that the third term of (4.20) is less than $\delta/3$ with probability larger than $1 - \varepsilon/3$, for any $N \geq N_{03}$.

This completes the proof of (ii).

(iii) An immediate consequence of (ii) □

Remark 23. *In the proof of Proposition 9 it is used only the fact that the random variables* $(d_{ij})_{j=1}^N$ *are independent, for any* i. *Thus, the same result holds also for a more general class* *than the Erdos-Renyi random graph. A particular example is games on random directed graphs. Furthermore, there is no need for all the players to assume the same connection probability* c_N *or exactly the same random graph model. Thus, the "common prior" assumption can be weakened.*

Example 11. *If the function* g *has the form* $g(z) = z^2$*, the strategies given by* (4.17) *can be computed explicitly. It holds:*

$$h(\theta, \bar{u}) = (\theta + \bar{u})/2. \tag{4.21}$$

Equation (4.18) *implies that* $\bar{u} = L/2$ *and the set of strategies given by:*

$$u^i = \theta_i/2 + L/4, \tag{4.22}$$

constitute an ε*-PAN equilibrium for large* N*, if* $c_N \in \omega\left(\frac{\ln N}{\sqrt{N}}\right)$. ■

The following proposition studies the case of low connectivity.

Proposition 10. *For any integer* μ*, if* $c_N \in o\left(\frac{1}{N^{\mu/(\mu-1)}}\right)$ *then* $C_a(m) \leq \mu$ *and the ensemble of games is asymptotically simple.*

Proof: If $c_N \in o\left(\frac{1}{N^{\mu/(\mu-1)}}\right)$ then with probability approaching 1, as $N \to \infty$, the random graph has no connected components having more than μ nodes ([Bol98] ch. 4).

We shall show that there is a Nash equilibrium such that each player uses only the information contained in her connected component. Consider a player p_i and the connected component in the graph which contains i denoted by $\bar{G}_i$. Consider also the game with $\left|\bar{G}_i\right|$ players among the players of $\bar{G}_i$ denoted by $g^{\bar{G}_i}$, assuming that the actions of the players are restricted to belong to $[0, L]$.

Each of the games $g^{\bar{G}_i}$, due to the convexity of the function $g(\cdot)$, satisfy the conditions of Theorem 1 of [Ros65]. Thus, it has a Nash equilibrium. This equilibrium, due to (4.19), is also a Nash equilibrium for the corresponding game with unrestricted strategies. Thus, with probability approaching 1 as $N \to \infty$, there exists a Nash equilibrium of the original game such that each player uses only the knowledge of her connected component. Thus, $C_a(m) \leq \mu$ and the game is asymptotically simple. □

Remark 24. *The upper bound for the complexity function* $C_a(m)$ *increases as* μ *increases and* $1/N^{\mu/(\mu-1)}$ *approaches* $1/N$. □

The following proposition deals with the intermediate connectivity case.

Proposition 11. *If* $c_N \in \omega\left(\frac{1}{N}\right)$ *and* $c_N \in o\left(\frac{\ln N}{N}\right)$ *then the ensemble of games is asymptotically complex.*

Proof: Due to the fact that $c_N \in \omega\left(1/N\right)$, the maximum number of edges adjacent to a node grows unbounded with N. Thus, it suffices to show that $\bar{n}(m) \geq 1$, for some integer m. To contradict, assume that $\bar{n}(m) = 0$.

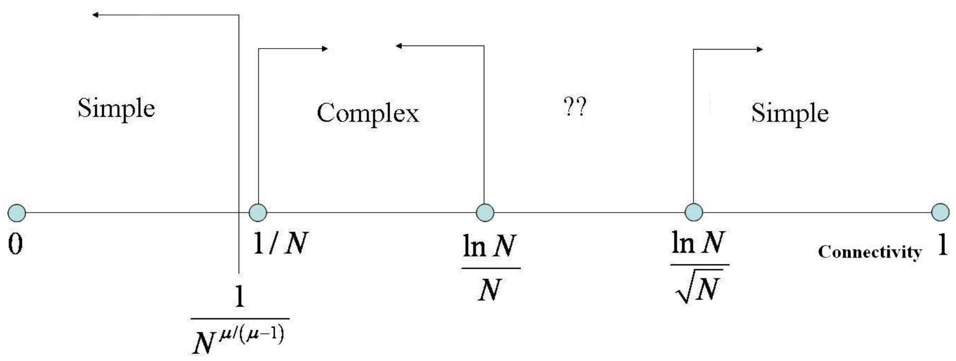


Figure 4.2: The complexity of the ensemble of games for the various connectivity intervals.

Due to the fact that $c_N \in o(\ln N/N)$ there exists an isolated node with probability approaching 1 as $N \to \infty$ [Bol98]. In fact, the expected number of such nodes grows unbounded with N.

For such a node the optimal cost is 0. The function $g(\cdot)$ is strictly convex and $g(0) = g'(0) = 0$. Thus, for any $\varepsilon > 0$ there exists a $\delta = \delta(\varepsilon) > 0$ with $\delta(\varepsilon) \to 0$ as $\varepsilon \to 0$, such that $g(z) < \varepsilon$ implies $|z| < \delta$.

The assumption $\bar{n}(m) = 0$ implies that $I^i = \{\theta_i\}$ and that:

$$|\gamma(\theta) - \theta| < \delta, \tag{4.23}$$

if $2^{-m} < \varepsilon$.

It is not difficult to see that with probability approaching 1 as $N \to \infty$ there exists a player p_i such that $\theta_i < L/8$ and $\sum_{j \in \mathcal{N}_i(G)} \gamma(\theta_j) > L/4$. For such a player if m is large enough, a strategy satisfying (4.23) is not 2^{-m} optimal. This fact contradicts $\bar{n}(m) = 0$. □

The results proved are summarized in the following corollary. The situation is depicted graphically in Figure 4.2

Corollary 4. *If* $c_N \in o\left(\frac{1}{N^{\mu}(\mu-1)}\right)$ *for some integer* μ *or if* $c_N \in \omega\left(\frac{\ln N}{\sqrt{N}}\right)$ *the ensemble of games is asymptotically simple. If* $c_N \in \omega\left(\frac{1}{N}\right)$ *and* $c_N \in o\left(\frac{\ln N}{N}\right)$ *then the ensemble of games is asymptotically complex.* □

4.4.2 Static Games on Small World Networks

This example studies a game involving players interacting on a small world network. The small world network model was first introduced in [WS98]. However, for simplicity we use the model described in [NW99]

Let us first describe a stochastic model for the interaction graph. There are N players placed on a circle. Initially each one of them is connected with $2k_N$ players closer to her. Then, each one of the remaining edges appears in the graph with probability p_N independently of the existence of the other edges. It is required that $N \gg k = k_N \gg \ln N \gg 1$ and $p_N = 2k_N/(N - 2k_N)$.

Let us denote by G the network. The cost function for each player is given by:

$$J^i = (u^i - \theta^i)^2 + \left(u^i - \frac{1}{4k} \sum_{j \in \mathcal{N}_i(G)} u^j \right)^2, \tag{4.24}$$

where $\mathcal{N}_i(G)$ consists of the players in the neighborhood of player p_i.

We assume that the types θ^i are random and they are given by:

$$\theta^i = \left[1 + sin \left(\frac{2\pi i}{N} \right) \right] r^i, \tag{4.25}$$

where r^i are i.i.d. random variables uniformly distributed in $[0, 1]$.

We first derive a set of strategies which will be shown to be $\varepsilon-$ PAN, for large N, under specified conditions. In order to do so, let us consider two quantities: $\bar{u}^i$ and $\bar{u}$. The first stands for the local mean of the actions of the players that are close to player i on the circle. The second quantity stands for the mean of the actions of all the players. The following approximation will be shown to be valid:

$$\bar{u}^i \simeq \frac{1}{2k} \sum_{\substack{|i-j| \leq k \\ j \neq i}} u^j \text{ and } \bar{u} \simeq \frac{1}{2k} \sum_{\substack{j \in \mathcal{N}_i(G) \\ |i-j| > k}} u^j. \tag{4.26}$$

Using these approximations the cost functions satisfy:

$$J^i \simeq (u^i - \theta^i)^2 + \left(u^i - \frac{\bar{u}^i + \bar{u}}{2} \right)^2. \tag{4.27}$$

The approximate best response map is given by:

$$u^i = \frac{2\theta^i + \bar{u}^i + \bar{u}}{4} \tag{4.28}$$

Taking the local mean around i we obtain:

$$\bar{u}^i = \frac{2\bar{\theta}^i}{3} + \frac{\bar{u}}{3}, \tag{4.29}$$

where

$$\bar{\theta}^i = \frac{1 + sin \left(\frac{2\pi i}{N} \right)}{2}. \tag{4.30}$$

Taking the mean of $\bar{u}^i$ we obtain $\bar{u} = \bar{\theta} = 1/2$. Thus, the desired strategy is given by:

$$u^i = \frac{\theta^i}{2} + \frac{sin \left(\frac{2\pi i}{N} \right)}{12} + \frac{1}{4}. \tag{4.31}$$

The strategy given by (4.31) depends on player's type and on *local* as well as emphglobal averages.

Proposition 12. *If* $p_N \in \omega\left(\frac{\ln(N)}{\sqrt{N}}\right)$ *and* $p_N \in o(1)$ *then the set of strategies given by* (4.31) *are* ε*-PAN for large* N*. Furthermore, the ensemble of games described is asymptotically simple.*

Proof. It is sufficient to show that for any positive constant δ, it holds:

$$P\left(\left|\bar{u}^i - \frac{1}{2k}\sum_{\substack{|i-j|\le k\\ j\neq i}} u^j\right| < \delta,\ \text{for any } i = 1,\ldots,N\right) > 1-\delta, \tag{4.32}$$

$$P\left(\left|\bar{u} - \frac{1}{2k}\sum_{\substack{j\in\mathcal{N}_i(G)\\ |i-j|>k}} u^j\right| < \delta,\ \text{for any } i = 1,\ldots,N\right) > 1-\delta, \tag{4.33}$$

for large N.

It holds:

$$\left|\bar{u}^i - \frac{1}{2k}\sum_{\substack{|i-j|\le k\\ j\neq i}} u^j\right| \le \frac{1}{2k}\left|\sum_{\substack{|i-j|\le k\\ j\neq i}} \frac{2r^j-1}{2}\right| + \frac{7\pi k}{6N}.$$

Theorem 4 and Lemma 9 imply that (4.67) holds for large N.

For each player i, we divide set of players j such that $|j-i|>k$ to $\bar{\lambda}$ equal circle sectors $\Lambda_1,\ldots,\Lambda_{\bar{\lambda}}$. For each sector l we choose a representative player λ_l. It holds:

$$\left|\bar{u} - \frac{1}{2k}\sum_{\substack{j\in\mathcal{N}_i(G)\\ |i-j|>k}} u^j\right| \le \left|\bar{u} - \frac{1}{\bar{\lambda}}\sum_{l=1}^{\bar{\lambda}} \bar{u}^{\lambda_l}\right| + \frac{1}{\bar{\lambda}}\sum_{l=1}^{\bar{\lambda}}\left|\bar{u}^{\lambda_l} - \frac{\bar{\lambda}}{2k}\sum_{j\in\Lambda_l} u^j d_{ij}\right|, \tag{4.34}$$

where $d_{ij}=1$ if there is an edge among i and j and 0 otherwise.The first term is less than $\delta/2$ with probability 1, assuming that $\bar{\lambda}$ is large. For the second term it holds:

$$\left|\bar{u}^{\lambda_l} - \frac{\bar{\lambda}}{2k}\sum_{j\in\Lambda_l} u^j d_{ij}\right| = \frac{1}{\frac{N-2k}{\bar{\lambda}}\frac{2k}{N-2k}}\left|\sum_{j\in\Lambda_l}(p_k\bar{u}^{\lambda_l} - u^j d_{ij})\right|$$

$$\le \frac{1}{\frac{N-2k}{\bar{\lambda}}\frac{2k}{N-2k}}\left|\sum_{j\in\Lambda_l}\left[p_k\left(\frac{1}{2}+\frac{\sin\phi_l}{3}\right) - \left(\frac{r_j}{2}+\frac{(1+6r_j)\sin\phi_l}{12}+\frac{1}{4}\right)d_{ij}\right]\right| +$$
$$+\frac{1}{\frac{N-2k}{\bar{\lambda}}\frac{2k}{N-2k}}\left|\sum_{j\in\Lambda_l} d_{ij}\frac{1}{\bar{\lambda}}\right|$$

where $\phi_l = \frac{2\pi\lambda_l}{N}$. Theorem 4 can be applied to both terms of the last inequality. Then Lemma 9 completes the proof. □

Remark 25. *Several extensions are possible. For example, different network model, the LQ dynamic counterpart or static non-quadratic games can be studied using the same tools.* □

4.4.3 LQ games on Random Graphs

This section describes an opinion dynamics game, involving players having some amount of stubbornness, i.e. the players tend to insist to their initial (intrinsic opinion) (ex. [GS13]). A large number N of players interact on a random graph $G = (V, E)$ having a connection probability c_N, i.e. each link have a probability to exist equal to c_N independent of the existence of the other links. The state variable of player p_i, x^i_k represents her opinion at time step k. The type θ_i is her initial (intrinsic) opinion and $x^i_0 = \theta_i$. The random variables θ_i are i.i.d. with uniform distribution in $[0, L]$.

Each player has the ability to influence her own opinion in order to come closer to the mean value of her neighbors or closer to her initial opinion. Furthermore, the state variables are influenced by random disturbances. The dynamics is given by:

$$x^i_{k+1} = x^i_k + u^i_k + w^i_k, \tag{4.35}$$

where u^i_k is the control variable of player p_i and w^i_k are zero mean i.i.d. Gaussian random variables with variance σ^2.

The cost functions are given by:

$$J^i = E\left\{\sum_{k=0}^{\infty}\rho^k\left[\left(x^i_k - \frac{1}{|\mathcal{N}_i(G)|}\sum_{j\in N_i(G)} x^j_k\right)^2 + \right.\right.$$
$$\left.\left. +s\left(x^i_k - \theta_i\right)^2 + r(u^i_k)^2\right]\middle|\mathcal{F}^s\right\}, \tag{4.36}$$

where $\rho \in (0, 1)$ is the discount factor, $s \geq 0$ the amount of stubbornness and r a positive constant.

We then study when it is possible to have an ε-PAN set of strategies assuming that each player has only statistical information. A set of approximate optimal control problems is thus stated. For player p_i the approximate optimal control problem is:
Minimize:

$$J^{i,a} = E\left\{\sum_{k=0}^{\infty}\rho^k\left[\left(x^i_k - \bar{\theta}\right)^2 + s\left(x^i_k - \theta_i\right)^2 + r(u^i_k)^2\right]\right\}, \tag{4.37}$$

where $\bar{\theta} = E[\theta_j] = L/2$,
Subject to (4.35).

Using the change of variables:

$$\tilde{x}^i_k = x^i_k - \theta^{i,f}, \tag{4.38}$$

where $\theta^{i,f} = \frac{\bar{\theta}+s\theta_i}{1+s}$, the optimal control problem becomes the following LQ problem:
Minimize:

$$J^{i,a} = E\left\{\sum_{k=0}^{\infty}\rho^k\left[(1+s)(\tilde{x}_k^i)^2 + r(u_k^i)^2\right]\right\} + \\ + \frac{s(\bar{\theta}-\theta_i)^2}{(1+s)(1-\rho)}, \tag{4.39}$$

Subject to:

$$\tilde{x}_{k+1}^i = \tilde{x}_k^i + u_k^i + w_k^i. \tag{4.40}$$

The control law which minimizes the approximate optimal control problem is given by:

$$u_k^i = -\rho\frac{K}{\rho K + r}\tilde{x}_k^i, \tag{4.41}$$

where K is the positive solution of the Riccati equation:

$$K = \rho K - \frac{\rho^2 K^2}{\rho K + r} + (1+s). \tag{4.42}$$

The closed loop dynamic equation for player p_i is given by:

$$\tilde{x}_{k+1}^i = a\tilde{x}_k^i + w_k^i, \tag{4.43}$$

where $a = 1 - \rho K/(\rho K + r)$ and $a \in (0,1)$.

The following proposition identifies a class of games where the set of strategies described by (4.41) is ε-PAN.

Proposition 13. *Assume that* $c_N \in \omega(\frac{\ln N}{\sqrt{N}})$. *Then:*

(i) The set of strategies given by (4.41) *constitute an* ε*-PAN set of strategies, for large* N.

(ii) It holds $C_a(m) = 1$ *and the sequence of the ensembles of games is asymptotically simple.*

Proof: See Appendix 4.8.2 □

If the stubbornness of the players is zero, then the long time average of the opinions of the players reach a consensus.

Proposition 14. *If* $s = 0$ *and the players use the strategies given by* (4.41) *then:*

$$P(\lim_{T\to\infty}\frac{1}{T}\sum_{k=0}^{T}x_k^i = \bar{\theta}) = 1. \tag{4.44}$$

Proof: See Appendix 4.8.3. □

4.5 Static Games on Organized Structures

4.5.1 Quadratic Games on Lattices

In this section, we study quadratic games on μ-dimensional lattices. It is shown that the ANIC function is polynomial with degree equal to μ.

There are $N = N_0^\mu$ players on a μ-dimensional lattice (hypercube). For each node i a set of coordinates $(c_1, \ldots, c_\mu)$ is introduced, indicating the place of the node in the lattice. Each coordinate satisfies $c_\nu \in \{1, \ldots, N_0\}$. Each player interacts with her immediate neighbors, i.e. a player with coordinates $(c_1, \ldots, c_\mu)$ interacts with every player with coordinates $(c_1, \ldots, c_\nu \pm 1, \ldots, c_\mu)$, for $\nu = 1, \ldots, \mu$. For the nodes on the faces of the hypercube, the convention $N_0 + 1 \equiv 1$ is used.

Each player p_i has a type $\theta_i \in [-L, L]$. The cost of player p_i is given by:

$$J^i = a\left(u^i - \frac{1}{2^\mu}\sum_{j\in\mathcal{N}_i(G)} u^j\right)^2 + \left(u^i - \theta_i\right)^2, \tag{4.45}$$

where a is a positive constant.

In order to describe the ensemble of games it remains to determine a probability structure on the types of the players. We assume that θ_i are i.i.d. random variables with uniform distribution.

Let us consider the following iterative scheme:

$$\begin{aligned} z^i(t+1) &= \frac{a}{(a+1)2^\mu}\sum_{j\in\mathcal{N}_i^n(G)} z^j(t) + \frac{1}{a+1}\theta_i, \\ z^i(0) &= 0. \end{aligned} \tag{4.46}$$

Equation (4.46) corresponds to the best response of player p_i if the other players use $u^j = z^j$.

It will be shown that the ANIC of the ensemble is at most polynomial using the fact that the mapping:

$$T : z(t) \mapsto z(t+1), \tag{4.47}$$

where $z(t) = [z^1(t), \ldots, z^N(t)]$, is a contraction (Lipschitz with a constant less than 1).

Proposition 15. *(i) For any $\varepsilon > 0$, there exists an $n \in \mathbb{N}$ such that the set of strategies $u^i = z^i(n)$ constitute an ε-PAN equilibrium.*

(ii) The ensemble of games has an ANIC satisfying $C_a(m) \in O(m^\mu)$.

Proof: (i) It holds:

$$\begin{aligned} &J_i(z^i(t), z^{-i}(t)) - \min_u\{J_i(u, z^{-i}(t))\} = \\ &\qquad = (a+1)(z^i(t) - z^i(t+1))^2 \end{aligned} \tag{4.48}$$

The mapping $T : (\mathbb{R}^N, \|\cdot\|_\infty) \to (\mathbb{R}^N, \|\cdot\|_\infty)$ is contractive, with a Lipschitz constant $a/(a+1)$. Hence,

$$\|z(t+1) - z(t)\| \leq L \left(\frac{a}{a+1}\right)^t.$$

Therefore,

$$(a+1)L^2 \left(\frac{a}{a+1}\right)^{2n} < \varepsilon, \tag{4.49}$$

implies that $u^i = z^i(n)$ is ε-PAN equilibrium.

(ii) Using (4.49) with $\varepsilon = 2^{-m}$ we have:

$$\bar{n}(m) \leq \frac{m}{2\log_2\left(\frac{a+1}{a}\right)} + \frac{2\log_2 L + \log_2(a+1)}{2\log_2\left(\frac{a+1}{a}\right)}.$$

Furthermore, $|\mathcal{N}_i^{\bar{n}}| < (2\bar{n})^\mu$. Thus, $C_a(m) \in O(m^\mu)$. □

A polynomial lower bound can also be derived.

Proposition 16. *The asymptotic complexity function $C_a(m)$ satisfies $C_a(m) \in \Omega(m^\mu)$.*

Proof: Consider a game in the ensemble and a set of actions $(u^i)_i$. Due to the contractivity of T, there exists a unique Nash equilibrium of the game. Denote by $(u^{i,N})_i$ that equilibrium.

Let us use (4.46) with $z^i(0) = u^i$. Due to (4.48), if $(u^i)_i$ is an ε-Nash equilibrium then:

$$\|z(1) - z(0)\|_\infty < \sqrt{\frac{\varepsilon}{a+1}}.$$

Thus, due to contractivity of T, it holds:

$$\max_i\{|u^i - u^{i,N}|\} < \sqrt{(a+1)\varepsilon}.$$

The unique Nash equilibrium can be expressed as:

$$u^{i,N} = \sum_{c'_1,\ldots,c'_\mu \in \mathbb{Z}} b^{c'_1,\ldots,c'_\mu} \theta^{c_1+c'_1,\ldots,c'_\mu+c'_\mu},$$

where $c_1, \ldots, c_\mu$ are the coordinates corresponding to player p_i. It is not difficult to show that the constants b satisfy:

$$b^{c'_1,\ldots,c'_\mu} > \frac{1}{a+1}\lambda^{|c'_1|+\cdots+|c'_\mu|},$$

where $\lambda = \frac{a}{(a+1)2^\mu}$.

Consider now a set of strategies in the form $u^i = \gamma(\bar{I}^{i,n-1})$. Consider a player p_i with coordinates $c_1, \ldots, c_\mu$. Then, with a probability larger than $1/2$ the player p_j with coordinates $c_1+n, \ldots, c_\mu$ has a type $|\theta_j| > L/2$. Thus, with probability larger than $1/4$, it holds $|u^i - u^{i,N}| > \lambda^n/(a+1)$.

Therefore, using an information neighborhood of order $n-1$, an $\lambda^{2n}/(a+1)^3$ - PAN equilibrium is not attainable. Hence:

$$\bar{n}(m) > m\frac{\ln 2}{-2\ln\lambda} - \frac{3\ln(a+1)}{-2\ln\lambda}.$$

Thus, $C_a(m) \in \Omega(m^\mu)$. □

Corollary 5. *The ANIC function satisfies* $C_a(m) \in \Theta(m^\mu)$.

Remark 26. *The properties proved do not depend on the assumption that the hypercube has the same length* N_0 *in all the dimensions nor on the nonlocal topological properties of the lattice.* □

The result about the upper bound of the ANIC function can be generalized to ensembles of games on graphs with known maximum degree, using exactly the same arguments.

Proposition 17. *Consider an ensemble of games with cost function* (4.45) *and interaction graphs which with probability one have maximum degree less than* μ. *Then,* $C_a(\mu) \in O(m^\mu)$.

Proof: The proof uses essentially the same ideas as Proposition 15, i.e. the contractivity of the best response maps is used. □

The bound of Proposition 15 is much sharper than the bound of Proposition 17, when applied to a Lattice, due to the fact that the Lattice is highly clustered.

4.5.2 A non-Quadratic Game on a Ring

In this subsection we study an example of an ensemble of games, where a PAN equilibrium can be obtained, using some form of cooperation among the players. The best response maps in this example are chaotic. Let us note that the use of chaotic maps is not unusual in the modeling of erratic behavior in economics (ex. [BD81]).

There are N players interacting on a ring. The type of each player has the form $\theta_i = (\xi_i, i)$ and ξ_i has two possible values 1 and 2. The cost function of each player, except player p_0, is given by:

$$J_i = (u^i - f^{\xi_i}(u^{i-1}))^2. \tag{4.50}$$

The functions f^1 and f^2 have the form:

$$f^i(z) = \begin{cases} z/a_i & \text{if } 0 \le z < a_i \\ \frac{1-z}{1-a_i} & \text{if } a_i < z \le 1 \end{cases},$$

where $a_1 = 1/3$ and $a_2 = 2/3$. The functions f^1 and f^2 are variants of the tent map [ASY96] and their plots are shown in figure 4.3. For the player p_0 the cost is:

$$J_0 = (u^0 - f^{\xi_0}(u^N))^2 + (u_0 - 1)^2. \tag{4.51}$$

In order to describe the ensemble it remains to determine a stochastic structure on ξ_i. We assume that ξ_i are i.i.d. random variables, taking values 1 and 2 with equal probabilities.

The best response map for player p_i is given by $u^i = f^{\xi_i}(u^{i-1})$. This map is not contractive. In fact it is chaotic. However, the following proposition shows that ANIC is at most linear.

Proposition 18. *The ANIC function of the ensemble of games satisfies* $C_a(m) \in O(m)$.

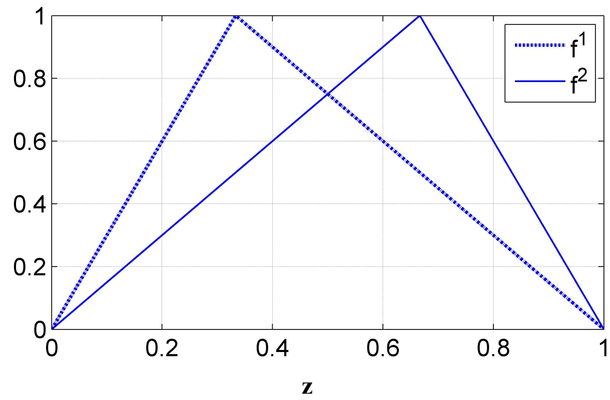


Figure 4.3: The plot of f^1 and f^2

Proof: The proof is constructive. It is not difficult to show that, for any positive integer μ, any finite sequence $s_1, \dots, s_\mu$ $s_j \in \{1, 2\}$ and any $z \in [0, 1]$ there exists a $\bar{z} \in [0, 1]$ such that $|z - \bar{z}| \leq \frac{1}{2}\left(\frac{2}{3}\right)^\mu$ and $f^{s_1} \circ \dots \circ f^{s_\mu}(\bar{z}) = 0.5$. Let us denote by $h_\mu(s_0, \dots, s_\mu, z)$ the minimal such point $\bar{z} \in [0, 1]$.

In order to construct the strategies of the players we consider distinct cases. For a player p_i, $i \notin \{0, (\mu - 1)[N/\mu], \mu[N/\mu]\}$ such that $i \equiv 0(mod\mu)$:

$$u^i = h_\mu(\xi_i, \dots, \xi_{i+\mu-1}, 0.5). \tag{4.52}$$

For the player $p_{(\mu-1)[N/\mu]}$:

$$u^{(\mu-1)[N/\mu]} = h_{N-(\mu-1)[N/\mu]}(\xi_i, \dots, \xi_{i+\mu-1}, 0.5). \tag{4.53}$$

For any player p_i, such that $i \not\equiv 0(mod\mu), i < (\mu - 1)[N/\mu]$:

$$u^i = f^{\xi_i}(f^{\xi_{i-1}}(\dots f^{\mu[i/\mu]}(0.5))). \tag{4.54}$$

For the players $p_{(\mu-1)[N/\mu]+1}, \dots, p_{N-1}$:

$$u^i = f^{\xi_i}(f^{\xi_{i-1}}(\dots f^{(\mu-1)[N/\mu]}(0.5))). \tag{4.55}$$

For the player p_0:

$$u^i = h_\mu(\xi_0, \dots, \xi_{\mu-1}, 0.75). \tag{4.56}$$

It is not difficult to see that the set of strategies (4.52)- (4.56) constitute an $\frac{1}{4}\left(\frac{4}{9}\right)^\mu$-PAN set of strategies. Thus, $C_a(m) \leq m$ and the proof is complete. □

Remark 27. *The game has a lot of Nash equilibria. In a Nash equilibrium, the following equations hold:*

$$u^0 = \frac{1}{2} + \frac{f^{\xi_0}(f^{\xi_{N-1}}(\dots(f^{\xi_1}(u^0))))}{2} \tag{4.57}$$

$$u^i = f^{\xi_i}(u^{u-1}). \tag{4.58}$$

The equation (4.57) *has approximately N/2 solutions. The strategies of the players for any Nash equilibrium are given by* (4.58). *Thus, a full knowledge of the information is needed. The ε-PAN set of strategies, described in the proof of Proposition 18 is far from any Nash equilibrium.* □

Remark 28. *The ε-PAN set of strategies is in some sense cooperative. Particularly, a player p_i, $i = l\mu$ can improve her performance based on her own information. However, these agents help the others have a predictable best response with local information only and thus to behave optimally. If such a player changes her action to the optimal response, then she would expect that the other players would also use their best responses. Due to the chaoticity of the maps f^1, f^2, we would expect that for a long amount of time the best responses would not converge. This would make the situation worse for these players.* □

4.6 LQ Games on a Ring

4.6.1 A finite horizon LQ game on a Ring

This section studies an example of a game with a known number of players $N = \nu$ lying on a ring having interactions only with their nearest neighbors. Specifically, there is a set of players $p_1, \ldots, p_N$ and each player p_i has a connection with p_{i-1} and p_{i+1}, where the convention $N + l \equiv l$ is used. Each agent p_i has a type $\theta_i \in [0, 1]$. The random variables θ_i are independent and distributed uniformly on $[0, 1]$.

The dynamics of the state vector of the player i is described by the following equation:

$$x_{k+1}^i = ax_k^i + u_k^i + w_k^i, \tag{4.59}$$

where x^i is the scalar state variable of the agent i, u_k^i the control variable of the agent i and w_k^i are zero mean i.i.d. random variables with finite second moments. The initial conditions are given by $x_0^i = w_{-1}^i$.

The cost function for the player i is given by:

$$J^i = E\left\{\left(z_T^i\right)^2 + \sum_{k=0}^{T-1}\left[\left(z_k^i\right)^2 + r_i\left(u_k^i\right)^2\right]\right\}, \tag{4.60}$$

where $r_i = (1 + \theta_i)/2$ and

$$z_k^i = x_k^i - \lambda\left(x_k^{i-1} + x_k^{i+1}\right). \tag{4.61}$$

Thus, a Linear Quadratic game with coupling only through the cost functions is considered.

Remark 29. *The random variables θ_i are defined on $(\mathcal{E}, \mathcal{A}, Q)$ and the disturbances w_k^i on $(\Omega, \mathcal{F}, P)$.*

It will be shown that the ANIC is at most linear under some specified conditions. To do so, we shall use simultaneous dynamic programming from the time step $k = T-1$, backwards to $k = 0$. Consider the last step of the simultaneous dynamic programming. The cost functions are given by:

$$J^i_{T-1} = E\left[(z^i_{T-1})^2 + \left(x^i_T - \lambda\left(x^{i-1}_T + x^{i+1}_T\right)\right)^2 \right. \\ \left. + r_i\left(u^i_{T-1}\right)^2 |x_{T-1}\right]. \tag{4.62}$$

It is not difficult to see that the equilibrium condition is given by:

$$u^i_{T-1} = W^i_{T-1}((u^j_{T-1})^N_{j=1}), \tag{4.63}$$

where the mapping $W^i_{T-1}((u^j_{T-1})^N_{j=1})$ is given by:

$$W^i_{T-1}((u^j_{T-1})^N_{j=1}) = \frac{\lambda}{1+r_i}\left[u^{i-1}_{T-1} + u^{i-1}_{T-1}\right] + \\ + \frac{a}{1+r_i}\left[-x^i_{T-1} + \lambda(x^{i-1}_{T-1} + x^{i+1}_{T-1})\right]. \tag{4.64}$$

The proof of the following Proposition 19 is based on the contractivity of the following mapping:

$$W_{T-1}((u^j_{T-1})^N_{j=1}) = \left(W^i_{T-1}((u^j_{T-1})^N_{j=1})\right)^N_{i=1}. \tag{4.65}$$

Analogous mappings are then defined for the other time steps $T-2, \ldots, 0$.

Proposition 19. *For small coupling constant λ, the ANIC of the ensemble of games described by* (4.59) *and* (4.60) *is at most linear, i.e. $C_a(m) \in O(m)$.*

Proof. For simplicity reasons the proof is given for $T = 2$. The proof starts at $k = 1$ and then moves backwards to $k = 0$. If $|\lambda| < 1/2$, then the mapping given by (4.65) is contractive for the infinity norm. Consider the feedback strategies obtained after m iterations of (4.65) with zero initial strategies. These strategies have the form:

$$u^{i,m}_1 = \sum_{l=-m}^{m} k^{i,l,m}_1 x^{i+l}_1. \tag{4.66}$$

Then the equation (4.65) is also a contraction in the space of the vectors of feedback gains with the infinity norm, i.e. the mapping $(k^{i,l,m}_1)^N_{i=1} \mapsto (k^{i,l,m+1}_1)^N_{i=1}$ is a contraction. Therefore:

$$\left\|(k^{i,l,m+1}_1)^N_{i=1} - (k^{i,l,m}_1)^N_{i=1}\right\|_\infty \leq \frac{(2\lambda+1)a}{1.5}\left(\frac{2\lambda}{1.5}\right)^m \tag{4.67}$$

Before going back to the step $k = 0$, let us compute the form of the cost functions J^i_1 when in the last step the strategies with feedback gains $(k^{i,l,m}_1)^N_{i=1}$ are applied. Equation (4.62) implies:

$$J^i_1 = \left(\sum_{l=-m}^{m} \xi^{i,l}_1 x^{i+l}_1\right)^2 + \left(\sum_{l=-m}^{m} k^{i,l,m}_1 x^{i+l}_1\right)^2 + (z^i_1)^2 + C_i,$$

where:

$$\xi_1^{i,l} = a\delta_{l,0} - \lambda a(\delta_{l,1}+\delta_{l,-1}) + k_1^{i,l,m} - \\ - \lambda(k_1^{i+1,l-1,m} + k_1^{i-1,l+1,m}). \tag{4.68}$$

It is not difficult to show that it holds:

$$J_1^i = \sum_{l_1,l_2} q_1^{i,l_1,l_2} x_1^{i+l_1} x_1^{i+l_2}, \tag{4.69}$$

where $q_1^{i,l_1,l_2} = q_1^{i,l_2,l_1}$, $|q_1^{i,l_1,l_2}| < M\beta^{|l_1|+|l_2|}$ and $\beta > 0$. Furthermore, $\beta \to 0$ as $\lambda \to 0$. Let us now go back one step to $k = 0$ and assume that the players at time step $k = 1$ will follow the strategies with feedback gains given by $k_1^{i,l,m}$. The cost functions have the form:

$$J_0^i = (z_0^i)^2 + C_i + r_i(u_0^i)^2 + \sum_{l_1,l_2} q_1^{i,l_1,l_2} x_1^{i+l_1} x_1^{i+l_2} \tag{4.70}$$

The equilibrium condition is, thus, given by:

$$u_0^i = -\frac{1}{(q_1^{i,0,0} + r_i)} \left[\sum_{l_2} a q_1^{i,0,l_2} x^{i+l_2} + \sum_{l_2 \neq 0} q_1^{i,0,l_2} u_0^{i+l_2} \right],$$

and the mapping in the feedback gains by:

$$k_0^{i,l,m'+1} = -\frac{1}{(q_1^{i,0,0} + r_i)} \left[a q_1^{i,0,l} + \sum_{l_2 \neq 0} q_1^{i,0,l_2} k_1^{i+l_2,l-l_2,m'} \right] \tag{4.71}$$

For small β, the mapping (4.71) is contractive with Lipschitz constant $2M\beta/(1-\beta)$. Therefore:

$$\left\| (k_0^{i,l,m'+1})_{i=1}^N - (k_0^{i,l,m'})_{i=1}^N \right\|_\infty \leq aM \left(\frac{2M\beta}{1-\beta} \right)^{m'}. \tag{4.72}$$

To complete the proof, let us introduce some quantities. Let $\bar{J}_1^{i,m}(x_1)$ be the minimum cost to go for player i at time step $k = 1$, assuming that the other players use the strategies given by (4.66), $J_1^{i,m}(x_1)$ be the cost to go if all the players use the strategies given by (4.66) and:

$$J_0^{i,m}(x_0, u_0^i) = (z_0^i)^2 + (u_0^i)^2 + E[J_1^{i,m}(x_1)|x(0)], \\ \bar{J}_0^{i,m}(x_0, u_0^i) = (z_0^i)^2 + (u_0^i)^2 + E[\bar{J}_1^{i,m}(x_1)|x(0)],$$

where we assume that the other players use the strategies given by (4.66).

The proof is based on the following facts:

Fact 1. *Let* $\gamma > (2\lambda/1.5)^2$. *For large* m *it holds:*

$$J^{i,m}(x_1) \leq \bar{J}^{i,m}(x_1) + \gamma^m(1 + \|x_1\|^2_{\infty,i,m}),$$

where $\|x_1\|_{\infty,i,m} = max\{|x_1^j| : j = i-m, \ldots, i+m\}$.

The proof of Fact 1 is immediate from equations (4.62) *and* (4.67).

Fact 2. *Let* $\gamma > (2\lambda/1.5)^2$. *For large* m *it holds:*

$$\min_{u_0^i} J^{i,m}(x_0, u_0^i) \leq \min_{u_0^i} \bar{J}^{i,m}(x_0, u_0^i) + \gamma^m(1 + \|x_0\|^2_{\infty,i,m}).$$

To prove Fact 2, let us observe that:

$$J_0^{i,m}(x_0, u_0^i) = (z_0^i)^2 + (u_0^i)^2 + J_1^{i,m}(f(x_0, u_0^i)) + c_0,$$

where $f(x_0, u_0^i))$ *is the expected value of* x_1 *given* x_0 *if the player* i *uses* u_0^i *and the other players use the strategies given by* (4.66). *Furthermore, it holds:*

$$\bar{J}_0^{i,m}(x_0, u_0^i) = (z_0^i)^2 + (u_0^i)^2 + \bar{J}_1^{i,m}(f(x_0, u_0^i)) + \bar{c}_0.$$

Fact 1 implies that $|c_0 - \bar{c}_0| < \gamma^m$, *for large* m. *Denoting by* $v(x_0)$ *the value of* u_0^i *that minimizes* $\bar{J}_0^{i,m}(x_0, u_0^i)$ *we have:*

$$\begin{aligned}\min_{u_0^i} J_0^{i,m}(x_0, u_0^i) &\leq J_0^{i,m}(x_0, v)\\ &\leq \min_{u_0^i} \bar{J}_0^{i,m}(x_0, u_0^i) + 2\gamma^m(1 + \|x_0\|^2_{\infty,i,m}).\end{aligned}$$

And using a small abuse of notation (choosing a smaller value for γ*) we conclude to Fact 2.*

Fact 3. *Let* $m > (2M\beta/(1-\beta))^2$. *For large* m *it holds:*

$$\begin{aligned}J_0^{i,m}(x_0, \sum_{l=-m}^{m} k_0^{i,l,m} x_0^{i+l}) &\leq\\ &\leq \min_{u_0^i} J_0^{i,m}(x_0, u_0^i) + \gamma^m(1 + \|x_0\|^2_{\infty,i,m}).\end{aligned}$$

The proof of Fact 3 is immediate from (4.72).

For a small value of λ, there exist a constant $\tilde{\gamma} > max\{(2M\beta/(1-\beta))^2, (2\lambda/1.5)^2\}$ and $\tilde{\gamma} < 1$ such that Facts 2 and 3 apply with $\gamma = \tilde{\gamma}$. Thus, if $\bar{\gamma} > \tilde{\gamma}$, for large m it holds:

$$E\left[J_0^{i,m}(x_0, \sum_{l=-m}^{m} k_0^{i,l,m} x_0^{i+l})\right] \leq E\left[\min_{u_0^i} J_0^{i,m}(x_0, u_0^i)\right] + \bar{\gamma}^m$$

Thus, the strategies given by the feedback gains $k_0^{\cdot,\cdot,\cdot}$ and $k_1^{\cdot,\cdot,\cdot}$ are $\bar{\gamma}^m$ - fine for large m. Thus, the ANIC function is at most linear. □

4.6.2 A Numerical Study of a LQ Game on a Ring

In this section we study numerically LQ games on large rings. We assume that there are N players placed on a ring, such that each player p_i interacts with the players p_{i+1} and p_{i-1}. The convention $N + 1 \equiv 1$ is used as before. Each of the players has her own dynamic equation given by:

$$x_{k+1}^i = x_k^i + \lambda(x_k^{i-1} + x_k^{i+1}) + u_k^i, \tag{4.73}$$

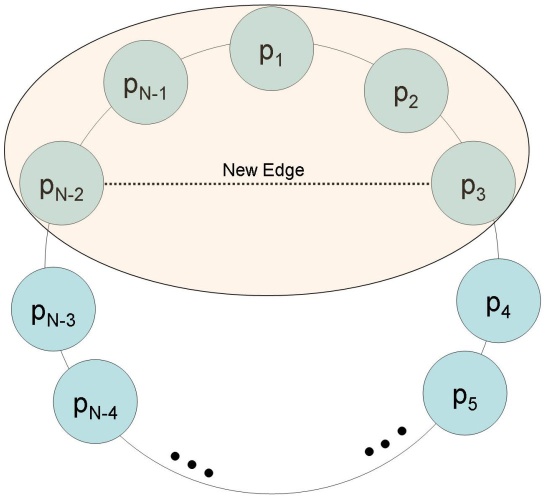


Figure 4.4: Player p_1 considers only the players $p_2, p_3, p_{N-1}, p_{N-2}$ and himself and assumes the existence of a new edge between players p_3 and p_{N-2}.

where λ is a constant describing the coupling through the state equation and x_0^i are zero mean i.i.d. random variables with variance equal to 1. The cost functions are given by:

$$J_i = \sum_{k=1}^{\infty} \left[(x_k^i)^2 + s \left(x_k^i - \frac{x_k^{i+1} + x_k^{i-1}}{2} \right)^2 + r(u_k^i)^2 \right], \tag{4.74}$$

where s denotes the coupling through the cost functions.

A very simple technique is used to obtain approximate equilibrium policies using small information neighborhoods. Assume that the players have information neighborhoods of order n. The approximate equilibrium policy is derived using the following procedure:

Step 1: Each player p_i considers a reduced ring game with $2n+1$ players, i.e. the players $p_{i-n}, \dots p_{i+n}$. That is, she assumes the existence of an edge between players p_{i-n} and p_{i+n}. Figure 4.4 shows the reduced game that player p_1 perceives in case where $n = 2$.

Step 2: Player p_i computes a Nash equilibrium of the reduced game. Let us denote by $\gamma^{i,n}$ the Nash strategy of the player i in the reduced ($2n+1$ players) game.

Step 3: Apply $\gamma^{i,n}$ in the original N players game.

We then compare $J^i(\gamma^{i,n}, u\gamma{-i,n})$ and $\min_\gamma J^i(\gamma, \gamma^{-i,n})$. Figures 4.5, 4.6 and 4.7 illustrate the difference $J^i(\gamma^{i,n}, \gamma^{-i,n}) - \min_\gamma J^i(\gamma, \gamma^{-i,n})$ in logarithmic scale for different values of λ and s. The simulation was performed with $r = 1$ and $N = 101$.

Remark 30. *Figures 4.5-4.7 illustrate that the complexity is approximately at most linear. The reason for this low complexity in the current example is that the Nash strategies of the reduced order games $\gamma^{i,n}$ assign low gains to distant players. Note that none of the reasons for low complexity existing in the previous examples, is present. That is, this low complexity is not due to a law of large numbers, the best response maps are not contractive and there is no cooperation among the players.* □

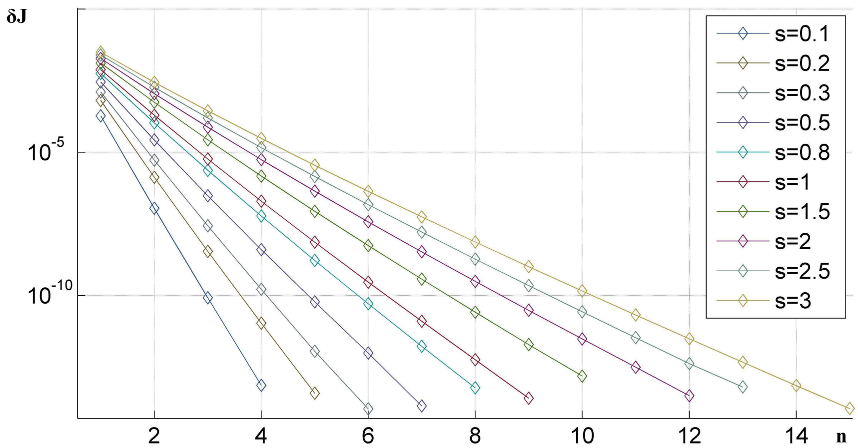


Figure 4.5: The difference $J^i(u^{i,n}, u^{-i,n}) - \min_u J^i(u, u^{-i,n})$ for $\lambda = 0$ as a function of n

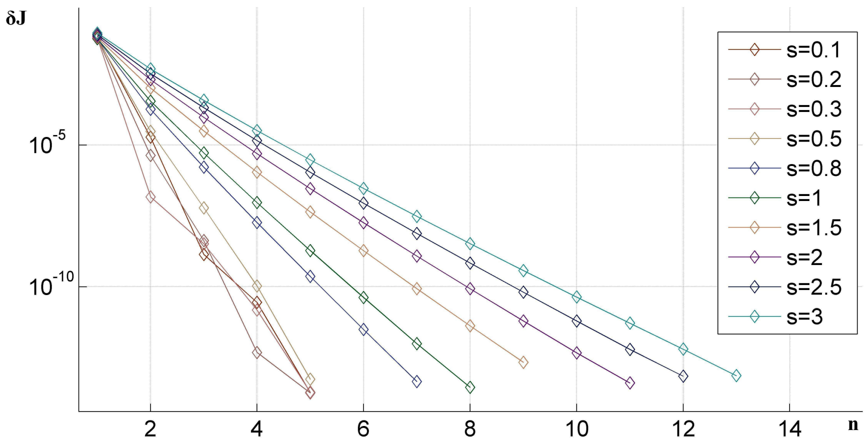


Figure 4.6: The difference $J^i(u^{i,n}, u^{-i,n}) - \min_u J^i(u, u^{-i,n})$ for $\lambda = 0.2$ as a function of n

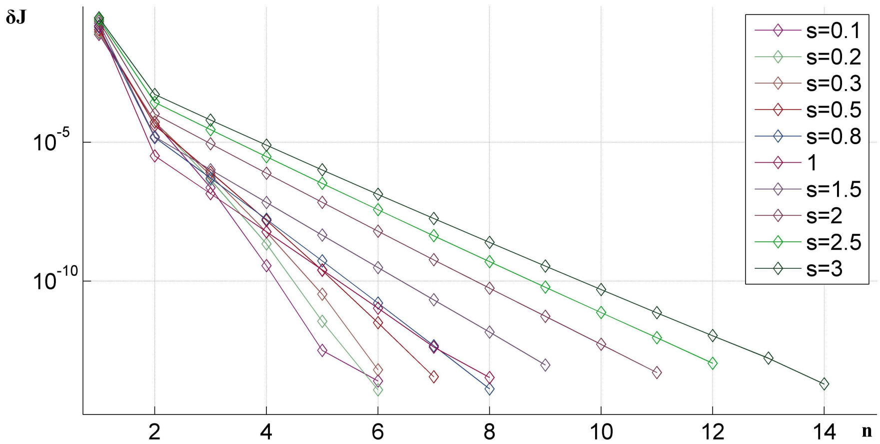


Figure 4.7: The difference $J^i(u^{i,n}, u^{-i,n}) - \min_u J^i(u, u^{-i,n})$ for $\lambda = -0.2$ as a function of n

4.7 Conclusion

Games on Large Networks of interacting agents were considered. A framework involving ensembles of games instead of a single game was introduced, motivated by ideas from the statistical physics. The Probabilistic Approximate Nash equilibrium concept was then defined and compared to the Bayesian Nash equilibrium. Some properties were also proved. The NIC and ANIC functions were defined to quantify the informational complexity of an ensemble of games.

Several special cases were then analyzed. Static games on large Erdos-Renyi random graphs were shown to have different complexity in different regions of the connectivity parameter. A class of LQ games on Erdos-Renyi random graphs were then proved to be simple for the high connectivity regime and an ε-PAN set of strategies was computed. Static games on lattices were then considered and it was shown, using contraction mapping ideas, that they have polynomial complexity where the leading term of the polynomial has the same exponent as the dimension of the lattice. A class of games on ring having chaotic best response maps was then shown to have at most linear complexity, using strategies involving some form of cooperation among the players. LQ games on a ring were finally considered numerically and found to have approximately linear complexity, using a reduction to a smaller game.

In all the cases with low complexity, strategies constituting probabilistic approximate Nash equilibrium where found. The reasons for low complexity in the examples analyzed were laws of large numbers, contractive best response maps, cooperation among the players and low gains to distant players.

4.8 Appendix: Omitted Proofs

4.8.1 Some Probability Inequalities

The following results will be repeatedly used throughout the proof of Propositions and.. 13.

Theorem 4 (Bernstein Inequality)**.** *Let $X_1, \dots, X_N$ be zero mean, independent random variables such that $|X_i| \leq M$. Denoting by $\bar{\sigma}^2 = \frac{1}{N}\sum_{i=1}^N Var\{X_i\}$, for any $t > 0$ it holds:*

$$P\left(\left|\frac{1}{N}\sum_{i=1}^N X_i\right| \geq t\right) \leq 2\exp\left(-\frac{Nt^2}{2\bar{\sigma}^2 + 2Mt/3}\right) \tag{4.75}$$

Lemma 9. *For a finite or infinite sequence of events $A_1, A_2, \dots$ it holds:*

$$P(\bigcap_{i=1,2,\dots} A_i) \geq 1 - \sum_{i=1,2,\dots} P(A_i^c).$$

Theorem 4, called Bernstein inequality, falls into the class of concentration inequalities [Sri]. Lemma 9, has immediate proof.

Lemma 10. *Let X_{ij}, $i, j \in \mathbb{N}$ be a set of zero mean random variables absolutely bounded by a constant M. Assume that $X_{i1}, X_{i2}, \ldots$ are mutually independent. Assume also that for any $N \in \mathbb{N}$ there is a constant c_N such that $c_N \in \omega(\sqrt{\ln N/N})$ Then, for any $\varepsilon_1 > 0$, it holds:*

$$P\left(\left|\frac{1}{Nc_N}\sum_{j=1}^{N} X_{ij}\right| < \varepsilon_1, \textit{ for any } i = 1, 2, \ldots\right) > 1 - \varepsilon_1,$$

for large N.

Proof: Applying Bernstein inequality with $t = \varepsilon_1 c_N$, we obtain:

$$P\left(\left|\frac{1}{Nc_N}\sum_{j=1}^{N} X_{ij}\right| > \varepsilon_1,\right) \leq 2\exp\left(-\frac{Nc_N^2\varepsilon_1^2}{2M^2 + 2Mc_N\varepsilon_1/3}\right)$$

Applying Lemma (9) we have:

$$P\left(\left|\frac{1}{Nc_N}\sum_{j=1}^{N} X_{ij}\right| < \varepsilon_1, \text{ for any } i = 1, 2, \ldots\right) \geq$$
$$\geq 1 - 2N\exp\left(-\frac{Nc_N^2\varepsilon_1^2}{2M^2 + 2Mc_N\varepsilon_1/3}\right)$$

The fact that $c_N^2 \gg \ln N/N$ completes the proof □

Lemma 11. *Consider an Erdos Renyi random graph with connection probability c_N. Let d_{ij} be a random variable with $d_{ij} = 1$ if there is an edge between i and j and $d_{ij} = 0$ otherwise. Then, if $c_N \in \omega(\ln N/\sqrt{N})$ for any $\delta_1, \varepsilon_1 > 0$ it holds:*

$$P\left(\left|\frac{Nc_N}{|\mathcal{N}_i(G)|} - 1\right| < \delta_1, \textit{ for every } i\right) > 1 - \varepsilon_1,$$

for large N.

Proof: For any $\delta_1 > 0$ there is an $\delta_2 = \delta_2(\delta_1) > 0$ such that: $\left|\frac{|\mathcal{N}_i(G)|}{Nc_N} - 1\right| < \delta_2$ implies $\left|\frac{Nc_N}{|\mathcal{N}_i(G)|} - 1\right| < \delta_1$. Furthermore,

$$\frac{|\mathcal{N}_i(G)|}{Nc_N} - 1 = \frac{1}{Nc_N}\sum_{j=1}^{N}(d_{ij} - c_N).$$

Applying Lemma 10 with $X_{ij} = d_{ij} - c_N$ we conclude to the desired result. □

4.8.2 Proof of Proposition 13

(i) It holds:

$$J^i \leq J^{i,a} + E\left\{\sum_{k=0}^{\infty}\rho^k\left[\left(\bar{\theta} - \frac{1}{|N_i(G)|}\sum_{j\in N_i(G)} x_k^j\right)^2\right]\middle|\mathcal{F}^s\right\}$$

The second term of the right hand side is less than the summation of the following terms

$$E\left[\sum_{k=0}^{\infty}\rho^k\left[\bar{\theta}-\frac{1}{N}\sum_{j=1}^{N}(\theta^{j,f}+a^k x_0^j)\right]^2\middle|\mathcal{F}^s\right] \tag{4.76}$$

$$E\left[\sum_{k=0}^{\infty}\rho^k\left[\frac{1}{N}\sum_{j=1}^{N}x_k^j-\frac{1}{N}\sum_{j=1}^{N}(\theta^{j,f}+a^k\tilde{x}_0^j)\right]^2\middle|\mathcal{F}^s\right] \tag{4.77}$$

$$E\left[\sum_{k=0}^{\infty}\rho^k\left[\frac{1}{Nc_N}\sum_{j=1}^{N}(d_{ij}-c_N)x_k^j\right]^2\middle|\mathcal{F}^s\right] \tag{4.78}$$

$$E\left[\sum_{k=0}^{\infty}\rho^k\left[\left(\frac{Nc_N}{|N_i^1(G)|}-1\right)\frac{1}{N}\sum_{j=1}^{N}x_k^j\right]^2\middle|\mathcal{F}^s\right] \tag{4.79}$$

We shall show that the expressions (4.76) - (4.79) are small for all $i=1,\ldots,N$ with high probability if N is large enough. In the expressions (4.76) - (4.79), the expectation and the summation over k are interchangeable due to Bepo Levi theorem [Bil08]. The terms (4.76) and (4.77) are common among the players.

The term (4.76)

It holds:

$$\bar{\theta}-\frac{1}{N}\sum_{j=1}^{N}(\theta^{j,f}+a^k x_0^j)=\frac{s+a^k}{1+s}\left[\bar{\theta}-\sum_{j=1}^{N}\theta_j\right]$$

Thus the term (4.76) is bounded by:

$$\frac{1}{1-\rho}\left[\bar{\theta}-\sum_{j=1}^{N}\theta_j\right]^2,$$

which due to the weak law of large numbers ([Bil08]) is smaller than $\bar{\varepsilon}$ with probability larger than $1-\bar{\varepsilon}$.

The term (4.77)

Fixing k, it holds:

$$X_k=\frac{1}{N}\sum_{j=1}^{N}(x_k^j-\theta^{j,f}-a^k\tilde{x}_0^j)=\frac{1}{N}\sum_{j=1}^{N}\sum_{t=0}^{k}a^t w_{k-t-1}^j$$

Due to independence we have $E[X_k^2]\leq\frac{\sigma^2}{N(1-a)}$. Thus,

$$\sum_{k=0}^{\infty}E[X_k^2]\leq\frac{\sigma^2}{N(1-a)(1-\rho)}.$$

Hence, term (4.77) is less than $\bar{\varepsilon}$ for large N.

The term (4.78)

It holds:

$$x_k^j = a^k x_0^j + \theta^{j,f} + \sum_{t=0}^{k-1} a^t w_{k-t-1}^j. \tag{4.80}$$

Denoting by $X^{ij} = \frac{d_{ij}-c_N}{Nc_N}$, $Y_k^j = a^k(\theta_j - \bar{\theta}) - \frac{\bar{\theta}+s\theta_j}{1+s}$ and $\xi_k^j = \sum_{t=0}^{k-1} a^t w_{k-t-1}^j$, for the term (4.78) we have:

$$\sum_{k=0}^{\infty} \rho^k E\left[\left[\frac{1}{Nc_N}\sum_{j=1}^{N}(d_{ij}-c_N)x_k^j\right]^2 \middle| \mathcal{F}^s\right] =$$

$$= \sum_{k=0}^{\infty} \rho^k E\left[\left[\sum_{j=1}^{N} X^{ij}(Y_k^j + \xi_k^j)\right]^2 \middle| \mathcal{F}^s\right] \tag{4.81}$$

The random variables X^{ij} and Y_k^j are $\mathcal{F}^s$ measurable and ξ_k^j are zero mean and independent for any fixed k. Thus,

$$E\left[\left[\sum_{j=1}^{N} X^{ij}(Y_k^j + \xi_k^j)\right]^2 \middle| \mathcal{F}^s\right] =$$

$$= \left[\sum_{j=1}^{N} X^{ij}Y_k^j\right]^2 + \sum_{j=1}^{N}\left[(X^{i,j})^2 E\left[(\xi_k^j)^2 \middle| \mathcal{F}^s\right]\right] \leq$$

$$\leq \left[\sum_{j=1}^{N} X^{ij}Y_k^j\right]^2 + \frac{\sigma^2}{1-a^2}\sum_{j=1}^{N}\left[(X^{i,j})^2\right] \tag{4.82}$$

For any fixed k, applying Lemma 10 to the set of random variables $Nc_N X^{ij}Y_k^j$ with $\varepsilon_1 = (1-\rho)\bar{\varepsilon}/2$ we have:

$$P\left(\left[\sum_{j=1}^{N} X^{ij}Y_k^j\right]^2 < ((1-\rho)\bar{\varepsilon}/2)^2 < (1-\rho)\bar{\varepsilon}/2\right) >$$

$$> 1-(1-\rho)\bar{\varepsilon}/2 > 1-\bar{\varepsilon}/2.$$

Using the fact that $|Nc_N X^{ij}| \leq 1$ and $c_N \in \omega(1/\sqrt{N})$, the last term of the right hand side of the inequality (4.82) is smaller than $\bar{\varepsilon}/2$ with probability 1 for large N.

Thus, due to (4.81), the term (4.78) is less than $\bar{\varepsilon}$ with probability larger than $1-\bar{\varepsilon}$.

The term (4.79)

Using equation (4.80) we have:

$$E\left[\sum_{k=0}^{\infty} \rho^k \left[\frac{1}{N}\sum_{j=1}^{N} x_k^j\right]^2 \middle| \mathcal{F}^s\right] \leq \frac{1}{1-\rho}\left[4L^2 + \frac{\sigma^2}{N(1-a^2)}\right].$$

The right hand side of the inequality is less than $5L^2/(1-\rho)$, for large N.

Using Lemma 11 and the fact that $\left(\frac{Nc_N}{|N_i^1(G)|}-1\right)^2$ is $\mathcal{F}^s$ measurable, we conclude that the term (4.79) is less than $\bar{\varepsilon}$ with probability larger than $1-\bar{\varepsilon}$, for large N.

The choice $\bar{\varepsilon}=\varepsilon/4$ completes the proof.

(ii) Immediate. ■

4.8.3 Proof of Proposition 14

For any $i=1,\ldots,N$, it holds $\theta^{i,f}=\bar{\theta}$. The random variable $\sum_{k=0}^{T} a^k x_0^i/T$ converges to 0 almost surely. Thus, due to (4.80), it remains to show that the sequence of random variables:

$$X_T^i=\frac{1}{T}\sum_{k=1}^{T-1}\sum_{t=0}^{k-1} a^t w_{k-t-1}^i$$

converges almost surely to 0. The random variable X_T^i can be written as:

$$X_T^i=\frac{1}{T}\sum_{\nu=0}^{T-2}\sum_{k=\nu+1}^{T-1} a^{k+1-\nu} w_\nu^i.$$

Hence, X_T^i is zero mean and Gaussian satisfying $Var(X_T^i)<\frac{\sigma^2}{(1-a)}\frac{1}{T}=\sigma_T^2$. Thus,

$$P(|X_T^i|>1/l)<\exp\left(-1/(l\sigma_T)^2\right).$$

Denote by $B_{T,l}=\{\omega\in\Omega:|X_T^i|>1/l\}$. It is not difficult to see that $\sum_{T=1}^{\infty}P(B_{T,l})<\infty$. Therefore, using the 1st Borel-Cantelli Lemma [Bil08], we have: $P(\limsup_T B_{T,l})=0$. Hence,

$$P\left(\{\omega: X_T^i\to 0\}^c\right)=P\left(\bigcup_{l=1}^{\infty}\limsup_T B_{T,l}\right)=0.$$

Thus, $X_T^i\to 0$ almost surely and the proof is completed. ■

Chapter 5

Dynamic Rules and Cheating in Games. Application to Electricity Markets

This chapter studies Dynamic Game situations with incomplete structural information, motivated by problems arising in electricity market modeling. Some Adaptive / Learning strategies are considered as an expression of the Bounded Rationality of the participants of the game. The Adaptive strategies are typically not in Nash equilibrium. Thus, the possibility of cheating, i.e. the use of the dynamic rule of the opponent in order to manipulate her, appears. In order to assess Adaptive / Learning strategies, two criteria are stated: Firstly, how far the cost of each player is from the cost of her best response in the sense of the Nash equilibrium. Secondly, we consider the case where the first player follows the adaptive strategy and the second player implements the best response to the first player. Then, the criterion depends on the difference of the cost of the first player comparing with the cost in case where both players follow their adaptive control laws. This difference may be positive or negative.

We then examine a smaller class of strategies, called the pretender strategies, where each player acts as if she had different, not real, preferences. It turns out that under certain technical conditions, if only one player is pretending, she can achieve the same cost as if she were the Stackelberg leader. The situation where all the players are pretending is then considered and an auxiliary game, called the pretenders game, is introduced. The effects of adaptation and cheating, when the number of players in the game becomes large, are examined in a simple example. It turns out that, the incentive to cheat is quite small, in the case where there is a large number of relatively weak players.

Models for competitive Electricity Markets are then considered. Particularly, a Cournot oligopoly model, a supply function model and a mechanism form the literature (Rasouli Teneketzis mechanism) are studied. It turns out that cheating some times increases competition, other times enhances cooperation and in the case of Rasouli-Teneketzis mechanism makes the system not working at all.

5.1 Introduction

In a number of real game situations there is a number of decision makers that interact strategically over time but each one of them has only a partial knowledge of the intentions of the others. A particular example is an electricity market where several producer firms are competing repeatedly over time and each firm knows its own costs but not the costs of the other firms ex. [GCF$^+$12], [SVP02]. Such strategic interactions over time can be described by dynamic games with incomplete structural information.

Quite often is very difficult to find a Nash equilibrium for dynamic or repeated games with incomplete information. Two difficult problems may arise; the first is due to the “Witsenhausen effect" [Wit68], i.e. the current action of each player affects the future state estimation of the other players. The second is due to the “Dual Control effect" [Fel60], i.e. that the current action of a player affects the quality of his own future parameter estimation. Due to the later difficulty, the Optimization problem has not been solved analytically even for the single player (control) case, except only of a few special cases [Wit02].

In this context, it has been proposed that the players, instead of trying to find the equilibrium strategies, they use some simple deterministic or stochastic dynamic rules which determine the future actions of the players based on their past actions. The use of such rules, usually, reflect the bounded rationality [Sim72], [Sel01] of the participants of the game, i.e. their inability to solve very difficult problems. These rules can be models of learning, adaptation, evolution or imitation [FL98], [You04], [SLB08] ch. 7, [LT11], [CC83] [Pap88], [CC83]. Most of the research on dynamic rules focuses on convergence issues. However, a set of such strategies is not typically in Nash equilibrium.

Are these adaptive/learning strategies a reasonable prediction for the evolution of the game? Particularly, if the players knew that they are going to implement those strategies, would they stick on them? When a player is implementing an adaptive control law, she may be viewed by another player as a system under control. That is, the other player may “cheat”, i.e. use the knowledge of the adaptation law of the first player to manipulate her. The topic of this chapter is to study phenomena like “cheating” and the implications that may have to the costs of the participants of the game.

At first, two criteria to assess dynamic rules are stated. The first is based on the opportunities that appear; that is, how far is a dynamic rule from being optimal. We then assume that the other player has adopted the best response with respect to the former player’s strategy. The second criterion depends on how much the former player may loose, in that case. The assessment, as well as the computational problems arise are illustrated through a very simple numerical example. Then a smaller class of “cheating” strategies is considered, called the “pretender’s” strategies. The cheating player implements her learning/adaptation rule as if she had a different, not real type (preferences). The case, where all the players are pretending, is then studied and some possible limit points are identified, using an auxiliary game called the “pretenders’ game”. The relationship of cheating and Stackelberg leadership is also examined. The phenomena described are illustrated in a simple class of quadratic opinion games. The effects of cheating are, finally, studied in competitive electricity market models.

The rest of this chapter is organized as follows: In section 2, the criteria for the assessment of the dynamic rules are stated. The pretenders’ strategies are studied in section 3. The opinion quadratic game examples are given in Section 4. Section 5 studies some electricity market games. Particularly, Section 5.1 considers a Cournot duopoly, Section 5.2 a supply function model and Section 5.3 a mechanism from the literature. Finally, Section 6 concludes.

5.2 Adaptive Strategies and Assessment

In this section, we define two criteria in order to assess a set of adaptive strategies. At first, a general form of strategies is considered. For simplicity reasons, the criteria are stated for two player games.

Let us first define formally the game. There are two players of types θ_1, θ_2. Each player knows her own type and θ_1, θ_2 are part of a random vector $\bar{\theta} = [\theta_1 \;\; \theta_2 \;\; \theta]^T \in \bar{\Theta} = \Theta_1 \times \Theta_2 \times \Theta$ having a commonly known distribution.

The dynamics have the form:

$$x_{k+1} = f(x_k, \bar{\theta}, u_k^1, u_k^2, w_k), \tag{5.1}$$

where u_k^1, u_k^2 are the action variables of the players and w_k a random disturbance.

The cost functions have the form:

$$J_i = E\left[\sum_{k=0}^{T} \rho^k L_i(x_k, \theta_i, u_k^1, u_k^2)\right], \tag{5.2}$$

where $\rho \in (0, 1]$ is a discount parameter and T can have a finite or infinite value. An alternative formula for the costs is:

$$J_i = \lim_{T\to\infty} E\left[\frac{1}{T}\sum_{k=0}^{T} L_i(x_k, \theta_i, u_k^1, u_k^2)\right]. \tag{5.3}$$

Equations (5.1), (5.2) or (5.1), (5.3) can describe dynamic games as well as repeated static games.

Each player receives at each time step an information vector according to:
Information Structure 1: $\bar{I}_k^{i,new} = (x_k, u_{k-1}^i)$, or
Information Structure 2: $\bar{I}_k^{i,new} = (x_k, u_{k-1}^i, u_{k-1}^{-i})$.
The information that each player possesses at time step k has the form $I_k^i = (\bar{I}_0^{i,new}, \ldots, \bar{I}_k^{i,new})$.

Let us now describe a general form of strategies of the players: $s^i = (\gamma_1^i, \gamma_2^i, \dots)$, $i = 1, 2$, where γ_k^i is a function having the form:

$$\gamma_k^i : (x_0, x_1, \dots, x_k, \theta_i) \mapsto u_k^i \in U^i.$$

We shall focus on "state feedback" strategies, where all the previous information is used only for "adaptation", assuming that the players have full access to the state vector. Specifically,

$$u_k^i = \gamma_k^i(x_k, \theta_i, \hat{\theta}_k^i), \tag{5.4}$$

where $\hat{\theta}_k^i$ is the adapted parameter of player i. We assume that the adapted parameter evolves according to a dynamic equation:

$$\hat{\theta}_{k+1}^i = \phi^i(\hat{\theta}_k^i, \bar{I}_{k+1}^{i,new}, \theta_i). \tag{5.5}$$

For infinite horizon games, the following property is quite interesting see for example [YP94]:
Property 1: The adapted values $\hat{\theta}_k^1, \hat{\theta}_k^2$ converge to some limits $\hat{\theta}_\infty^1, \hat{\theta}_\infty^2$, such that the feedback (no memory) strategies $\gamma_k^i(x_k, \theta_i, \hat{\theta}_\infty^i),\ i = 1, 2$ constitute a Perfect Nash equilibrium for the complete information game.

The criteria depend on the best response of each player to the opponent's strategy. Thus, in order to state the criteria, the best response $\bar{s}^i$, $i = 1, 2$ of each player given the strategy of the other player s^{-i}, is considered.

The first criterion states that each player does not have a lot to gain from moving to her best response. We call it the opportunity criterion.
Opportunity Criterion *with parameter* a: For some $a > 0$ it holds:

$$J^i(s^i, s^{-i}) \leq J^i(\bar{s}^i, s^{-i}) + a, \tag{5.6}$$

for any θ_i, $i = 1, 2$.

The second criterion states that if the other player moves to her best response, the first player does not have a lot to lose. Let us call it the conservative criterion.
Conservative Criterion *with parameter* b: For some $b > 0$, it holds:

$$J^i(s^i, \bar{s}^{-i}) \leq J^i(s^i, s^{-i}) + b, \tag{5.7}$$

for any θ_i, $i = 1, 2$.

Definition 10. *A pair of strategies s^1, s^2 is a, b - not sensitive to cheating if either the Opportunity Criterion with parameter a or the Conservative Criterion with parameter b holds.*

Remark 31. *Definition 10 borrows some ideas form satisficing based decision making [Sim72], where the values of a, b have roles related to the satisfactory levels.*

Remark 32. *Definition 10 corresponds to conservative players. Particularly, it states that each player believes either that the opponent will not have enough motivation to change her strategy or that if she has, this change would not increase the cost of the former player a lot. For less conservative players, the "either, or" of the definition should be replaced by "and". An alternative definition would involve any "better response" $\tilde{s}^i$ instead of the best response $\bar{s}^i$.*

Remark 33. *A class of games which satisfy Definition 1 is Team Games. More generally, if $max|J_1(s_1, s_2) - J_2(s_2, s_2)| \leq max\{a, b\}$, then the game is a, b not sensitive to cheating.*

The verification of Definition 10 is not easy due to the fact that the optimal control problems involved are quite difficult.

5.3 Pretenders Strategies

In this section we focus on a special class of cheating strategies called the pretenders' strategies. Particularly, the cheating player pretends to have a false type, probably in a response of the other players' actions. For simplicity, we assume that the θ is commonly known, i.e. $\Theta = \{\theta\}$. In what follows, only games with infinite horizon will be considered.

The general form of a pretender's strategy that corresponds to (5.4) is:

$$u_k^i = \gamma_k^i(x_k, \theta_k^{i,pr}, \hat{\theta}_k^i), \tag{5.8}$$

where the pretended type $\theta_k^{i,pr}$ is given as an output of a system:

$$z_{k+1}^i = \phi^{i,pr}(z_k^i, \bar{I}_{k+1}^{i,new}, \theta_i), \tag{5.9}$$

$$\theta_k^{i,pr} = \psi^i(\theta_i, z_k^i). \tag{5.10}$$

Equations (5.8)-(5.10) represent a cheating player who pretends to have a type that depends on her real type and a new, probably augmented, adapted parameter z_k^i. That is, in order to pretend adequately, it is probably useful to accumulate more information.

In what follows, we consider the stationary cases where one of the players or both the players are pretending. These stationary outcomes are be possible limit points of learning/adaptation rules, when the players are pretending.

5.3.1 Optimal Stationary Pretending

We consider the possible limit points of the pretending strategies, assuming games with infinite horizon and long run average cost. We assume that only player 1 is pretending. In the spirit of Property 1, we analyze the following situation. Player 1 has revealed all the useful information for θ_2 and the pretended type of player 1 has converged to $\theta_\infty^{1,pr}$. Player 2 reacts to a player of type $\theta_\infty^{1,pr}$. Furthermore, the pair of strategies $\gamma^1(x_k, \theta_\infty^{1,pr}, \hat{\theta}_\infty^1), \gamma^2(x_k, \theta_2, \hat{\theta}_\infty^2)$ constitute a Perfect Nash equilibrium for the game with full information and types $\theta_\infty^{1,pr}, \theta_2$.

In order to define the Optimal Stationary Pretending, the following assumption is made:
Assumption 1: For any pair of types θ_1, θ_2, there exist a unique Perfect Nash equilibrium of the full information game. Let us denote by $\gamma^{i,N}_{\theta_1,\theta_2}(x_k),\ i = 1, 2$ the pair of strategies constituting the Nash equilibrium. Assumption 1 is not unusual in static or dynamic games.

The optimal pretending for the player 1 is given by:

$$\theta^{1,pr}_{\infty} = arg \min_{\tilde{\theta}_1 \in \Theta_1} J_1(\gamma^{1,N}_{\tilde{\theta}_1,\theta_2}, \gamma^{2,N}_{\tilde{\theta}_1,\theta_2}) \tag{5.11}$$

It is interesting to compare the cost that the cheating player 1 attains with the cost of the full information game having 1 as Stackelberg leader. Let us denote by $\gamma^{i,S}_{\theta_1,\theta_2},\ i = 1, 2$, a pair of strategies constituting a feedback Stackelberg equilibrium with 1 as leader.

Proposition 20. *If player 1 pretends optimally, then:*

$$J_1(\gamma^{1,S}_{\theta_1,\theta_2}, \gamma^{2,S}_{\theta_1,\theta_2}) \leq J_1(\gamma^{1,N}_{\theta^{1,pr}_{\infty},\theta_2}, \gamma^{2,N}_{\theta^{1,pr}_{\infty},\theta_2}). \tag{5.12}$$

An equality is attained is there exist a $\tilde{\theta}_1 \in \Theta_1$ *such that* $\gamma^{1,S}_{\theta_1,\theta_2} = \gamma^{1,N}_{\tilde{\theta}_1,\theta_2}$.

Proof. The proof is immediate. □

Remark 34. *Proposition 1 roughly says that if there is enough uncertainty and only one player pretends, then:*
"The pretender becomes the leader".

5.3.2 The Pretenders' Game

The case where both players are pretending is then analyzed. Under the Assumption 1, an auxiliary game, called the pretenders' game can be defined.

Definition 11. *For the game described by* (5.1), (5.3)*, under assumption 1, the corresponding pretenders' game involves the same players, the actions of player* i *is* $\theta^{i,pr}$ *and the cost is given by:*

$$\bar{J}^i(\theta^{i,pr}, \theta^{-i,pr}) = J^i(\gamma^{i,N}_{\theta^{i,pr},\theta^{-i,pr}}, \gamma^{-i,N}_{\theta^{i,pr},\theta^{-i,pr}}, \theta_i), \tag{5.13}$$

where θ_i *is the actual type.*

An interesting point is the Nash equilibrium of the pretenders' game. This point can serve as a prediction of the outcome of the original game.

Definition 12. *For the game described by* (5.1), (5.3)*, under assumption 1, a Nash pretenders' outcome is the pair of strategies* $(\gamma^{1,N}_{\theta^{1,pr}_{\infty},\theta^{2,pr}_{\infty}}, \gamma^{2,N}_{\theta^{1,pr}_{\infty},\theta^{2,pr}_{\infty}})$*, where* $(\theta^{1,pr}_{\infty}, \theta^{2,pr}_{\infty})$ *is an equilibrium of the corresponding pretenders' game.*

The Nash pretenders' outcome is interesting as a possible limit point of a learning/adaptive algorithm, where each of the players pretends dynamically to have a false type.

Remark 35. *The definitions and the reasoning of Section 3 can be extended to the many players case.*

Remark 36. *Let us comment on the relationship with the work in [AAEA10]. This work identifies an interesting phenomenon in routing games with partially altruistic players: in some cases "When a user increases its degree of cooperation while other users keep their degree of cooperation unchanged, leads to performance improvement of that user". To be more precise, if J^i is the "selfish" cost of player i, [AAEA10] considers the case where the players are partially altruistic, that is each player tries to minimize:*

$$\tilde{J}^i = (1 - a_i)J^i + a_i J^{-i}, \tag{5.14}$$

where a_i is the degree of cooperation. It is then shown numerically that, when the players apply strategies constituting a Nash equilibrium of the game described by (5.14)*, it is possible J^i to decrease when a_i increases. This phenomenon is called the " Paradox in cooperation". In this context the notion of the price of unilateral altruism was introduced [AM11].*

Consider such a situation where the degree of cooperation of player 2 is commonly known, whereas the degree of cooperation of player 1 is known only to player 1 and that there is a "Paradox in cooperation". Then, player 1 has an incentive to pretend to be more altruistic than she actually is. Furthermore, even if player 2 knew the actual degree of cooperation of player 1, she has an incentive to accept that player 1 is more altruistic, otherwise she is going to induce a less altruistic behavior to player 1 and thus hurt herself. Finally, player 2 cannot discriminate among the case where player 1 is trying to reach a Nash equilibrium and she is more competitive and the case where player 1 is less competitive and pretends.

5.4 Quadratic Opinion Games

In this section, a very simple class of repeated, quadratic games where exact results are easily obtained, is analyzed. The results of this section help us to get some intuition.

5.4.1 Two Player Games

A two players opinion game, is studied. The cost function is given by (5.3) and the instantaneous costs by:

$$L_1 = (u^1 - \theta_1)^2 + (u^1 - u^2)^2, \tag{5.15}$$

$$L_2 = (u^2 - \theta_2)^2 + (u^2 - u^1)^2, \tag{5.16}$$

where $(\theta_1, \theta_2) \in \mathbb{R}^2$. We assume an information structure of type 2. The full information (static) game has a unique Nash equilibrium:

$$u^{i,N} = \frac{2}{3}\theta_i + \frac{1}{3}\theta_{-i}, \;\; i = 1, 2 \tag{5.17}$$

Several adaptive (iterative) techniques for the incomplete information game were studied in [Pap86]. Probably, the simplest one is the best response map:

$$u_k^i = (\theta_i + \hat{\theta}_k^i)/2, \tag{5.18}$$

$$\hat{\theta}_k^i = u_{k-1}^{-i}. \tag{5.19}$$

If both players follow their best response maps, their actions will converge to the Nash equilibrium of the full information static game.

Let us then analyze the situation of player 1 cheating against player 2 and player 2 following (5.18), (5.19). Due to the fact that the map $\theta_1 \mapsto u^{1,N}$ in (5.17) is onto, Proposition 1 applies. Thus, the feedback Stackelberg cost for player 1 is feasible through pretending.

The optimal pretending for player 1 is given by:

$$\theta_\infty^{1,pr} = \frac{6}{5}\theta_1 - \frac{1}{5}\theta_2. \tag{5.20}$$

Player 1 can use several ways to learn θ_2 in order to implement her cheating policy. One way is to use only the last iteration. Particularly:

$$z_{k+1}^{1,1} = z_k^{1,2}, \tag{5.21}$$

$$z_{k+1}^{1,2} = u_k^1, \tag{5.22}$$

$$\theta_k^{1,pr} = \frac{6}{5}\theta_1 - \frac{1}{5}(2u_{k-1}^2 - z_k^{1,1}). \tag{5.23}$$

An alternative way is to use Recursive Least Squares (RLS).

Figure 1 shows the action trajectories when no player, one player and both players are pretending. The parameters are $\theta_1 = 1$ and $\theta_2 = -1.3$. If no player is pretending, then the dynamic rules lead to the Nash equilibrium. If a single player is pretending, the dynamic rule converges to the Stackelberg equilibrium having the pretending player as a leader. Finally, in the case, where both players are pretending the dynamic rule converges to the Nash pretenders' outcome.

Example 12. *In this example we study a dynamic game which represents the repeated version of a static Stackelberg game. Particularly, player* 1 *acts at time instants* $1, 3, 5, \dots$ *and player* 2 *at* $2, 4, 6, \dots$. *The feedback Nash equilibrium of the dynamic game corresponds to the Stackelberg equilibrium of the static game.*

The instant costs are given by:

$$L^1(k) = (x_k - 1)^2 + \theta_1(x_k - u_k^2)^2, \quad k = 2, 4, \dots \textit{ and 0 otherwise}, \tag{5.24}$$

$$L^2(k) = (u_k^2 - 2)^2 + \theta_2(x_k - u_k^2)^2, \quad k = 2, 4, \dots \textit{ and 0 otherwise}, \tag{5.25}$$

and the state equation by:

$$x_{k+1} = \begin{cases} u_k^1 & \textit{if } k = 1, 3, 5, \dots \\ \emptyset & \textit{if } k = 2, 4, 6 \end{cases} \tag{5.26}$$

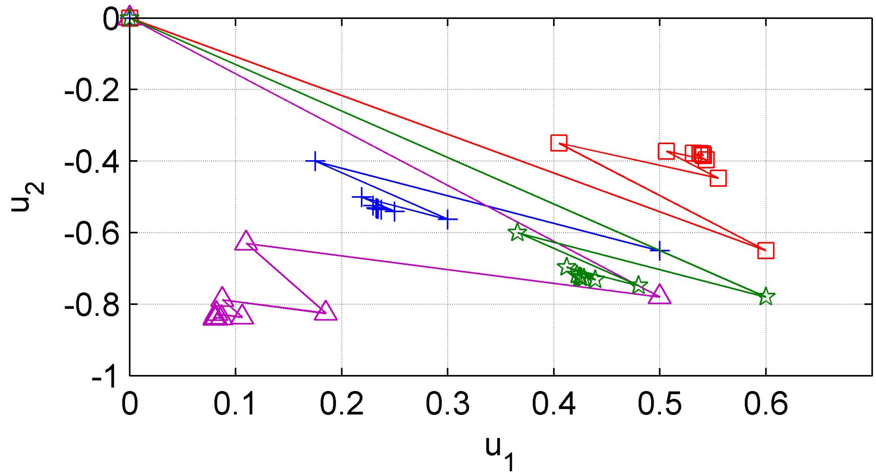


Figure 5.1: The blue line corresponds to best response with no pretending players, the red with player 1 pretending, the purple with player 2 and the green with both players pretending.

The Nash equilibrium (which coincides with the Stackelberg equilibrium of the static game) is given by:

$$u_k^1 = \frac{(1+\theta_2)^2 + 2\theta_1}{(1+\theta_2)^2 + \theta_1}, \quad k = 1, 3, 5 \ldots \tag{5.27}$$

$$u_k^2 = \frac{2 + \theta_2 x_k}{1 + \theta_2}, \quad k = 2, 4, 6 \ldots \tag{5.28}$$

A simple dynamic rule that leads to the Nash equilibrium is given by:

$$u_k^1 = \frac{(1+\hat{\theta}_1)^2 + 2\theta_1}{(1+\hat{\theta}_1)^2 + \theta_1}, \quad k = 1, 3, 5 \ldots \tag{5.29}$$

$$u_k^2 = \frac{2 + \theta_2 x_k}{1 + \theta_2}, \quad k = 2, 4, 6 \ldots \tag{5.30}$$

where $\hat{\theta}_1$ is the least squares estimate of θ_2 based on (5.28).

A very simple cheating rule for player 2 is to use (5.30) *with $\theta_k^{2,pr} = (\varepsilon - 1)$ in the place of θ_2, where ε is a small positive constant. Using these rules, after two time steps, the actions of the players will converge to:*

$$u^1 = \frac{\varepsilon^2 + 2\theta_1}{\varepsilon^2 + \theta_1}, \tag{5.31}$$

$$u^2 = \frac{\varepsilon^2 + \varepsilon + 2\theta_1}{\varepsilon^2 + \theta_1}. \tag{5.32}$$

This pair of strategies approaches $(2, 2)$, as ε approaches 0. Furthermore, the cost of player 2 approaches 0.

Remark 37. *In the Example 12, the Stackelberg leader behaves approximately as the follower wants her to behave. The situation is quite similar with an inverse Stackelberg game with player 2 as the leader [HLO82]. Thus, player 2 can achieve a very strong form of leadership and in* *this sense, we may conclude that:*

"At least in some cases, regarding leadership, the best use of the information available is much more important than the commitment or the timing of the game."

Remark 38. *In the Example 12, player* 2 *can achieve approximately the same cost, as if she was acting on player* 1*' behalf. An interesting question arises: "In which sense does this outcome expresses the 'free will' of the player 1"?*

5.4.2 Many Players Games

We then move to a slightly more general case, where there are N symmetrically interacting players. The instantaneous costs have the form:

$$L_i = (u^i - \theta_i)^2 + \left(u^i - \frac{1}{N-1}\sum_{j\neq i} u^j\right)^2. \tag{5.33}$$

The unique feedback Nash equilibrium of the full information game has the form:

$$u^i = \frac{N\theta_i + \sum_{j\neq i}\theta_j}{2N-1} \tag{5.34}$$

We then consider the pretenders' game. The costs are given by:

$$\tilde{J}_i = \frac{1}{(2N-1)^2}\left[\left(N\theta^{i,pr} + \sum_{j\neq i}\theta^{j,pr} - (2N-1)\theta^i\right)^2 + \left((N-1)\theta^{i,pr} - \sum_{j\neq i}\theta^{j,pr}\right)\right]. \tag{5.35}$$

The equilibrium of the pretenders' game is characterized by the following equations:

$$\theta^{i,pr} = \frac{2N^2 - N}{2N^2 - 2N + 1}\theta^i - \frac{1}{2N^2 - 2N + 1}\sum_{j\neq i}\theta^{j,pr}. \tag{5.36}$$

Hence, the equilibrium of the pretenders' game is given by:

$$\theta^{i,pr} = \left(1 + \frac{1}{2N}\right)\theta^i - \frac{1}{2N(N-1)}\sum_{j\neq i}\theta^j \tag{5.37}$$

Remark 39. *Equation* (5.37) *shows that, when there is a symmetric interaction, the players tend to pretend less as the number of the players increases. Furthermore, the value of pretending decreases and the Nash pretenders outcome approaches the Nash equilibrium, as the number of players increases.*

5.5 Electricity Market Models

In this section three competitive market models are analyzed. The first considers a Cournot duopoly and analytical results for the Nash pretenders outcome are presented. It is shown that pretending enhances the competition among the players. The second example is a linear supply function duopoly. A simple pretending algorithm is proposed and numerical results show that the pretending is cooperative and the price increases. The third example is a mechanism from the literature [RT14]. It is shown numerically that cheating may make the mechanism not working at all.

5.5.1 A Cournot Duopoly

In this subsection, we study a Cournot duopoly of electric power producing firms. Each player decides the amount of energy that produces. The instantaneous cost of the players is given by:

$$L_i(q_1, q_2) = C_i(q_i) - pq_i, \tag{5.38}$$

where $C_i(q_i)$ is the production cost, p is the price and q_i is the quantity produced by player i. We assume constant marginal cost, i.e. $C_i(q_i) = c_i q_i$ and linear demand curve, i.e. $p = A - B(q_1 + q_2)$. Furthermore, we assume that the quantities produced are non-negative and that $c_i < A, i = 1, 2$.

The Nash equilibrium has one of the following forms:

$$\begin{aligned} q_1^N &= \frac{A - 2c_1 + c_2}{3B}, \\ q_2^N &= \frac{A - 2c_2 + c_1}{3B}, \end{aligned} \tag{5.39}$$

if $A + c_2 > 2c_1$ and $A + c_1 > 2c_2$,

$$\begin{aligned} q_1^N &= \frac{A - c_1}{2B}, \\ q_2^N &= 0, \end{aligned} \tag{5.40}$$

if $c_2 > c_1$ and $A + c_1 < 2c_2$, or

$$\begin{aligned} q_1^N &= 0, \\ q_2^N &= \frac{A - c_2}{2B}, \end{aligned} \tag{5.41}$$

if $c_2 > c_1$ and $A + c_2 < 2c_1$. The form (5.39) is more interesting than the other forms of the equilibrium, due to the fact that both players are producing in the equilibrium. Thus, we assume that $A + c_2 > 2c_1$ and $A + c_1 > 2c_2$.

Let us first consider the case where only player 1 is pretending. The optimal stationary pretending for player 1 is given by:

$$c_1^{pr} = \begin{cases} \frac{6c_1 - A - c_2}{4}, & \text{if} \quad 2c_1 + A > 3c_2 \\ 2c_2 - A, & \text{if} \quad 2c_1 + A \leq 3c_2 \end{cases} \tag{5.42}$$

Remark 40. *If $A + c_2 > 2c_1$ and $A + c_1 > 2c_2$ then it holds $c_1^{pr} < c_1$. Thus, producer 1 pretends to be cheaper than she actually is. Furthermore, in the second brunch of* (5.42) *producer 1 is pretending to have a cost such that the second producer is not producing. In both cases, the pretending player improves her cost, while player 2 gets worse.*

Table 5.1: Symmetric Cournot duopoly outcomes with cheating

Pretends	q_1	q_2	L_1	L_2	p
No player	0.4	0.4	-0.16	-0.16	1.2
Player 1	0.6	0.3	-0.18	-0.09	1.1
Player 2	0.3	0.6	-0.09	-0.18	1.1
Both	0.48	0.48	-0.1152	-0.1152	1.04

We then consider the case where both the players are pretending. We further assume that no player is interested to pretend to have a production cost, such that it makes the other player not to produce, i.e. that it holds $2c_1 + A > 3c_2$ and $2c_2 + A > 3c_1$. Then, the Nash pretenders outcome is given by:

$$\begin{aligned} c_1^{pr} &= \frac{24c_1 - 3A - 6c_2}{15}, \\ c_2^{pr} &= \frac{24c_2 - 3A - 6c_1}{15}, \end{aligned} \tag{5.43}$$

Remark 41. *(i) It is not difficult to see that under the aforementioned assumptions, the pretended production costs are smaller than the actual. Thus, pretending enhances the competition.*

(ii) For higher values of A, the players tend to seem cheaper, i.e. if the demand increases the producers appear to have a lower cost.

(iii) The cost that each player seems to have increases with a factor 1.6 (grater than 1) with respect to the actual cost.

Example 13. *This example studies a symmetric Cournot duopoly. We assume that $A = 2$, $B = 1$ and $c_1 = c_2 = 0.8$. The quantities, the costs, as well as the price is illustrated in the Table 5.1, when no player, one player or both players are pretending. Table 5.1 shows that pretending is a non-cooperative action and that the Nash pretenders' outcome is worse for both of the players than the Nash equilibrium. However, the price is smaller if one or both of the players are pretending and thus, the consumers are improving.*

Example 14.

5.5.2 Linear Supply Function Model

In this subsection, we assume that each player proposes a linear supply function, that is at each price level an energy quantity proportional to that price. We also consider linear demand of the form $q = A - Bp$. The instantaneous cost of each player is given by:

$$L_i = c_i q_i^2 - p q_i. \tag{5.44}$$

Each player chooses a constant u_i such that $q_i = u_i p$. Simple manipulations show that:

$$p = \frac{A}{u_1 + u_2 + B}, \tag{5.45}$$

and:

$$L_i = A^2 \frac{c_i u_i^2 - u_i}{(u_1 + u_2 + B)^2}. \tag{5.46}$$

The best response maps are given by:

$$u_1^+ = \frac{B + u_2}{2c_1 u_2 + 2c_1 B + 1}, \tag{5.47}$$

$$u_2^+ = \frac{B + u_1}{2c_2 u_1 + 2c_2 B + 1} \tag{5.48}$$

It is easy to see that the best response map is contractive with respect to the infinity norm. Thus, these equations can be used repeatedly to find the Nash equilibrium.

A simple cheating algorithm for player 1 is then described.

Algorithm 3

1: Choose $c_0^{1,pr} = c_1$.
2: At time step k choose u_1 according to (5.47) with $c_k^{1,pr}$ in the place of c_1.
3: Estimate c_2 from the last M measurements according to (5.48) using least squares.
4: Set $c_{k+1}^{1,pr} = c_k^{1,pr}$
5: For some small constant ε, with probability $1 - \varepsilon$, jump to Step 2.
6: Implement repeatedly (5.47)-(5.48) with the estimated value of c_2 on a grid of several c_1 values.
7: Choose $c_{k+1}^{1,pr}$ the value of c_1 which minimizes the actual cost and jump to Step 2.

A similar pretending algorithm may be used by player 2.

Example 15. *Assume that $A = 1$, $B = 0.5$, $c_1 = 1$, $c_2 = 1.2$. Player 1 implements the Algorithm 3 and player 2 uses the best response until time step 100 and the Algorithm 3 afterwards. The results are illustrated in the Figures 5.2 - 5.4.*

Remark 42. *The example shows that the pretending is a cooperative action, i.e. each player prefers the other player to cheat than not to cheat. Both the players pretend to have a higher cost than her actual cost.*

Example 16. *In this example, we assume that the production costs are commonly known and player 1 pretends to be partially altruistic. Thus, the cost of player 1, as perceived by player 2 has the form:*

$$L_1 = A^2 \frac{(1-a)(c_1 u_1^2 - u_1) + a(c_2 u_2^2 - u_2)}{(u_1 + u_2 + B)^2}, \tag{5.49}$$

where a is the amount of altruism.

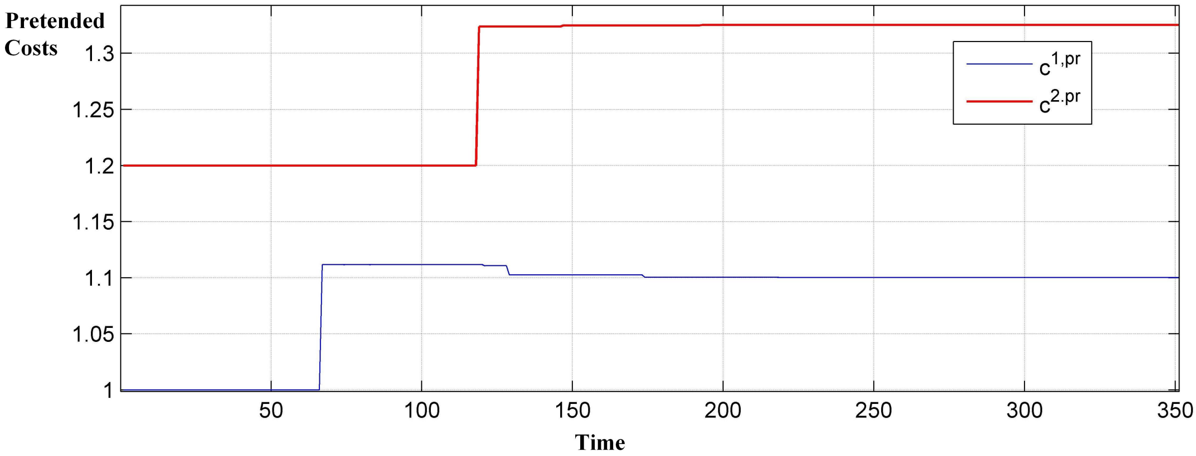


Figure 5.2: The pretended production costs $c^{1,pr}, c^{2,pr}$

.

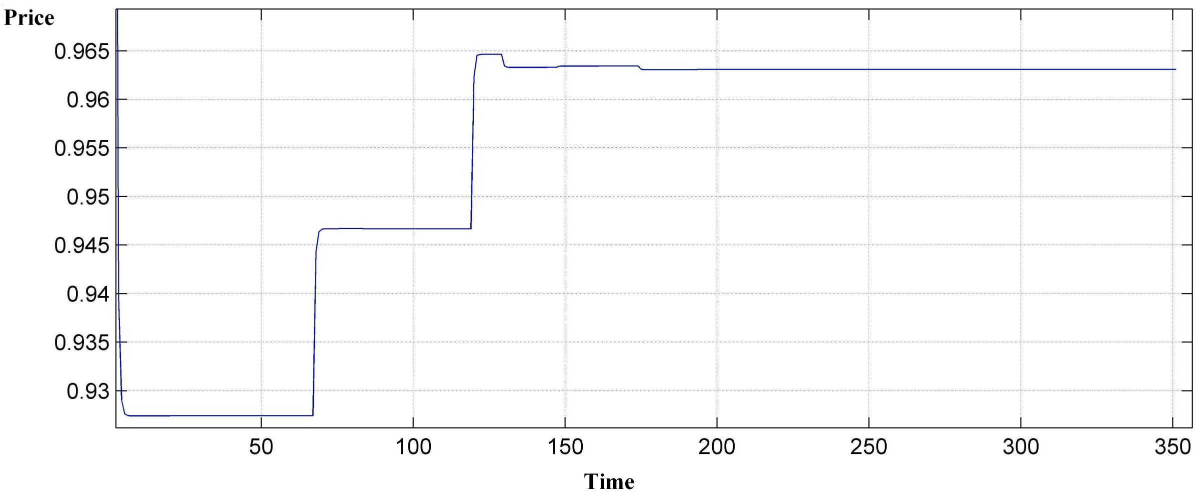


Figure 5.3: The time evolution of the price

.

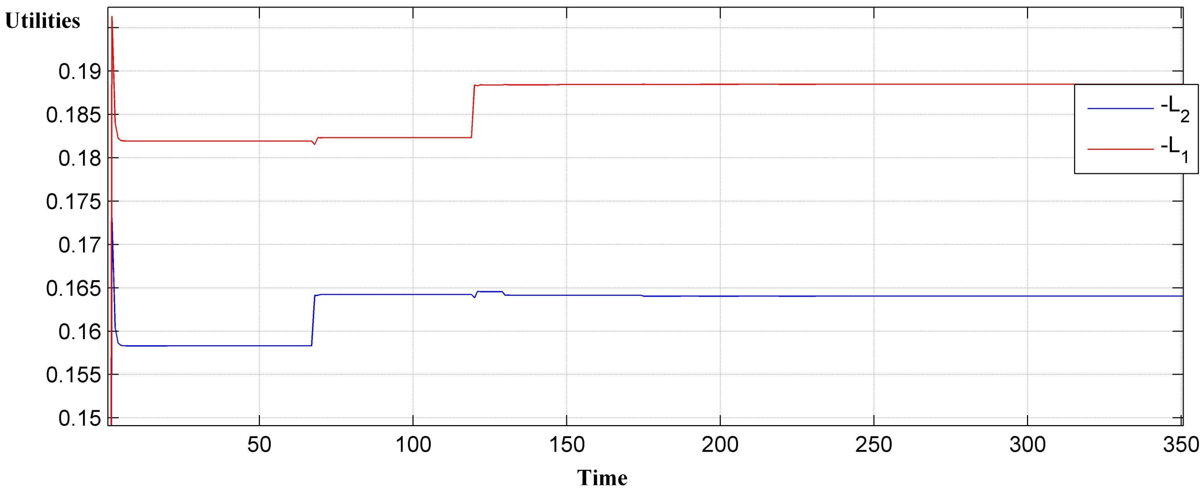


Figure 5.4: The time evolution the utilities $(-L_i)$

.

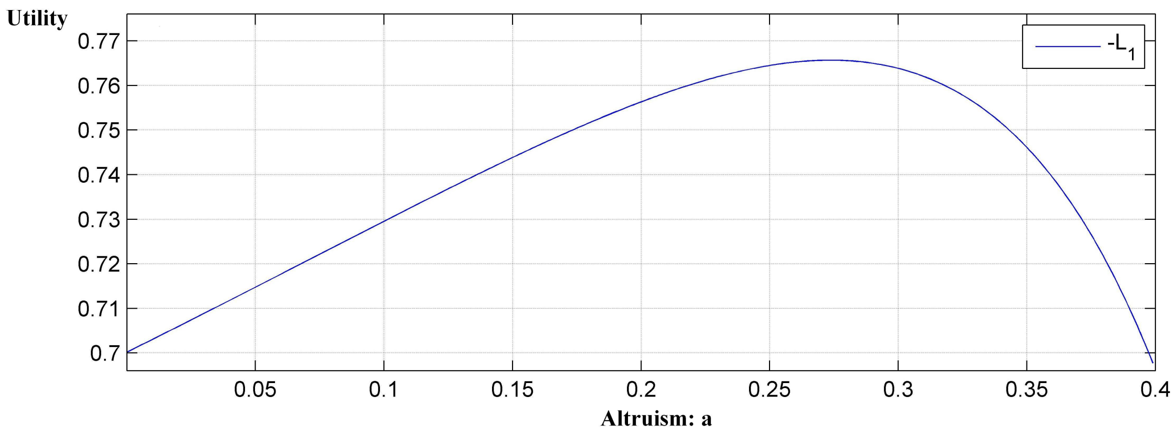


Figure 5.5: The utility of player 1 ($-L_1$) for various values of a.

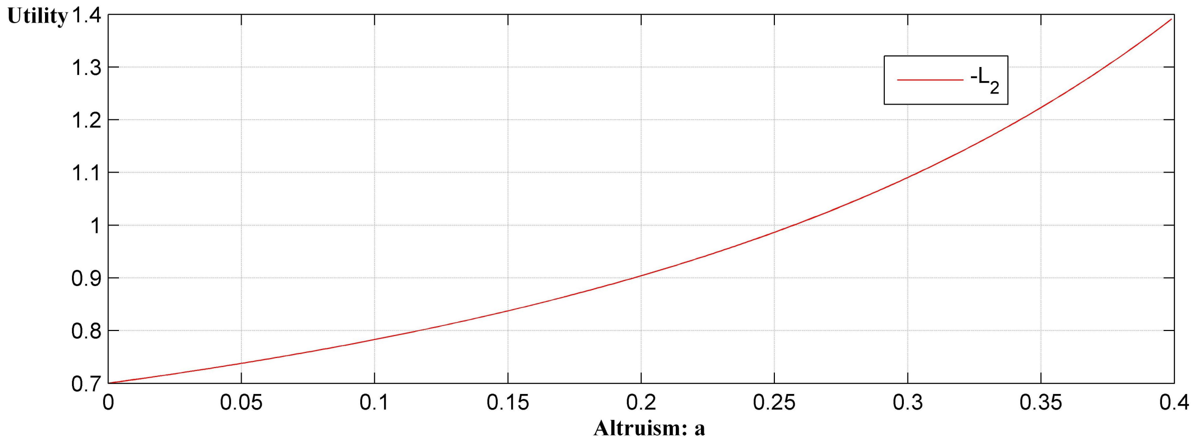


Figure 5.6: The utility of player 2 ($-L_2$) for various values of a.

The constants have values $c_1 = c_2 = A = 1, B = 0.1$. Figures 5.5, 5.5 illustrate the utilities and Figure 5.7 illustrates the production quantities of the players for various values of the amount of altruism a.

The optimal value for a corresponds to a player which is interested about the other player 27.5% and for itself 72.5%. The increase in the utility of player 1 is 9.35%.

5.5.3 The Rasouli-Teneketzis Mechanism

In this subsection we analyze an example of an electricity market having many competing producers. At first, a mechanism described in [RT14] is reviewed. An algorithm to converge to the Nash equilibrium is then proposed and tested numerically. The ability of the participants to cheat is then studied.

There are N energy producers $P_1, \ldots, P_N$. Each one of them produces an amount q_i of energy and proposes a price p_i. Let us denote the total demand by $D_0 > 0$. The cost of the production is given by a convex, increasing function $C_i : \mathbb{R}_0^+ \to \mathbb{R}_0^+$ with $C(0) = 0$. The utility of each player is given by:

$$u_i(q_i, t_i) = t_i - C_i(q_i), \tag{5.50}$$

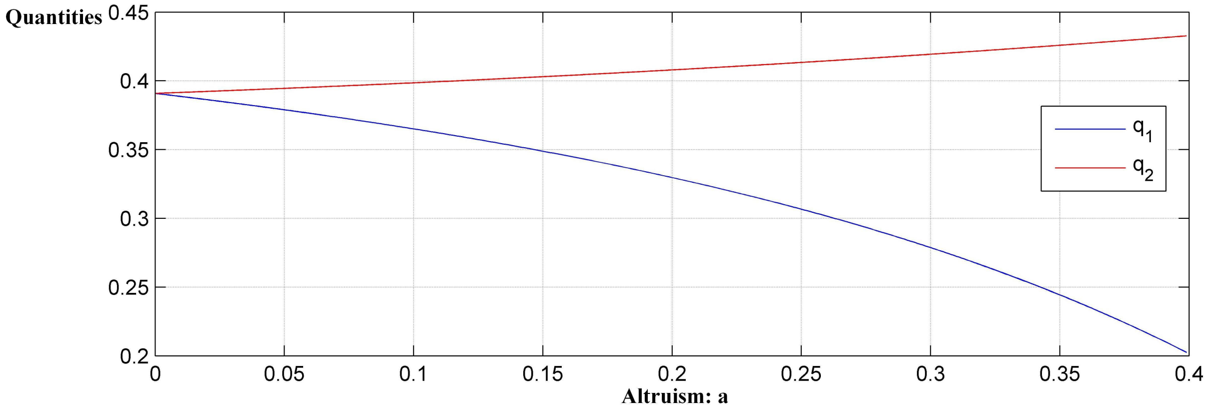


Figure 5.7: The quantities for various values of a

.

where t_i is the payment to the producer i by the mechanism. The payment t_i is given by:

$$t_i = p_{i+1}q_i - (p_i - p_{i+1})^2 - 2p_i\zeta^2, \tag{5.51}$$

where $p_{N+1} \equiv p_1$ and:

$$\zeta = D_0 - \sum_{i=1}^{N} q_i. \tag{5.52}$$

[RT14] study the nontrivial Nash equilibria of the game, i.e. the Nash equilibria such that it does not hold $q_1 = \cdots = q_N = 0$. It is shown that in any such equilibrium the demand is exactly covered, the proposed prices are identical and the total energy is produced with the smallest total cost (the sum of C_i is minimum). Furthermore a non-trivial Nash equilibrium always exits.

We then propose a very simple algorithm. This algorithm is found numerically to converge asymptotically to a non-trivial Nash equilibrium of the game. In what follows, we assume a very simple form for the costs of the players:

$$C_i = a_i q_i^2 \tag{5.53}$$

The energy production at the time step k is the best response i.e.:

$$q_i^k = arg \min_{q_i}\{p_{i+1}^{k-1} q_i - C_i(q_i^k)\} \tag{5.54}$$

The proposed prices evolve according to the following rule. If $\zeta < 0$, increase the prices by a constant step δ, i.e. $p_i^{k+1} = p_i^k + \delta_1$, where δ_1 is a positive scalar. If $\zeta \leq 0$, the new values of the proposed prices move towards the best responses:

$$p_i^{k+1} = (1-\delta_2)p_i^k + \delta_2 arg \min_{p_i}\{(p_i - p_{i+1}^k)^2 + 2p_i(\zeta^k)^2\}, \tag{5.55}$$

where $\delta_2 \in (0, 1)$.

It is not difficult to show that if the algorithm converges then it converges to a non-trivial Nash equilibrium of the game. To do so observe that the trivial Nash equilibrium is not a fixed

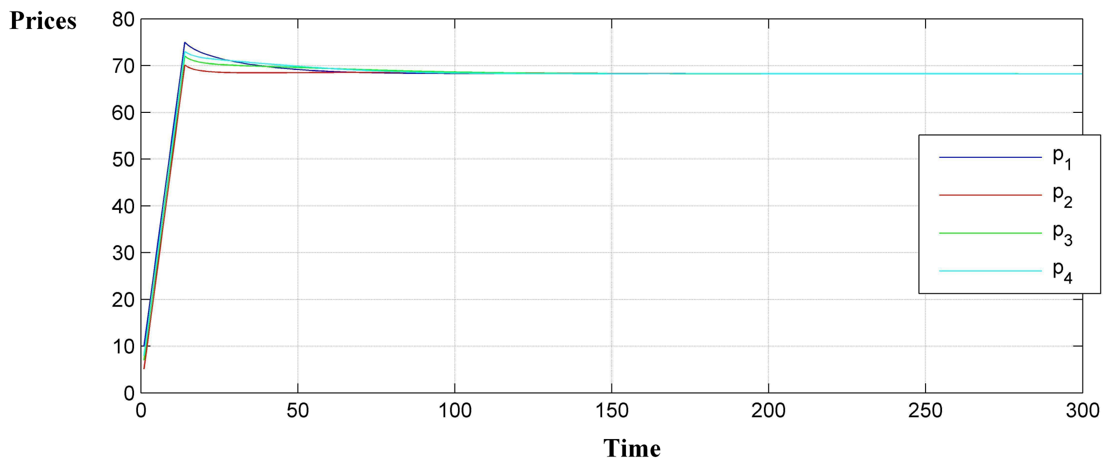


Figure 5.8: Prices evolution

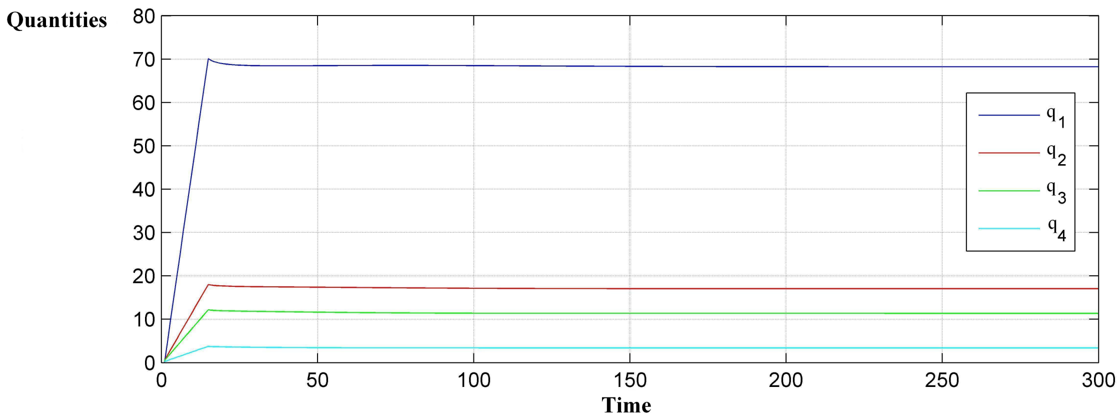


Figure 5.9: Production evolution

point of the proposed dynamics and that in any fixed point the energy productions q_i and the prices p_i are a best response of the actions of the other players.

Figures 5.8 and 5.9 illustrate the evolution of the energy production and the proposed prices in an example with four producers having parameters: $a_1 = 0.5, a_2 = 2, a_3 = 3, a_4 = 10$.

A very simple cheating strategy is then described. A cheating player i, implements the algorithm given by (5.54), (5.55) as if she had a false cost a_i^{pr}. The pretended type is updated rarely but periodically according to the rule:

$$a_i^{pr}(t+1) = a_i^{pr}(t) + \rho sat_{[-L,L]}\left(\frac{U^i(t) - U^i(t-1)}{a_i^{pr}(t) - a_i^{pr}(t-1)}\right). \tag{5.56}$$

A simple example with eight producers is illustrated in figures 5.10 and 5.11. We assume that only player 1 (having the smaller a_i) is cheating.

Figures 5.12 and 5.13 illustrate the situation when all the players are pretending. The price grows unbounded. So the mechanism in this case does not work at all.

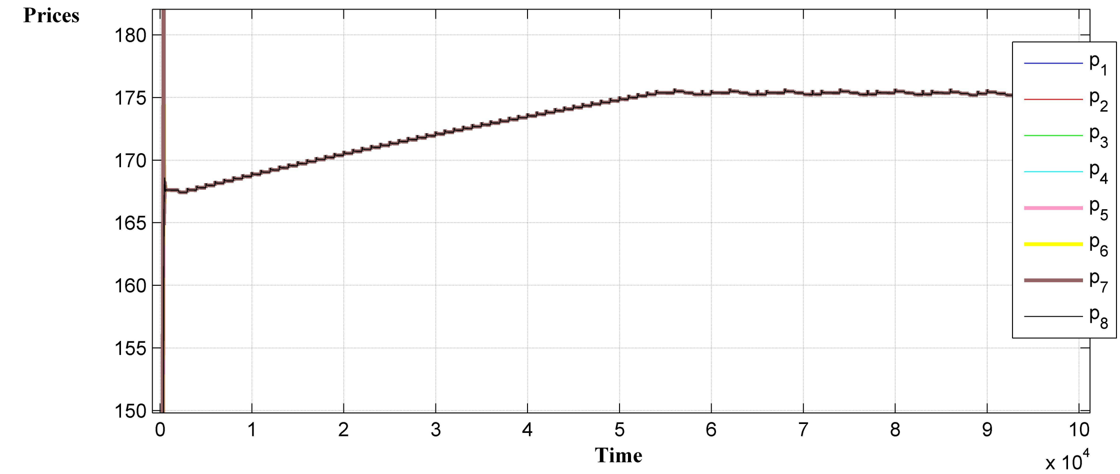


Figure 5.10: The evolution of the prices. Only player 1 is cheating

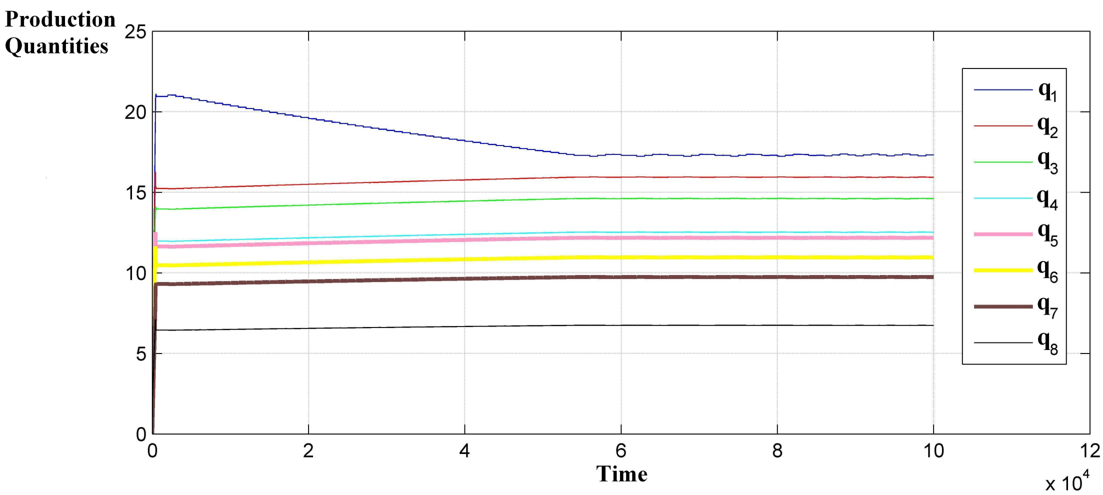


Figure 5.11: The evolution of energy production. Only player 1 is cheating

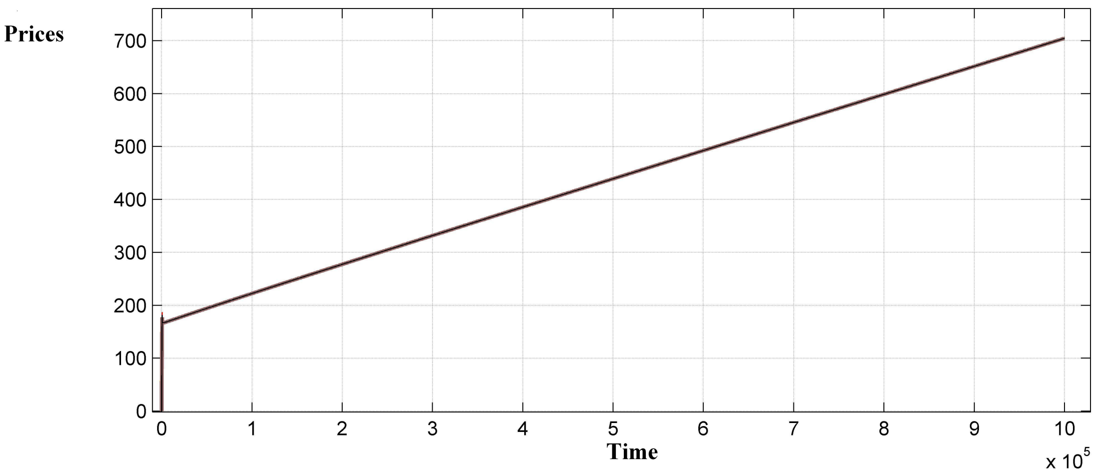


Figure 5.12: The evolution of the prices. All the players are pretending

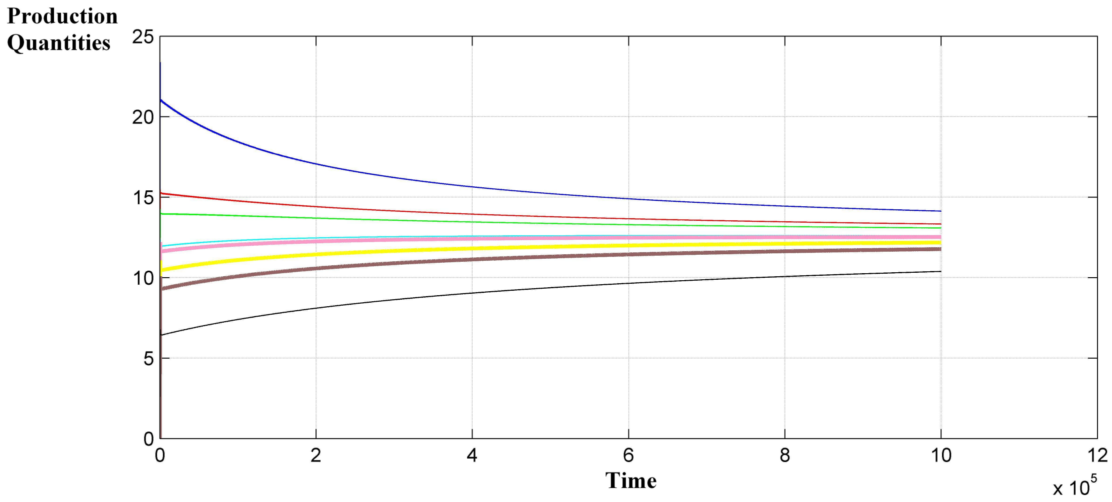


Figure 5.13: The evolution of the energy production. All the players are pretending

5.6 Conclusion

We studied the possibility of cheating the dynamic rules in dynamic/repeated game situations. Tow criteria were stated in order to assess dynamic rules. We then concentrated to the smaller class of pretenders' strategies and defined the auxiliary pretenders' game. It was shown that if there is enough uncertainty and there is only one player pretending, this player can get the same outcome as if she was a Stackelberg leader.

We then moved to the study of some special cases. At first we studied quadratic repeated games and the relations of pretending and leadership. We then moved to simple models of electricity markets. A Cournot duopoly was first analyzed and it was shown that pretending enhances the competition. We then moved to a Linear supply function model and it was shown that pretending is a cooperative action. Finally, the effects of cheating on the Rasouli-Teneketzis mechanism were examined. It was shown that if all the players are allowed to cheat in a certain manner, the prices diverge and thus the system does not work at all.

Κεφάλαιο 6

Επίλογος: Μελλοντικές Κατευθύνσεις

Στη διατριβή αυτή μελετήθηκαν διάφορα προβλήματα στη θεωρία των δυναμικών παιγνίων. Αναφορικά με το πρώτο μέρος, δείξαμε ότι η μελέτη των γραμμικών τετραγωνικών παιγνίων στα οποία η είσοδος των παικτών είναι τυχαία μπορεί να αναχθεί στη μελέτη συζευγμένων προβλημάτων βέλτιστου ελέγχου για MJLS. Παρόμοια προσέγγιση μπορεί να χρησιμοποιηθεί και σε άλλες περιπτώσεις τυχαίας συμμετοχής στο παιχνίδι. Για παράδειγμα, μπορεί παίκτες να μπαίνουν και να βγαίνουν τυχαία. Το τεχνικό μέρος που παρουσιάστηκε στο Κεφάλαιο 2, μπορεί επίσης να έχει πολλές εφαρμογές και επεκτάσεις, π.χ. συστήματα με τυχαίες χρονικές καθυστερήσεις, αβέβαια συστήματα κ.λπ.

Στο δεύτερο μέρος εξετάστηκαν παίγνια στα οποία οι παίκτες αλληλεπιδρούν πάνω σε ένα μεγάλο δίκτυο και ορίζονται έννοιες πιθανοτικής προσεγγιστικής ισορροπίας και πολυπλοκότητας. Υπάρχουν διάφορες ενδιαφέρουσες επεκτάσεις. Αρχικά, μπορούν να μελετηθούν εναλλακτικοί ορισμοί της πολυπλοκότητας. Μια άλλη κατεύθυνση είναι η μελέτη παιγνίων στα οποία η αλληλεπίδραση εξαρτάται από την κατάσταση του κάθε παίκτη. Επίσης, θα είχε ενδιαφέρον να εξεταστεί και η στοχαστική εξέλιξη του γραφήματος των αλληλεπιδράσεων και πιθανόν η εξάρτηση από μια αλυσίδα Markov. Σε αυτή την κατεύθυνση, θα μπορούσε να συνδυαστεί και με τα μοντέλα τη τυχαίας συμμετοχής παικτών, καθώς και με τη μελέτη των χρονικά εξελισσόμενων στοχαστικών δικτύων (π.χ. Scale Free Networks). Υπάρχουν τέλος και πολλές πιθανές εφαρμογές, όπως σε ευφυή δίκτυα, διάδοση ιδεών ή προϊόντων σε κοινωνικά δίκτυα κ.λπ.

Στο τρίτο μέρος εξετάστηκε η δυνατότητα εξαπάτησης καθώς και οι πιθανές της επιπτώσεις, στην περίπτωση που οι παίκτες εφαρμόζουν κάποιους δυναμικούς κανόνες. Είναι ενδιαφέρον να εξεταστούν οι διάφοροι δυναμικοί κανόνες ως προς τις δυνατότητες που δίνουν για εξαπάτηση και ενδεχομένως να σχεδιαστούν άλλοι δυναμικοί κανόνες που θα προσφέρουν λιγότερες τέτοιες δυνατότητες. Ένα άλλο μελλοντικό ερευνητικό βήμα μπορεί να είναι η εξέταση πιο πρακτικών μοντέλων των αγορών ηλεκτρικής ενέργειας, λαμβάνοντας υπ όψη τη γνώση που έχει ο κάθε παίκτης. Μια άλλη ερευνητική κατεύθυνση είναι να μελετηθούν δυναμικοί κανόνες εκμάθησης/προσαρμογής σε παίγνια στα οποία οι παίκτες αλληλεπιδρούν σε ένα μεγάλο δίκτυο. Σε αυτή την περίπτωση έχει ενδιαφέρον η σχέση που έχει η δυνατότητα εξαπάτησης με την 'κεντρικότητα' (centrality) που έχει ο κάθε παίκτης.

Bibliography

[AAEA10] A. P. Azad, E. Altman, and R. El-Azouzi. Routing games: From egoism to altruism. In *Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt), 2010 Proceedings of the 8th International Symposium on*, pages 528–537. IEEE, 2010.

[AB06] M. Aghassi and D. Bertsimas. Robust game theory. *Mathematical Programming*, 107(1-2):231–273, 2006.

[AKFJ95] H. Abou-Kandil, G. Freiling, and G. Jank. On the solution of discrete-time markovian jump linear quadratic control problems. *Automatica*, 31:765 – 768, 1995.

[AM11] A. P. Azad and J. Musacchio. Unilateral altruism in network routing games with atomic players. In *Network Games, Control and Optimization (NetGCooP), 2011 5th International Conference on*, pages 1–8. IEEE, 2011.

[Arv02] W. Arveson. *A Short Course on Spectral Theory*. Springer-Verlag, New York, 2002.

[ASY96] Kathleen T Alligood, Tim D Sauer, and James A Yorke. *Chaos*. Springer, 1996.

[AU05] I. Arribas and A. Urbano. Repeated games with probabilistic horizon. *Mathematical Social Sciences*, 50(1):39–60, 2005.

[Aum64] R. J. Aumann. Markets with a continuum of traders. *Econometrica*, 32:39–50, 1964.

[Bar12] M. Bardi. Explicit solutions of some linear-quadratic mean field games. *Networks and Heterogeneous Media*, 7(2):243–261, 2012.

[Bas73] T. Basar. On the relative leadership property of stackelberg strategies. *Journal of Optimization Theory and Applications*, 11(6):655–661, 1973.

[Bas74] T. Basar. A counterexample in linear-quadratic games: existence of nonlinear nash solutions. *Journal of Optimization Theory and Applications*, 14(4):425–430, 1974.

[Bas75] T. Basar. On the uniqueness of the nash solution in linear-quadratic differential games. *International Journal of Game Theory*, 5:65–90, 1975.

[BB90] V. J. Baston and F. A. Bostock. An evasion game on a finite tree. *SIAM J. Control Optim.*, 28:671–677, 1990.

[BB92] J. Bergin and D. Bernhardt. Anonymous sequential games with aggregate uncertainty. *Journal of Mathematical Economics*, 21(6):543–562, 1992.

[BB10] S. Bhattacharya and T. Basar. Graph-theoretic approach to connectivity maintenance in mobile networks in the presence of a jammer. In *Proceedings of the IEEE Conference on Decision and Control*, 2010.

[BBM09] M.-F. Balcan, A. Blum, and Y. Mansour. The price of uncertainty. In *Proceedings of the 10th ACM conference on Electronic commerce EC '09*, New York, NY, USA, 2009.

[BCAZ06] C. Ballester, A. Calvo-Armengol, and Y. Zenou. Who's who in networks: wanted the key player. *The Annals of Applied Probability*, 74(5):1403 – 1417, 2006.

[BD81] Jess Benhabib and Richard H Day. Rational choice and erratic behaviour. *The Review of Economic Studies*, pages 459–471, 1981.

[BD01] W. Brock and S. Durlauf. Discrete choice with social interactions. *Review of Economic Studies*, 68:235–260, 2001.

[Ber07] D. P. Bertsekas. *Dynamic Programming and Optimal Control, Vol 2*. Athena Scientific, Belmont Massachusetts, 2007.

[BGFB94] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan. *Linear Matrix Inequalities in System and Control Theory*. SIAM books, Philadelphia, 1994.

[Bha61] B. H. Bhaurucha. *On the stability of randomly varying systems*. PhD thesis, Univ. of California, Berkeley, 1961.

[BHO06] A. Bisin, U. Horst, and O. Ozgur. Rational expectations equilibria of economies with local interactions. *Journal of Economic Theory*, 127:74–116, 2006.

[Bil08] Patrick Billingsley. *Probability and measure*. John Wiley & Sons, 2008.

[Blu93] L. E. Blume. The statistical mechanics of strategic interaction. *Games Econ. Behav.*, 5:387–424, 1993.

[BO99] T. Basar and G. J. Olsder. *Dynamic Noncooperative Game Theory*. SIAM, 1999.

[BO11] A. Bisin and O. Ozgur. Dynamic linear economies with social interactions. *Levine's Working Paper*, 2011.

[Bol98] B. Bollobás. *Random graphs*. Springer, 1998.

[Bou05] E. K. Boukas. *Stochastic switching systems: analysis and design*. Birkhauser, Berlin, 2005.

[BP93] B. F. Beidas and G. P. Papavassilopoulos. Convergence analysis of asynchronous linear iterations with stochastic delays. *Parallel Computing*, 19(3):281 – 302, 1993.

[BS73] F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637 – 654, 1973.

[BS75] W. P. Blair and D. D. Sworder. Feedback control of a class of linear discrete systems with jump parameters and quadratic cost criteria. *International Journal of Control*, 21(5):833 – 841, 1975.

[BS80] Y. Balasko and K. Shell. The overlapping-generations model, i: The case of pure exchange without money. *Journal of Economic Theory*, 23:281–306, 1980.

[BSYY14] A. Bensoussan, J. Sung, P.p Yam, and S. P. Yung. Linear-quadratic mean field games. *arXiv preprint arXiv:1404.5741*, 2014.

[BTB+12] D. Bauso, H. Tembine, T. Basar, et al. Robust mean field games with application to production of an exhaustible resource. In *Proceedings of 7th IFAC symposium on robust control design, Aalborg, Denmark*, pages 454–459, 2012.

[BTV03] V. D Blondel, J Theys, and A. A. Vladimirov. An elementary counterexample to the finiteness conjecture. *SIAM Journal on Matrix Analysis and Applications*, 24(4):963–970, 2003.

[BZ11] T. Basar and Q. Zhu. Prices of anarchy, information, and cooperation in differential games. *J. Dynamic Games and Applications*, 1:50–73, 2011.

[CC83] Y. M. Chan and JB Cruz. Decentralized stochastic adaptive nash games. *Optimal Control Applications and Methods*, 4(2):163–178, 1983.

[CF93] O. L. V. Costa and M. D. Fragoso. Stability results for discrete-time linear systems with Markovian jumping parameters. *J. Math. Analysis Applic.*, 179(1):154 – 178, 1993.

[CF95] O. L. V. Costa and M. D. Fragoso. Discrete-time LQ-optimal control problems for infinite markov jump parameter systems. *IEEE Transactions on Automatic Control*, 40(12):2076– 2088, 1995.

[CFM05] O. L. V. Costa, M. D. Fragoso, and R. P. Marques. *Discrete-time Markovian Jump linear systems*. Springer-Verlag, London, 2005.

[CWC86] H. J. Chizeck, A. S. Willsky, and D. Castanon. Discrete-time markovian-jump linear quadratic optimal control. *International Journal of Control*, 43(1):213 – 231, 1986.

[DDM11] C. Daskalakis, A. G. Dimakis, and E. Mossel. Connectivity and equilibrium in random games. *The Annals of Applied Probability*, 21(3):987–1016, 06 2011.

[DE10] J. Kleinberg D. Easley. *Networks, Crowds, and Markets: Reasoning About a Highly Connected World*. Cambridge University Press, Cambridge, UK, 2010.

[Dud02] Richard M Dudley. *Real analysis and probability*, volume 74. Cambridge University Press, 2002.

[Fel60] A. A. Feldbaum. Dual control theory. *Automation and Remote Control*, 21(9):874–1039, 1960.

[FJAK96] G. Freiling, G. Jank, and H. Abou-Kandil. On global existence of solutions to coupled matrix riccati equations in closed-loop nash games. *Automatic Control, IEEE Transactions on*, 41:264–269, 1996.

[FKB12] P. Frihauf, M. Krstic, and T. Basar. Nash equilibrium seeking in noncooperative games. *Automatic Control, IEEE Transactions on*, 57(5):1192–1207, 2012.

[FKM90] D. Fudenberg, D. M. Kreps, and E. S. Maskin. Repeated games with long-run and short-run players. *Review of Economic Studies*, 57(4):555–573, 1990.

[FL98] D. Fudenberg and D. K. Levine. *The theory of learning in games*, volume 2. MIT press, 1998.

[FL02] Y. Fang and K. A. Loparo. Stochastic stability of jump linear systems. *IEEE Transactions on Automatic Control*, 47(7):1204 – 1208, 2002.

[FL09] Drew Fudenberg and David K Levine. Learning and equilibrium. *Annu. Rev. Econ.*, 1(1):385–420, 2009.

[FT91a] D. Fudenberg and J. Tirole. Game theory. *MIT Press Books*, 1, 1991.

[FT91b] D. Fudenberg and J. Tirole. *Game Theory*. MIT Press, 1991.

[FY03] Dean P Foster and H Peyton Young. Learning, hypothesis testing, and nash equilibrium. *Games and Economic Behavior*, 45(1):73–96, 2003.

[FY+06] Dean P Foster, H Peyton Young, et al. Regret testing: Learning to play nash equilibrium without knowing you have an opponent. 2006.

[GC10] R. A. Gupta and Mo-Yuen Chow. Networked control system: Overview and research trends. *Industrial Electronics, IEEE Transactions on.*, 57(7):2527 – 2535, 2010.

[GCF+12] S. A. Gabriel, A. J. Conejo, J. D. Fuller, B. F. Hobbs, and C. Ruiz. *Complementarity modeling in energy markets*, volume 180. Springer, 2012.

[GGJ+10] A. Galeotti, S. Goyal, M. O. Jackson, F. Vega-Redondo, and L. Yariv. Who's who in networks: wanted the key player. *The Review of Economic Studies*, 77(1):218–244, 2010.

[GJC10] J. Grossklags, B. Johnson, and N. Christin. The price of uncertainty in security games. *Economics of Information Security and Privacy*, pages 9–32, 2010.

[GLL10] O. Gueant, J. M. Lasry, and P. L. Lions. *Mean field games and applications*. Springer, Paris-Princeton Lectures on Mathematical Finance, 2010.

[GNLB14] A. Gupta, A. Nayyar, C. Langbort, and T. Basar. Common information based markov perfect equilibria for linear-gaussian games with asymmetric information. *arXiv preprint arXiv:1401.4786*, 2014.

[GS13] J. Ghaderi and R. Srikant. Opinion dynamics in social networks: a local interaction game with stubborn agents. In *American Control Conference (ACC), 2013*, pages 1982–1987. IEEE, 2013.

[GZCN09] G. Ghoshal, V. Zlatić, G. Caldarelli, and M. E. J. Newman. Random hypergraphs and their applications. *Physical Review E*, 79(6):066118, 2009.

[HAK10] S. P. Hsu, A. Arapostathis, and R. Kumar. On controlled markov chains with optimality requirement and safety constraint. *International Journal on Innovative Computing, Information and Control*, 6(6):2497–2511, 2010.

[HBSW07] L.S. Hu, T. Bai, P. Shi, and Z. Wu. Sampled-data control of networked linear control systems. *Automatica*, 43(5):903 – 911, 2007.

[HCM07] M. Huang, P. E. Caines, and R. P. Malhamé. Large-population cost-coupled lqg problems with nonuniform agents: Individual-mass behavior and decentralized ε-nash equilibria. *Automatic Control, IEEE Transactions on*, 52(9):1560–1571, 2007.

[HCM12] M. Huang, P. E. Caines, and R. P. Malhamé. Social optima in mean field lqg control: centralized and decentralized strategies. *Automatic Control, IEEE Transactions on*, 57(7):1736–1751, 2012.

[HKM02] J. Huang, I. Kontoyiannis, and S. P. Meyn. The ODE method and spectral theory of Markov operators. In *Stochastic Theory and Control*, Lecture Notes in Control and Information Sciences. Springer, Berlin, 2002.

[HLO82] Y. C. Ho, P. B. Luh, and G. J. Olsder. A control-theoretic view on incentives. *Automatica*, 18(2):167–179, 1982.

[HMC05] M. Huang, R. P. Malhame, and P. E. Caines. Nash equilibria for large-population linear stochastic systems of weakly coupled agents. In *Analysis, Control and Optimization of Complex Dynamical Systems*, chapter 9, pages 215–252. Springer, New York, 2005.

[HMC06] M. Huang, R. P. Malhame, and P. E. Caines. Large population stochastic dynamic games: closed-loop mckean-vlasov systems and the nash certainty equivalence principle. *Communications in Information and Systems*, 6(2):221–251, 2006.

[HS98] Josef Hofbauer and Karl Sigmund. *Evolutionary games and population dynamics*. Cambridge University Press, 1998.

[Hua01] K. Huang. *Introduction to statistical physics*. CRC Press, 2001.

[Hua10] M. Huang. Large-population l.q.g. games involving a major player: the nash certainty equivalence principle. *SIAM J. Control Optimization*, 48(5):3318–3353, 2010.

[Jac08] M. O. Jackson. *Social and Economic Networks*. Princeton University Press, Princeton, N.J, 2008.

[JC88] Y. Ji and H. J. Chizeck. Controllability, observability, and discrete-time markovian jump linear quadratic control. *International Journal of Control*, 48(2):481 – 498, 1988.

[JC90] Y. Ji and H. J. Chizeck. Jump linear quadratic gaussian control: Steady-state solution and testable conditions. *Control Theory Adv. Technol.*, 6(3):289 – 318, 1990.

[JPAK12] M. Jungers, G. P. Papavassilopoulos, and H. Abou-Kandil. Feedback nash strategy for games with player-dependent time horizons. In *15th IFAC workshop on Control Applications of Optimization*, Rimini, Italy,, September 13–16, 2012.

[JR88] B. Jovanovic and R. W. Rosenthal. Anonymous sequential games. *Journal of Mathematical Economics*, 17(1):77–87, 1988.

[JY99] S. Jorgensen and D. W. K. Yeung. Inter- and intragenerational renewableresource extraction. *Annals of Operations Research*, 88, 1999.

[JY05] S. Jorgensen and D. W. K. Yeung. An overlapping generations stochastic deferential game. *Automatica*, 41:69 –74, 2005.

[JZng] M. O. Jackson and Y. Zenou. Games on networks. In S. Zamir (Eds.) Young, editor, *Handbook of Game Theory, Vol. 4*. Elsevier Science, Amsterdam, forthcoming.

[Kan92] M. Kandori. Social norms and community enforcement. *Review of Economic Studies*, 59(1):63–80, January 1992.

[Kar75] A. F. Karr. Weak convergence of a sequence of markov chains. *Wahrscheinlichkeitstheorie und Verw. Gebiete*, 33:41–48, 1975.

[KH11] N. Karcanias and A. G. Hessami. System of systems and emergence part 1: Principles and framework. In *in Proceedings Fourth International Conference on Emerging Trends in Engineering & Technology*, 2011.

[Kha02] Hassan K Khalil. *Nonlinear systems*, volume 3. 2002.

[KL93] E. Kalai and E. Lehrer. Rational learning leads to nash equilibrium. *Econometrica: Journal of the Econometric Society*, pages 1019–1045, 1993.

[KLS01] M. J. Kearns, M. L. Littman, and S. P. Singh. Graphical models for game theory. In *Proceedings of the 17th Conference in Uncertainty in Artificial Intelligence*, UAI '01, 2001.

[KLY11] V. N Kolokoltsov, J.e Li, and W. Yang. Mean field games and nonlinear markov processes. *arXiv preprint arXiv:1112.3744*, 2011.

[KP] I. Kordonis and G. P. Papavassilopoulos. Games on large networks: Information and complexity.

[KP10] N. Ch. Kakogiannis and G. P. Papavassilopoulos. Games with players having long term and short term horizons. In *14th International Symposium of Dynamic Games*, Bannf, Alberta, Canada, July 19–23, 2010.

[KP11] I. Kordonis and G. P. Papavassilopoulos. A nash lqg game with an infinite horizon major player and many randomly entering minor players of different time horizons. In *Workshop of International Society of Dynamic Games, Padua, Italy*, 2011.

[KP12] I. Kordonis and G. P. Papavassilopoulos. Lq nash games with players participating in the game for random time intervals. In *5th International Symposium on Dynamic Games and Applications, Byšice, Czech Republic July 19-22*, 2012.

[KP13a] Ioannis Kordonis and George P Papavassilopoulos. Dynamic games among agents with partial information of the structure of the interactions graph: Decision making and complexity issues. In *Systems, Man, and Cybernetics (SMC), 2013 IEEE International Conference on*, pages 1709–1714. IEEE, 2013.

[KP13b] Ioannis Kordonis and George P Papavassilopoulos. Games on large random interaction structures: Information and complexity aspects. In *Decision and Control (CDC), 2013 IEEE 52nd Annual Conference on*, pages 1744–1749. IEEE, 2013.

[KP14a] I. Kordonis and G. P. Papavassilopoulos. On stability and lq control of mjls with a markov chain with general state space. *Automatic Control, IEEE Transactions on*, 59(2):535–540, 2014.

[KP14b] Ioannis Kordonis and George P Papavassilopoulos. Cheating in adaptive games motivated by electricity markets. In *Communications, Control and Signal Processing (ISCCSP), 2014 6th International Symposium on*, pages 194–197. IEEE, 2014.

[KP15] I. Kordonis and G. P. Papavassilopoulos. Lq nash games with random entrance. an infinite horizon major player and minor players of finite horizons. *Automatic Control, IEEE Transactions on*, PP(99), 2015.

[Kub85] C. S. Kubrusly. Mean square stability for discrete bounded linear systems in hilbert space. *SIAM Journal on Control and Optimization*, 23(1):19–29, 1985.

[KW82] D. Kreps and R. Wilson. Repeated games with long-run and short-run players. *Journal of Economic Theory*, 27:253–279, 1982.

[LHN09] M. Liu, D. W. C. Ho, and Y. Niu. Stabilization of markovian jump linear system over networks with random communication delay. *Automatica*, 45(2):416 – 421, 2009.

[Lig85] T. M. Liggett. *Interacting Particle Systems*. Springer, New York, 1985.

[LL07] J. M. Lasry and P. L. Lions. Mean field games. *Japanese Journal of Mathematics*, 2(1):229–260, 2007.

[LM99] D. Liberzon and A. S. Morse. Basic problems in stability and design of switched systems. *Control Systems, IEEE*, 19(5):59–70, 1999.

[LT11] S. Lasaulce and H. Tembine. *Game theory and learning for wireless networks: fundamentals and applications*. Academic Press, 2011.

[LZHMD14] F. L. Lewis, H. Zhang, K. Hengster-Movric, and Abhijit Das. *Cooperative control of multi-agent systems: optimal and adaptive design approaches*. Springer Publishing Company, Incorporated, 2014.

[ME10] M. Mesbahi and M. Egerstedt. *Graph theoretic methods in multiagent networks*. Princeton University Press, 2010.

[Mou98] G. V. Moustakides. Exponential convergence of products of random matrices, application to the study of adaptive algorithms. *International Journal of Adaptive Control and Signal Processing.*, 2(12):579– 597, 1998.

[MS78] J. W. Milnor and L. S. Shapley. Values of large games ii: Oceanic games. *Mathematics of Operations Research*, 3(4), 1978.

[MT93] S. P. Meyn and R. L. Tweedie. *Markov Chains and Stochastic Stability*. Springer-Verlag, London, 1993.

[Mye98] R. Myerson. Population uncertainty and poisson games. *International Journal of Game Theory*, 27(3), 1998.

[Mye13] R. B. Myerson. *Game theory*. Harvard university press, 2013.

[N$^+$50] J. F. Nash et al. Equilibrium points in n-person games. *Proceedings of the national academy of sciences*, 36(1):48–49, 1950.

[NAB09] K. C. Nguyen, T. Alpcan, and T. Basar. Stochastic games for security in networks with interdependent nodes. In *Proc. of Intl. Conf. on Game Theory for Networks (GameNets)*, 2009.

[NC13] M. Nourian and P. E. Caines. ε-nash mean field game theory for nonlinear stochastic dynamical systems with major and minor agents. *SIAM Journal on Control and Optimization*, 51(4):3302–3331, 2013.

[NCMH13] M. Nourian, P. E. Caines, R. P. Malhame, and M. Huang. Nash, social and centralized solutions to consensus problems via mean field control theory. *Automatic Control, IEEE Transactions on*, 58(3):639–653, 2013.

[NH12] S. L. Nguyen and M. Huang. Linear-quadratic-gaussian mixed games with continuum-parametrized minor players. *SIAM Journal on Control and Optimization*, 50(5):2907–2937, 2012.

[Nil98] J. Nilsson. *Real-Time Control Systems with Delays*. PhD thesis, Lund Institute of Technology, 1998.

[NW99] Mark EJ Newman and Duncan J Watts. Renormalization group analysis of the small-world network model. *Physics Letters A*, 263(4):341–346, 1999.

[OP88] G. J. Olsder and G. P. Papavassilopoulos. A markov chain game with dynamic information. *Journal of optimization theory and applications*, 59(3):467–486, 1988.

[OR94] M. J. Osborne and A. Rubinstein. *A course in game theory*. MIT press, 1994.

[Owe69] G. Owen. *Game Theory*. W.B. Saunders Company, 1969.

[PAKJ13] G. P. Papavassilopoulos, H. Abou-Kandil, and M. Jungers. A nash game with long-term and short-term players. In *52nd IEEE Conference on Decision and Control (invited)*, Firenze, Italy, December 10–13, 2013.

[Pan08] D. Panchenko. *18.175 Theory of Probability fall-2008*. Massachusetts Institute of Technology: MIT OpenCourseWare License: Creative Commons BY-NC-SA. http://ocw.mit.edu, 2008.

[Pap82] GP Papavassilopoulos. On the linear-quadratic-gaussian nash game with one-step delay observation sharing pattern. *Automatic Control, IEEE Transactions on*, 27(5):1065–1071, 1982.

[Pap86] G. P. Papavassilopoulos. Iterative techniques for the nash solution in quadratic games with unknown parameters. *SIAM journal on control and optimization*, 24(4):821–834, 1986.

[Pap88] G. P. Papavassilopoulos. Adaptive games. In *Stochastic processes in physics and engineering*, pages 223–236. Springer, 1988.

[Pap89a] G. P. Papavassilopoulos. Cooperative outcomes of dynamic stochastic nash games. In *IEEE Conf. Decision & Control*, pages 186–191, 1989.

[Pap89b] G. P. Papavassilopoulos. Learning algorithms for repeated bimatrix nash games with incomplete information. *Journal of optimization theory and applications*, 62(3):467–488, 1989.

[Pap12] G. P. Papavassilopoulos. University-students game. *Dynamic Games and Applications, Published Online DOI 10.1007/s13235-012-0066-3*, pages 1–32, 2012.

[Pap14] G. P. Papavassilopoulos. Limiting behavior of lq deterministic infinite horizon nash games with symmetric players as the number of players goes to infinity. *arXiv preprint arXiv:1403.3300*, 2014.

[PCJ80] G. P. Papavassilopoulos and J. B. Cruz Jr. Sufficient conditions for stackelberg and nash strategies with memory. *Journal of Optimization Theory and Applications*, 31(2):233–260, 1980.

[PMC79] G. P. Papavassilopoulos, J. V. Medanic, and J. B. Cruz. On the existence of nash strategies and solutions of riccati equations in linear quadratic games. *Journal of Optimization Theory and Applications*, 28:49–75, 1979.

[Rad80] R. Radner. Collusive behavior in noncooperative epsilon-equilibria of oligopolies with long but finite lives. *Journal of economic theory*, 22(2):136–154, 1980.

[Ros65] J. B. Rosen. Existence and uniqueness of equilibrium points for concave n-person games. *Econometrica: Journal of the Econometric Society*, pages 520–534, 1965.

[RT14] M. Rasouli and D. Teneketzis. Electricity pooling markets with strategic producers possessing asymmetric information ii: Inelastic demand. *arXiv arXiv: 1404.5539*, 2014.

[Sam58] P. Samuelson. An exact consumption loan model of interest with or without the social contrivance of money. *Journal of Political Economy*, 66(6), 1958.

[SBA99] P. Shi, E. K. Boukas, and R. K. Agarwal. Kalman filtering for continuous-time uncertain systems with Markovian jumping parameters. *IEEE Transactions on Automatic Control*, 44(8):1592 – 1597, 1999.

[SCJ73a] M. Simaan and J. B. Cruz Jr. Additional aspects of the stackelberg strategy in nonzero-sum games. *Journal of Optimization Theory and Applications*, 11(6):613–626, 1973.

[SCJ73b] M. Simaan and J. B. Cruz Jr. On the stackelberg strategy in nonzero-sum games. *Journal of Optimization Theory and Applications*, 11(5):533–555, 1973.

[Sel75] R. Selten. Reexamination of the perfectness concept for equilibrium points in extensive games. *International Journal of Game Theory*, 4:25–55, 1975.

[Sel01] R. Selten. What is bounded rationality. *Bounded rationality: The adaptive toolbox*, pages 13–36, 2001.

[SF07] G. Szabo and G. Fath. Evolutionary games on graphs. *Physics Reports*, pages 97–216, 2007.

[SH69a] A. W. Starr and Y. C. Ho. Further properties of nonzero-sum differential games. *Journal of Optimization Theory and Applications*, 3(4):207–219, 1969.

[SH69b] A. W. Starr and Y.C. Ho. Nonzero sum differential games. *Journal of Optimization Theory and Applications*, 3(3), 1969.

[SHPB12] Walid Saad, Zhu Han, H Vincent Poor, and Tamer Basar. Game-theoretic methods for the smart grid: An overview of microgrid systems, demand-side management, and smart grid communications. *Signal Processing Magazine, IEEE*, 29(5):86–105, 2012.

[Sim72] H. A. Simon. Theories of bounded rationality. *Decision and organization*, 1:161–176, 1972.

[Sky03] B. Skyrms. *Stag-Hunt Game and the Evolution of Social Structure*. Cambridge University Press, Cambridge, 2003.

[SLB08] Y. Shoham and K. Leyton-Brown. *Multiagent systems: Algorithmic, game-theoretic, and logical foundations*. Cambridge University Press, 2008.

[Smi82] John Maynard Smith. *Evolution and the Theory of Games*. Cambridge university press, 1982.

[SPG03] Y. Shoham, R. Powers, and T. Grenager. Multi-agent reinforcement learning: a critical survey. *Web manuscript*, 2003.

[Sri] K. Sridharan. A gentle introduction to concentration inequalities.

[SVP02] C. Skoulidas, C. Vournas, and G. P. Papavassilopoulos. Adaptive game modeling of deregulated power markets. 2002.

[Swo69] D. D. Sworder. Feedback control of a class of linear systems with jump parameters. *IEEE Transactions on Automatic Control*, 14(1):9 – 14, 1969.

[Tem11a] H. Tembine. Large-scale games in large-scale systems. In *Proceedings of the 5th International ICST Conference on Performance Evaluation Methodologies and Tools*, pages 9–17. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2011.

[Tem11b] H. Tembine. Large-scale games in large-scale systems. In *Proceedings of the 5th International ICST Conference on Performance Evaluation Methodologies and Tools*, ICST, Brussels, Belgium, Belgium, 2011.

[TH11] H. Tembine and M. Huang. Mean field difference games: Mckean-vlasov dynamics. In *Decision and Control and European Control Conference (CDC-ECC), 2011 50th IEEE Conference on*, pages 1006–1011. IEEE, 2011.

[TP85] M. Tu and G. P. Papavassilopoulos. On the informational properties of the nash solution of lqg dynamic games. *IEEE Transactions on Automatic Control*, 30:377–385, 1985.

[TZB+11] H. Tembine, Q. Zhu, T. Basar, et al. Risk-sensitive mean-field stochastic differential games. In *Proc. 18th IFAC World Congress*, 2011.

[War52] J. G. Wardrop. Road paper. some theoretical aspects of road traffic research. In *ICE Proceedings: Engineering Divisions*, volume 1, pages 325–362. Thomas Telford, 1952.

[WBVR05] G. Y. Weintraub, C L. Benkard, and B. Van Roy. Oblivious equilibrium: A mean field approximation for large-scale dynamic games. In *NIPS*, 2005.

[Wes01] D. B. West. *Introduction to graph theory*, volume 2. Prentice hall Upper Saddle River, 2001.

[Wit68] H. S. Witsenhausen. A counterexample in stochastic optimum control. *SIAM Journal on Control*, 6(1):131–147, 1968.

[Wit02] B. Wittenmark. Adaptive dual control. *Control Systems, Robotics and Automation, Encyclopedia of Life Support Systems (EOLSS), Developed under the auspices of the UNESCO*, 2002.

[WS98] Duncan J Watts and Steven H Strogatz. Collective dynamics of 'small-world'networks. *nature*, 393(6684):440–442, 1998.

[WZ12] B.-C. Wang and J. F. Zhang. Distributed control of multi-agent systems with random parameters and a major agent. *Automatica*, 48(9):2093–2106, 2012.

[You04] H. P. Young. *Strategic learning and its limits*. Oxford University Press, 2004.

[YP94] W. Y. Yang and G. P. Papavassilopoulos. Decentralised adaptive control in a game situation for discrete-time, linear, time-invariant systems. In *American Control Conference, 1994*, volume 3, pages 3394–3399. IEEE, 1994.

[YP11] D. W. K. Yeung and L. A. Petrosyan. Subgame consistent cooperative solution of dynamic games with random horizon. *Journal of Optimization Theory and Applications*, 150(1), 2011.

Σύντομο Βιογραφικό Σημείωμα: Ι. Κορδώνης

Ο Ι. Κορδώνης γεννήθηκε το 1986 και το 2004 εισήλθε στη σχολή ΗΜΜΥ του ΕΠΜ. Κατά τη διάρκεια των προπτυχιακών σπουδών του επέλεξε την κατεύθυνση Ηλεκτρονικής και Συστημάτων και είχε ιδιαίτερο ενδιαφέρον για τον Αυτόματο Έλεγχο. Στη διπλωματική του εργασία ασχολήθηκε με την εφαρμογή τεχνικών του μη γραμμικού ελέγχου σε βιοιατρικά προβλήματα.

Το 2009 εισήλθε στο διδακτορικό πρόγραμμα της σχολής ΗΜΜΥ του ΕΜΠ υπό την επίβλεψη του καθ. Γ.Π. Παπαβασιλόπουλου. Κατά τη διάρκεια των μεταπτυχιακών σπουδών του ενίσχυσε το θεωρητικό του υπόβαθρο παρακολουθώντας διάφορα μαθήματα. Ερευνητικά ασχολήθηκε με τη θεωρία των Δυναμικών Παιγνίων καθώς και ορισμένα προβλήματα Στοχαστικού ελέγχου.

Τα ερευνητικά του ενδιαφέροντα περιλαμβάνουν θεωρητικά θέματα στα Δυναμικά Παίγνια και το στοχαστικό έλεγχο και εφαρμογές σε αγορές ηλεκτρικής ενέργειας.