Εθνικο Μετσοβιο Πολυτεχνειο
Σχολη Ηλεκτρολογων Μηχανικων και Μηχανικων Υπολογιστων
Τομεας Τεχνολογιας Πληροφορικης και Υπολογιστων
Εργαστηριο Λογικης και Επιστημης Υπολογισμων

Ad hoc επίτευξη - και αξιολόγηση τεχνικών για επίτευξη - ευστάθειας σε διανομές αγαθών

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Στέλιος Δρακονταειδής

Επιβλέπων: Δημήτριος Φωτάκης
Καθηγητής Ε.Μ.Π.

Αθήνα, Δεκέμβριος του 2023

ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΤΕΧΝΟΛΟΓΙΑΣ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΥΠΟΛΟΓΙΣΤΩΝ

Ad hoc επίτευξη - και αξιολόγηση τεχνικών για επίτευξη - ευστάθειας σε διανομές αγαθών

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Στέλιος Δρακονταειδής

Επιβλέπων: Δημήτριος Φωτάκης
Καθηγητής Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την ... Νοεμβρίου 2023.

....	....	....
Δημήτριος Φωτάκης	Αριστείδης Παγουρτζής	Ευάγγελος Μαρκάκης
Καθηγητής Ε.Μ.Π.	Καθηγητής Ε.Μ.Π.	Καθηγητής Ο.Π.Α.

Αθήνα, Νοέμβριος του 2023

....

Στέλιος Δρακονταειδής

(Διπλωματούχος Ηλεκτρολόγος Μηχανικός & Μηχανικός Υπολογιστών Ε.Μ.Π.)

Περίληψη

Στην παρούσα διπλωματική εργασία εξετάζουμε διαφορετικούς τρόπους διανομής αγαθών σε δημοπρασίες χωρίς πληρωμές, με σκοπό την δημιουργία ευσταθών στιγμιοτύπων που επιτυγχάνουν ταυτόχρονα υψηλό κοινωνικό όφελος. Σχετικά με τον ορισμό της ευστάθειας, απλώς να αναφέρουμε αδρομερώς πως αναφερόμαστε σε διανομές αγαθών στους συμμετέχοντες στην δημοπρασία που καταλήγουν σε παρόμοια αποτελέσματα σε περιπτώσεις μοναδιαίων αποκλίσεων στα διανύσμα αξιολόγησης των αγαθών από τους παίκτες. Στο πλαίσιο της αναζήτησής μας, δοκιμάζουμε θεωρητικά αλλά και πειραματικά διάφορες ήδη γνωστές προσεγγίσεις, όπως τους κανόνες αναλογικής και αντίστροφα αναλογικής διανομής, την τεχνική εντροπικής κανονικοποίησης επί της συνάρτησης κοινωνικού οφέλους και την τεχνική της αντικατάστασης της αντικειμενικής συνάρτησης με την συνάρτηση κοινωνικού οφέλους Nash. Παράλληλα, πάλι κινούμενοι επάνω στην ίδια γραμμή έρευνας, προτείνουμε μια νέα κλάση αλγορίθμων διανομής ενός αντικειμένου, η οποία αποδεικνύουμε ότι πετυχαίνει μεγιστοποίηση κοινωνικού οφέλους σε μια οικογένεια γραμμικών προγραμμάτων που - θεωρώντας τους συμμετέχοντες στην δημοπρασία διατεταγμένους ως προς την αξιολόγησή τους για το αντικείμενο - επιβάλλουν διανομή του αντικειμένου για μεγιστοποίηση του κοινωνικού οφέλους με περιορισμούς στην διαφορά των διανεμηθέντων αξιών μεταξύ διαδοχικής αξίας συμμετεχόντων.

Λέξεις κλειδιά: Θεωρία παιγνίων, Δημοπρασίες, Μηχανισμοί Διανομής, Αλγοριθμική Δικαιοσύνη, Διαφορική Ιδιωτικότητα, Κανονικοποίηση, Κοινωνικό Οφελος Nash.

Abstract

In this dissertation we examine different ways to allocate items in auction settings without payments, in order to produce stable instances with high efficiency. On a high-level, with the term stability we mean a unilateral deviation on the agents' value vector, doesn't change the final allocation by a lot. In this context, we examine both theoretically and experimentally various - already known - approaches, and more specifically Proportional and Inverse Proportional Allocation Rules, Entropy Regularization (to the Social Welfare function) Technique and the maximization of the Nash Social Welfare function. At the same time, in the same line of work, we provide a new class of single-item allocation algorithms (named "Keep Consecutive Scales Tight") which produces - as we prove - optimal solutions to a family of linear programs that - considering an ordering of participants' valuations for the item - enforce constraints between the allocated values between participants with consecutive valuations.

Keywords: Game theory, Auctions, Allocation Mechanisms, Algorithmic Fairness, Differential Privacy, Regularization, Nash Social Welfare.

Ευχαριστίες

Αρχικά, θα ήθελα να ευχαριστήσω τον αναγνώστη που θα διαβάσει ολόκληρο το κείμενο των ευχαριστιών ή/και - πολύ περισσότερο - ολόκληρο το κείμενο της διπλωματικής! Πάνω σε αυτό, απλώς να πω παρενθετικά πως αξίζει να διαβάσει κανείς το πλήρες κείμενο γιατί μόνο έτσι θα "ανακαλύψει" τα διάσπαρτα αποσπάσματα που έγραφα όταν είχα πραγματικά όρεξη (συνήθως μετά από τρέξιμο ή τις πρώτες ώρες μελέτης όταν τύχαινε να είμαι μόνος μου στην ΜΟΠ) και τα οποία πραγματικά πρέπει σε κάποιο βαθμό να εντυπωθούν σε κάποιο συλλογικό υποσυνείδητο. 'Εκλεισε η παρένθεση.

Πρώτα και πάνω απ' όλα θέλω και οφείλω να ευχαριστήσω με όλη μου την ψυχή την μητέρα μου και την θεία μου για την υποστήριξή τους από την πρώτη στιγμή που είδα το φως του Ήλιου. Καταλαβαίνω ότι για πολλά χρόνια σας είχα "παγιδεύσει" σε μία κατάσταση άγχους και αναμονής. Ενώ πήγαινα προς - και τελικά έφτασα - ηλικίες που οι λιγότερο άξιοι (ή/και τολμηροί) βρίσκουν σύντροφο και συμβιβάζονται στο 9 με 5 και οι πραγματικά άξιοι διαπρέπουν, εγώ πάλευα με διάφορους εσωτερικούς μου δαίμονες και αυτοτιμωρούμουν μέχρι κάπως η ζωή να φτάσει στο μεταιχμιακό στάδιο του "συμβιβάσου ή πήγαινε να τα πάρεις όλα". Και αισθάνομαι πραγματικά πολύ άσχημα, γιατί αυτή η "μανιχαϊστική" προσέγγιση είναι ωραία για ταινία και ενδεχομένως και στην πραγματική ζωή, αλλά για κάποιον άνθρωπο ο οποίος είναι αυθύπρακτος, ζει οικονομικά ανεξάρτητα και δεν έχει από πίσω του ανθρώπους να δίνουν την ζωή τους και να ελπίζουν σε αυτόν. Μεγαλώνετε έναν πολύ δύσκολο άνθρωπο, με μεγάλη δυσκολία προσαρμογής σε κάθε μεγάλη μετάβαση στην ζωή του, που συχνά φλέρταρε με σχήματα μηδενισμού και μεγαλοϊδεατισμού χωρίς να υποστηρίζει με πράξεις έστω την εξέλιξη στην ζωή - ενώ μπορούσε, και αυτό είναι το χειρότερο -, και ο οποίος ταυτόχρονα είναι τόσο - μα τόσο - χαρισματικός, που θα ήταν κρίμα να τον κόβατε και να τον υποχρεώνατε να συμβιβαστεί. Και εκεί ακριβώς συνίσταται η παγίδα. Τελοσπάντων, όπως και να'χει, νομίζω πως η μεγάλη μετάβαση προς την μεγάλη μετάβαση (από το πανεπιστήμιο στο χάος της ζωής) σε μεγάλο βαθμό έγινε. Υποβέλτιστα, σε πάρα πολύ χρόνο, ψυχοφθόρα, κουραστικά, αλλά έγινε. Και τα 2 πιο σημαντικά πράγματα είναι ότι πρώτον είμαι πολύ πιο κοντά στο "πήγαινε να τα πάρεις όλα" από το "συμβιβάσου" και δεύτερον (και πιο σημαντικό), παρά τον χαμένο χρόνο, νομίζω πως έφτιαξα μια προσωπικότητα που εμπνέει την εκτίμηση και τον σεβασμό προσωπικοτήτων που εμπνέουν την δική μου εκτίμηση και σεβασμό.

Μιας και είμαστε σε αυτό το πλαίσιο εποπτικής θεώρησης της ζωής, μια σύντομη απεύθυνση (ναι Μπαμπινιώτη το είπα, τι θα κάνεις για αυτό;) προς τον μελλοντικό εαυτό μου. Η ζωή δεν είναι θεώρηση μέσα από σχήματα απολυτότητας. Ενδεχομένως, αυτό να βοηθάει όταν γίνεται στοχευμένα και - κυρίως - ελεγχόμενα για συγκεκριμένα χρονικά διαστήματα προκειμένου να πετύχεις κάτι ακραίο. Ή πιο σωστά. Για να μεγιστοποιήσεις τις πιθανότητες να πετύχεις κάτι ακραίο. Αλλά στην ζωή υφίστανται όλες οι ενδιάμεσες φασματικές αποχρώσεις. Η ζωή έχει καθημερινότητα. Η ζωή έχει tradeoffs. Μόνο αν "τουνάρεις" καλά τα πράγματα πετυχαίνεις. Δεν υπάρχει άλλος τρόπος. Και το πιο σημαντικό από όλα. Στο τέλος της μέρας, όπου και να φτάσεις, το μότο πρέπει να είναι αυτό που λέει ο Will.i.am : "I don't give a f***, that's my whole M.O.. I rock the whole globe with no problemo".

Στην συνέχεια, θα ήθελα να ευχαριστήσω ολόψυχα τον επιβλέποντά μου στην διπλωματική (και λίγο στην ζωή την τελευταία περίοδο) καθηγητή Δημήτρη Φωτάκη. Νομίζω πως η μεγαλύτερη φιλοφρόνηση (με την έννοια της περιγραφής των θετικών του χαρακτηριστικών και όχι με χροιά κολακείας) για εκείνον είναι απλά να περιγράψω την δουλειά του. Η (βασική) δουλειά του, λοιπόν, είναι να βοηθάει 10, 15, 20 (πλοτάρουμε στατιστικά για να δούμε πόσοι είναι ακριβώς στο παράρτημα στο τέλος της διπλωματικής) από τα καλύτερα μυαλά της Ελλάδας να βρουν τον δρόμο τους κάθε χρόνο. Και - μάλλον - το κάνει καλύτερα από κάθε άλλον καθηγητή στην Ελλάδα. Αυτό. Είμαι σχεδόν βέβαιος πως θα είστε με διαφορά ο πιο νευραλγικός άνθρωπος και παράγοντας για όλη την πορεία μου στα μονοπάτια της επιστήμης και μάλλον και της επαγγελματικής ζωής. Αν

βρεθεί άλλος άνθρωπος με σημαντικότερη συνεισφορά στην εξέλιξη του ''επαγγελματία'' Στέλιου, τότε μάλλον κάνω κάτι πολύ καλά (ή κάνω κάτι πολύ λάθος, γιατί ετεροπροσδιορίζομαι έντονα - όπως το δει κανείς!). Επίσης, - μάλλον - θα είστε και ο τέταρτος σημαντικότερος παράγοντας στην εξέλιξη του Στέλιου συνολικά ως άνθρωπος, μετά από την μαμά μου, την θεία μου και τον εγωισμό μου. Όπως σας είχα πει και κάποια στιγμή πριν τις καλοκαιρινές διακοπές (αλλά η υπερπροσαμογή στους ηλεκτρολόγους σας δυσκολεύει να το δείτε) σας σέβομαι και σας αγαπώ βαθιά. Ή πιο σωστά, σας αγαπώ και σας σέβομαι βαθιά. Παρόλο που με έχετε διαχειριστεί υποβέλτιστα αρκετές φορές λόγω των - περισσότερων από όσο θα έπρεπε - διαχειριστικών σας υποχρεώσεων και της διαφορετικότητάς μου στην προσέγγιση της γνώσης αλλά και στην προσέγγιση της προσέγγισης της γνώσης σε σχέση με τον δειγματοχώρο ανθρώπων με τους οποίους συναναστρέφεστε. Πάντως, εντέλει, νομίζω πως θα βρούμε ένα modus operandi που θα το κάνει να ''τσουλήσει'' και να ''τσουλήσει'' καλά. Πολύ καλά. Βασικά, για να είμαι ειλικρινής, το μοναδικό constraint νομίζω πως είναι ένα: Να γίνω πραγματικά σοβαρός άνθρωπος. Επίσης, αισθάνομαι το ηθικό χρέος να σας ζητήσω μια συγνώμη από καρδιάς γιατί είμαι ο φοιτητής που σας άγχωσε περισσότερο από κάθε άλλον, και αυτό άρχισε να φαίνεται όταν τα πράγματα πήγαιναν άσχημα και έγινε προφανές όταν τα πράγματα άρχισαν να πηγαίνουν καλά.
Ακόμη, θα ήθελα να ευχαριστήσω πολύ θερμά τον καθηγητή Άρη Παγουρτζή. Μετά τον κ.Φωτάκη, είστε ο άνθρωπος που με ενέπνευσε περισσότερο να ασχοληθώ με την Θεωρητική Πληροφορική (και σε συνδυασμό με το αβυσσαλέο μου μίσος για οτιδήποτε άλλο παίζει στην σχολή, κλείδωσε στο μυαλό μου το Corelab). Σας χρωστάω ένα καλό πέρασμα στις διαφάνειες της Κρυπτογραφίας προς το καλοκαίρι. Στην πραγματικότητα, δεν διάβασα ποτέ για το μάθημα (πέραν από το πρότζεκτ με τον Ουροβόρο), παρ'ότι μου άρεσε πάρα πολύ. Ο λόγος ήταν ένας κυρτός συνδυασμός επανάπαυσης λόγω Covid και επανάπαυσης επειδή εκείνη την περίοδο παρακολουθούσα και Θεωρία Ομάδων και πίστευα πως θα καλύψω τα τεχνικά κενά έμμεσα, διαβάζοντας για αυτό το μάθημα (κάτι που έγινε εν τέλει, αλλά ετεροχρονισμένα). Επίσης ήθελα να σας ζητήσω συγνώμη για την ασυνέπειά μου (και) σε ένα ερευνητικό πρότζεκτ. Δεν ήμουν έτοιμος - ψυχολογικά κυρίως - ούτε για αυτό, ούτε για την διπλωματική μου. Τώρα όμως είμαι (με εξαίρεση ένα recovery phase που σίγουρα θα χρειαστώ). Και από την στιγμή που σας χρωστάω ένα πρότζεκτ σημαίνει πως πρέπει να κάνουμε τουλάχιστον 2.

Φυσικά, θέλω πάρα πολύ - και όχι και οφείλω - να ευχαριστήσω τον καθηγητή Στάθη Ζάχο. Ο βασικός λόγος είναι επειδή με μύησε στα άδυτα της Επιστήμης της Θεωρίας Υπολογισμού και Πολυπλοκότητας και βέβαια ... Προφανώς όχι, έτσι!! Οι βασικοί λόγοι είναι 3. Πρώτον, διότι είστε η απαραίτητη νότα κοινωνικού αντικομφορμισμού στο Corelab. Είστε βασικά αυτό που λέει ο Καζαντζάκης ''δεν ελπίζω τίποτα, δεν φοβούμαι τίποτα, είμαι λεύτερος''. Παρότι σίγουρα ελπίζετε και ίσως φοβάστε, σίγουρα αυτά συμβαίνουν σε πολύ μικρότερο βαθμό σε σχέση με τα υπόλοιπα μέλη του Εργαστηρίου. Δεύτερον, είστε ένα παράδειγμα του πως πρέπει να διαχειρίζεται ένας σπουδαίος άνθρωπος το ''φθάσιμο στην Ιθάκη''. Η όρεξη με την οποία έρχεστε την τελευταία δεκαετία στο εργαστήριο μόνο και μόνο επειδή γουστάρετε (ντάξει και επειδή η άλλη επιλογή είναι η γυναίκα σας στο σπίτι), η ατάκα που μας είχατε πει πέρσι στο πρώτο μάθημα Αλγορίθμων, στην έξτρα ώρα του ΑΛΜΑ ''δεν είμαι τόσο σπουδαίος'' ως αντίδραση στο παρατεταμένο χειροκρότημα από τα μικρά όταν μπήκατε αργοπορημένα (κατά τα γνωστά!) στο αμφιθέατρο, η έγνοια σας να υπάρχει συνέχεια για το community στην διδασκαλία της Πολυπλοκότητας μέσα από τον κ.Ποτίκα και τόσα άλλα που δεν μπορώ να απαριθμήσω σε αυτό το στενό πλαίσιο δείχνουν έναν καταπληκτικό άνθρωπο. Μακάρι στα 70 μου να ξυπνάω το πρωί με διάθεση να ντυθώ Άγιος Βασίλης και να μοιράζω πτυχία στο κυλικείο των Ηλεκτρολόγων! Τέλος, - υψηλά συσχετισμένο με την αμέσως προηγούμενη πρόταση - να ξέρετε πως θα με βοηθήσετε όσο δεν φαντάζεστε στην παρουσίαση της διπλωματικής μου!
Αυτά τα λίγα (όχι!) με τους καθηγητές. Τώρα πάμε στα δύσκολα, λόγω της ασάφειας στα κριτήρια διάταξης. Στους συμφοιτητές! Η πρώτη επιλογή είναι εύκολη. Θέλω να ευχαριστήσω

τον έναν από τους 2 πιο ιδιαίτερους ανθρώπους που ξέρω, μαζί με μένα, Φοίβο (μόνο που σε σένα η ιδιαιτερότητα ενέχει πολύ λιγότερη αρνητική χροιά) για πολλά latex templates, 2 emails, ψύγματα (τουλάχιστον προς το παρόν) know-how και αρκετές δόσεις τρέλας, η οποία να ξέρεις είναι άκρως μεταδοτική. Για να ξεφύγω από τον αμιγώς τεχνοκρατικό λόγο που έχω υιοθετήσει από την αρχή, ας βάλω και μια αλληγορία. Όταν το σπίτι είναι ποτισμένο με οινόπνευμα, ένα σπίρτο χρειάζεται για να τυλιχθεί στις φλόγες. Επίσης, μέχρι το 2028 θα διδάσκεις στο MIT, αρκεί να μην το σκέφτεσαι. Θέλω πολύ, ακόμα, να ευχαριστήσω τον έτερο Καππαδόκη στου Ζωγράφου, Στέφανο, επειδή ήσουν πάντα (οκ, σχεδόν πάντα) το σίγουρο αυτί όταν ήθελα άμεσα να μοιραστώ κάτι που με απασχολούσε. Ο τύπος που θα σήκωνε το τηλέφωνο και θα βγαίναμε για καφέ ή για μπύρα στην Γαρδένια όταν ήμουν κουρασμένος, μπουχτισμένος, ξενερωμένος, ξενυχτησμένος ή ένας γραμμικός συνδυασμός αυτών. 'Ησουν η απαραίτητη τυχαιοποίηση για την ευρωστία στις σκέψεις μου. Και βέβαια, σ'ευχαριστώ και για μια σειρά από κουβαλήματα σε ότι άκυρο (η μάλλον απαραίτητο για την πληρότητα, τεχνική επάρκεια και δικαιώμα υπογραφής του ηλεκτρολόγου μηχανικού) εργαστήριο ηλεκτρονικής και τηλεπικοινωνιών έπαιζε στην σχολή. Στη συνέχεια, ευχαριστώ μέσα από την καρδιά μου τον Γιάννη και τον Νίκο, γιατί είναι πραγματικά 2 από τα πιο καλά και ξηγημένα παιδιά που ξέρω. Προσπαθήσατε πολύ να με βοηθήσετε πέρσι, σε μια περίοδο που κοίταζα για μεγάλο χρονικό διάστημα "κατάματα" την άβυσσο (αλλά ευτυχώς όχι αρκετά μεγάλο ώστε να με κοιτάξει κι αυτή, όπως είπε ο Νίτσε). Αυτό, δεν θα το ξεχάσω ποτέ, όπως και να πάει η ζωή. Ποτέ. Για να συμπληρωθεί η τετράδα των Ντάλτονς (χωρίς αυτό το "αστείο" να ενέχει υποδόριες δόσεις σχολιασμού συνυφασμένες με υψομετρικές διακυμάνσεις), θέλω να ευχαριστήσω πολύ τον Γιώργο. Γιώργαρε, θα το βγάλουμε το μαρούλι όλοι μαζί. Είτε θα μπαίνω στις κλήσεις ως comic relief και για να συζητάμε για κάνα paper, είτε θα μάθω web development και θα γράφω και κώδικα. Ελπίζω το πρώτο, φοβάμαι το δεύτερο. Πάντως, όπως και να'χει χρήσιμος θα είμαι. Φυσικά, θέλω πολύ να ευχαριστήσω την αγαπημένη μου Δανάη, κατά βάση για τον ίδιο λόγο που ευχαρίστησα τον Γιάννη και τον Νίκο. Μου έστελνες όταν φλέρταρα με την μη αναστρέψιμη αποτυχία στην ζωή και σ'ευχαριστώ πολύ για αυτό. Κυρία Μπάλλα, έπρεπε να επιμείνετε περισσότερο τότε, αλλά όπως και να'χει το εκτιμώ πραγματικά και τέτοιες συμπεριφορές δεν τις ξεχνάω. Σίγουρα πρέπει να ευχαριστήσω τον μοναδικό επιστήμονα αεκτζή που ξέρω, Άκη, για τόσες ωραίες (και τόσες όχι τόσο ωραίες) στιγμές που μοιραστήκαμε μαζί στην σχολή και στην ζωή. Η αγωνία αν θα την κάνει, επιτέλους, την μέτρηση το γ******* το οχτάρι, τα ξενύχτια για την εργασία στα Σήματα (που τελικά παρέδωσα με καθυστέρηση μόλις πέντε ετών), οι μπύρες που είναι απλώς η αφορμή για να δούμε μαζί τις τριάρες του Ολυμπιακού επί της Άεκ (ή και όχι, προχωράμε), οι ποιοτικές μας πολιτικές συζητήσεις που ξεκινούν με διαφωνία, συνεχίζουν με συμφωνία (για λίγο) και καταλήγουν σε διαφωνία και τόσα άλλα. Είσαι πολύ καλό παιδί και σε έχω στενοχωρήσει πολλές φορές με τα πιστόλια που ρίχνω. Συγνώμη!! Θα επανορθώσω να ξέρεις, μόλις γίνει negligible η πιθανότητα να γίνω ταμίας στην καφετέρια του ξαδέλφου μου. Από την περσινή φουρνιά των μικρών του κ.Φωτάκη, θα ήθελα να ευχαριστήσω την Κατερίνα και τον Ιάσονα. Κατερίνα, σ'ευχαριστώ πολύ για το ενδιαφέρον και νομίζω πως κάποια στιγμή πρέπει να κάνουμε (τουλάχιστον μία) συζήτηση. Ιάσονα, έδειξες κι εσύ ενδιαφέρον και επίσης με γλίτωσες από αρκετές αμήχανες διαδράσεις στο workshop του Κουτσουπιά (όχι ότι έλειψαν και αυτές βέβαια, κάθε άλλο). Σ'ευχαριστώ πολύ!

Ακόμη, θέλω να απολογηθώ, παρότι είναι κείμενο ευχαριστιών (αλλά και να ευχαριστήσω για την υπομονή σας να το παλέψετε - φυσικά με μεγάλη διακύμανση - μέχρι απλά να αποδεχτείτε ότι αυτός απλά δεν είναι σε φάση τώρα) στην Αλίκη, στην Άννα, στην Αργυρώ, στην Ευδοκία, στην Κατερίνα, στην Μαρία, στην Μελίνα, στην Ναταλία, στην Σοφία και στην Φρόσω (η διάταξη πάει με αλφαβητική σειρά για αποφυγή παρεξηγήσεων).

Οφείλω, τέλος, να ευχαριστήσω ολόκληρο το στελεχιακό δυναμικό των Google, Red bull, Monster, Stack Overflow, Special One, AB Βασιλόπουλος, Ρουμελιώτης και (τώρα τελευταία) Ονείρων Γεύσεις, Starbucks, ChatGPT, Resto και Coffee Island. Πραγματικά δεν θα έβγαινε πτυχίο χω-

ρίς εσάς! Δεν ξέρω αν μου έχετε στερήσει χρόνια ποιοτικής ζωής, αλλά μου έχετε γλιτώσει άπειρο χρόνο για ποιοτική φοιτητική ζωή, άσχετα αν - στην πραγματικότητα - ποτέ δεν την έκανα. Δεν φταίτε εσείς για αυτό.
Α επίσης, να απολογηθώ σε όσες φατσούλες ξέχασα να ευχαριστήσω. Αν ενδιαφέρεσαι πραγματικά για μένα, κάπως το ξέρω και σ'ευχαριστώ μέσα από την καρδιά μου για αυτό. Και να ξέρεις ότι είμαι από τους ανθρώπους που μια κακή πράξη ίσως την ξεχάσουν, ίσως και όχι. Αλλά μια καλή δεν θα την ξεχάσουν ποτέ. Για το κλείσιμο να απολογηθώ και στον αναγνώστη για το μακροσκελές κείμενο (στο είχα πει στην αρχή, αλλά αυτό δεν το κάνει καλύτερο). Είναι αλήθεια ότι δεν φημίζομαι για την λακωνικότητά μου και όσο η ζωή δεν με τιμωρεί για αυτό θα συνεχίσω να μην φημίζομαι για την λακωνικότητά μου. Και επίσης, πέραν της πολυλογίας μου, η αλήθεια είναι ότι το κείμενο των ευχαριστιών γράφτηκε με όχι πολλή όρεξη και κάπως γρήγορα.

Και όπως είπε ο Πασκάλ: "Συγνώμη που σου έγραψα τόσο μακροσκελές γράμμα, αλλά δεν βρήκα τον χρόνο να σου γράψω συντομότερο".

Contents

List of Figures

Κεφάλαιο α

Εισαγωγικό σημείωμα

Στα πλαίσια της παρούσας διπλωματικής εργασίας, επιχειρούμε μια αδρομερή αλλά όσο το δυνατόν πληρέστερη (δεδομένου ότι γίνεται στα πλαίσια της παρούσας διπλωματικής εργασίας) επισκόπηση σε διάφορες έννοιες και τεχνικές επίτευξης ευστάθειας, μέσα από το πρίσμα του σχεδιασμού μηχανισμών σε δημοπρασίες. Για να είμαστε πιο ακριβείς αλλά όχι τεχνικοί, μας ενδιαφέρει μόνο το πρώτο σκέλος του σχεδιασμού μηχανισμών σε δημοπρασίες, αυτό των κανόνων διανομής των αντικειμένων στους συμμετέχοντες. Για να αυξήσουμε το κίνητρο (αλλά και τον βαθμό τεχνικής δυσκολίας!) εισάγουμε κάποια έννοια δικαιοσύνης ανάμεσα στους συμμετέχοντες και επιχειρούμε να μετατοπίζουμε τις προς διανομή μάζες σε συμμετέχοντες που μεγιστοποιούν το κοινωνικό όφελος, τηρώντας ωστόσο πάντα τους περιορισμούς δικαιοσύνης στους οποίους έχουμε προδεσμευτεί.
Θα εμβαθύνουμε σε τεχνικά θέματα σύντομα, αλλά πριν από αυτό λίγη ιστορία.

Απλώς μια ιστορική αναφορά : Μια συνθήκη αντικρουόμενων συμφερόντων και ένα δίλημμα για την επιστήμη των αντικρουόμενων συμφερόντων και των διλημμάτων. Για λόγους πληρότητας (και επειδή νομίζουμε πως έχει ενδιαφέρον!) θεωρήσαμε σκόπιμο να κάνουμε μια πολύ σύντομη ιστορική αναφορά για την σύγχρονη Θεωρία Παιγνίων, καθώς και για τις καταβολές της στον αρχαίο κόσμο. Μια κλασσική διλληματική κατάσταση αναφερόμενοι σε ιστορικά θέματα έγκειται στο ποιος θα πρέπει να θεωρείται ο πατέρας της ιστορίας, ο Ηρόδοτος ή ο Θουκυδίδης; Από την μία ο Ηρόδοτος, έχει το πολύ ισχυρό επιχείρημα της παλαιότητας, καθώς ήταν εκείνος που πρωτοστάτησε στην καταγραφή των ιστορικών γεγονότων και στην προσπάθεια επεξήγησης των αιτιών που τα προκαλούν. Από την άλλη ο Θουκυδίδης έχει με το μέρος του την επιστημοσύνη, αφού ήταν ο πρώτος ιστορικός που απογαλακτίστηκε από την χρήση μύθων, μυθευμάτων και φημών κατά την αναζήτηση της ιστορικής αλήθειας.
Επεκτείνοντας τον αλληγορικό συλλογισμό, ας προσπαθήσουμε να εξετάσουμε ένα ζήτημα που ενέχει λιγότερους βαθμούς υποκειμενικότητας σε σχέση με την πατρότητα της Θεωρίας Παιγνίων. Ποιος παίζει τον ρόλο του Ηρόδοτου και ποιος του Θουκυδίδη στην Θεωρία Παιγνίων; Σχετικά με τον "Ηρόδοτο", ουσιαστικά αναφερόμαστε σε πρωτόλειες προσπάθειες μοντελοποίησης και βελτιστοποίησης αλληλεπιδράσεων μεταξύ λογικά σκεπτόμενων, στρατηγικών αντιπάλων. Αξιομνημόνευτες είναι οι αναφορές στον Κινέζο στρατηγό, συγγραφέα και φιλόσοφο Sun Tzu και στο βαβυλωνιακό Ταλμούδ. Σχετικά με τον Sun Tzu (544 - 496πΧ), προφανώς αναφερόμαστε στην σπουδαία πραγματεία του "Η Τεχνη του Πολέμου", όπου επιχειρείται η μοντελοποίηση της έννοιας των αντικρουόμενων συμφερόντων και μέσω αυτής δίδονται πρακτικές συμβουλές για στρατιωτικές καταστάσεις με αντικρουόμενα συμφέροντα. Εικάζεται ότι στα πιο στενά παιγνιοθεωρητικά

πλαίσια, σε αυτό το βιβλίο πρωτοεισήχθησαν οι έννοιες των minmax, κυρίαρχων και μεικτών στρατηγικών, ενώ δεν "κατάφερε" να ορίσει την έννοια των στρατηγικών ισορροπίας. Συνεπώς, ενώ άγγιξε την έννοια της συλλογιστικής σε επαναληπτικές διαδικασίες, αδυνατεί να δώσει την πλήρη διάσταση του φαινομένου (ιδίως της σύγκλισης σε ισορροπία). Ωστόσο, αν αναλογιστούμε το σπαργανώδες επίπεδο της ανθρωπότητας σε παιγνιοθεωρητικές και - εν γένει - μαθηματικές γνώσεις, παραμένει εκπληκτικό το βάθος σκέψης και το πλήθος των εννοιών τις οποίες κατάφερε να προσεγγίσει απλώς και μόνο εκτοξεύοντας βέλη διαίσθησης μέσα από την νοητική του φαρέτρα. Παραπέμπουμε τον ιστοριολάτρη (και με επαρκή ελεύθερο χρόνο!) αναγνώστη στην πολύ ενδιαφέρουσα και έξω από τα "τεχνικά ύδατα" στα οποία συνήθως κολυμπάμε, εργασία των Niou και Ordeshook [36]. Αναφορικά με το βαβυλωνιακό Ταλμούδ, πρόκειται για το βασικό κείμενο του ραββινικού Ιουδαϊσμού, μια συλλογή νόμων η οποία συντάχθηκε κατά κύριο λόγο τους 5 πρώτους αιώνες μΧ και αποτελεί την βάση του Ιουδαϊκού θρησκευτικού, ποινικού και αστικού κώδικα. Σε εκείνο το κείμενο συναντούμε μια πρωτόλεια αλλά και αυτοαναιρούμενη μορφή παιγνιοθεωτικού συλλογισμού, στο λεγόμενο πρόβλημα του συμβολαίου γάμου. Ειδικότερα, ένας άνδρας είναι παντρεμένος με 3 γυναίκες και έχει προσυμφωνήσει ότι σε περίπτωση θανάτου του, θα λάβουν 100, 200 και 300 αντίστοιχα. Αν η περιουσία του ανέρχεται στα 100, το Ταλμούδ ορίζει ισομοιρασμό. Αν ανέρχεται στα 300 ορίζει αναλογική διανομή, δηλαδή (50,100,150) ενώ αν ανέρχεται στα 200 ορίζει μια "παράλογη" μοιρασιά της μορφής (50,75,75). Παρά την "αδεξιότητα" αυτού του εγχειρήματος, παρατηρούμε πως από την αρχαιότητα η ανθρωπότητα αφουγκραζόταν θεμελιώδεις έννοιες δικαιοσύνης και αναλογικότητας.

Για τον "Θουκυδίδη" της Θεωρίας Παιγνίων, νομίζουμε πως τα ονόματα που έρχονται στο μυαλό με φυσικό τρόπο είναι 2, είναι σπουδαία και σχεδόν ασυναγώνιστα : John Von Neumann και John Nash! Από την μία, ο John Von Neumann ήταν ο άνθρωπος που θεμελίωσε την σύγχρονη Θεωρία Παιγνίων, με την ιδέα των μικτών στρατηγικών ισορροπίας σε παίγνια 2 παικτών μηδενικού αθροίσματος. Πέρα από την ιδέα της ύπαρξης, μεγάλη αξία έχει και η ιδέα της απόδειξης του, αφού η ιδέα της χρήσης του Brouwer's Fixed Point Theorem επηρρέασε την μετέπειτα εξέλιξη τόσο της Θεωρίας Παιγνιών όσο και της επιστήμης των Οικονομικών (αλλά και του έτερου θεμελιωτή John Nash). Επίσης, νευραλγικό είναι το έργο του "Theory of Games and Economic Behavior" ([43]) μαζί με τον Oskar Morgenstern, το οποίο αποτέλεσε το εφαλτήριο για την θεμελίωση της θεωρίας παιγνίων πάνω σε στέρεα μαθηματική βάση. Από την άλλη, ο - ελαφρώς - μεταγενέστερος John Nash θέτει και αυτός ισχυρή υποψηφιότητα για τον τίτλο του θεμελιωτή. Είναι, μάλλον ένας κυρτός συνδυασμός σπουδαιότητας έργου (κυρίως η απόδειξη ύπαρξης σημείων ισορροπίας σε πεπερασμένα παίγνια [33]) και έμπνευσης νέων ανθρώπων λόγω διασημότητας(ποιος δεν έχει δει την ταινία "Ένας Υπέροχος 'Ανθρωπος";) που του δίνει ισχυρό έρεισμα σε αυτό το άτυπο debate.

Κεφάλαιο β

Εκτεταμένη ελληνόγλωσση περίληψη

β.1 Εκτεταμένη ελληνόγλωσση περίληψη

Στο παρόν κεφάλαιο, θα επιχειρήσουμε να περιγράψουμε αδρομερώς - αλλά με πληρότητα - ολόκληρη την εικόνα της διπλωματικής μας εργασίας, τόσο σε επίπεδο τεχνικού υπόβαθρου όσο και πρωτότυπης παραχθείσας γνώσης, δηλαδή αποτελεσμάτων. Απλώς να διευκρινίσουμε πως θεωρούμε άσκοπο στο πλαίσιο της ελληνόγλωσσης περίληψης να αναλωθούμε σε τεχνικές λεπτομέρειες, επομένως θα επιχειρήσουμε να δώσουμε έμφαση στην διαισθητική προσέγγιση με διάσπαρτες αμιγώς τεχνικές πινελιές όταν κρίνεται πως ευθυγραμμίζονται με την κατεύθυνση της έμφασης στην διαισθητική προσέγγιση. Επίσης, μια διευκρίνισης, πιθανώς χρήσιμη, πιθανώς όχι. Για την κατασκευή ή/και παρουσίαση οποιουδήποτε μηχανισμού σε αυτό το κεφάλαιο (αλλά και γενικά σε όλη την διπλωματική εργασία) επικεντρωνόμαστε σε περιπτώσεις κλασματικής διανομής ενός αντικειμένου στους συμμετέχοντες στην διαδικασία διανομής (ή ισοδύναμα δημοπρασία χωρίς πληρωμές).

β.2 Αδρομερής περιγραφή των Αρχών για την προσέγγιση του Προβλήματος

β.2.1 Αρχή "Αναλογικού Κανόνα Διανομής"

Ως πρώτη αρχή/προσέγγιση προς παρουσίαση του προβλήματος της επίτευξης ευστάθειας κατά την διανομή αγαθών επιλέχθηκε η αρχή "Αναλογικού Κανόνα Διανομής". Οι βασικοί λόγοι είναι δύο. Πρώτον, υπήρξε η πρώτη μας επαφή και συνάμα το ερευνητικό μας έναυσμα για να ασχοληθούμε με μηχανισμούς διανομής αγαθών που πετυχαίνουν ευστάθεια (υπό την έννοια: παρεμφερείς είσοδοι δίνουν παρεμφερείς εξόδους). Συγκεκριμένα, η εργασία των Chawla και Jegadeesan [5] υπήρξε το εφαλτήριο για να κατανοήσουμε και στην συνέχεια να επινοήσουμε μηχανισμούς που μοιράζουν αγαθά με κάποια έννοια αναλογικότητας ως προς τις αξίες των παικτών για τα αγαθά αυτά. Επιπλέον, θεωρούμε πως η εν λόγω αρχή διανομής είναι η ευκολότερη εννοιολογικά από τις 3 και επομένως είναι εύλογο να ξεκινήσει κανείς παρουσιάζοντας αυτήν.
Ο πρώτος μηχανισμός που βασίζεται σε αυτήν την αρχή είναι ο Μηχανισμός Διανομής Ευθείας Αναλογικότητας. Η ιδέα είναι πραγματικά πολύ απλή: Διαμοίρασε αναλογικά προς τις αξίες ή προς μια γνησίως αύξουσα συνάρτηση των αξιών κλάσματα του αντικειμένου στους παίκτες. Ο μηχανισμός διανομής που βασίζεται σε αυτήν την αρχή παρουσιάζεται στον Αλγόριθμο 3. Το

βασικό πρόβλημα με αυτόν τον μηχανισμό είναι η ευρωστία του. Ειδικότερα, αφού κάθε παίκτης (μη-μηδενικής αξίας) θα λάβει μη-μηδενικό μέρος του προς διανομή αντικειμένου, όσο το πλήθος των παικτών τείνει προς το άπειρο, όλοι οι παίκτες θα λαμβάνουν ελάχιστο τμήμα του αντικειμένου και συνεπώς το συνολικό κοινωνικό όφελος θα μηδενίζεται. Για να επιλύσουν αυτό το πρόβλημα, οι Chawla και Jegadeesan επινόησαν έναν μηχανισμό ο οποίος αντί να μοιράζει αναλογικά προς τις αξίες, δίνει σε κάθε παίκτη ένα "εικονικό" αντικείμενο και "κόβει" από καθένα με ρυθμό ανάλογο προς το αντίστροφο της αξίας του (δηλ. σε υψηλής αξίας παίκτες κόβει "λίγο" και σε χαμηλής αξίας παίκτες κόβει "πολύ") μέχρις ότου το άθροισμα όλων των διανεμηθέντων κλασματικών μερών να ισούται με 1 (δηλ. να καταλήξουμε σε εφικτή λύση). Ο λόγος που αυτός ο μηχανισμός είναι πιο εύρωστος σε σχέση με τον ευθείας αναλογικότητας (υπό την έννοια ότι απογαλακτίζεται από το πλήθος των παικτών) είναι ότι κατά την διαδικασία "κοψίματος", δυνητικά αφήνει πολλούς παίκτες χαμηλής αξίας με μηδενικό κλάσμα του αντικειμένου. Για μια πιο τεχνική παρουσίαση αυτής της ιδέας, παραπέμπουμε τον ενδιαφερόμενο στον ψευδοκώδικα του Αλγορίθμου 4.

β.2.2 Αρχή "Επίλυσης Γραμμικού Προγράμματος με επιβολή περιορισμών δικαιοσύνης"

Η δεύτερη προσέγγιση του προβλήματός μας συνίσταται στην επίλυση ενός γραμμικού προγράμματος μεγιστοποίησης του κοινωνικού οφέλους (η "κλασσική" αντικειμενική συνάρτηση) με περιορισμό στις διαφορές των διαμοιραζόμενων αξιών μεταξύ παικτών διαδοχικής αξίας (θεωρώντας τους παίκτες διατεταγμένους ως προς την αξιολόγησή τους για το αντικείμενο). Για να γίνουμε πιο σαφείς, ουσιαστικά επιβάλλουμε ένα άνω φράγμα στην απόσταση της αξίας που θα πάρει ένα παίκτη σε σχέση με τον αμέσως "πλουσιότερο" και αμέσως "φτωχότερό" του (αν υπάρχουν). Αυτό το άνω φράγμα είναι λογικό να είναι μια συνάρτηση της απόστασης των αξιών των - προς επιβολή περιορισμών - παικτών. Για την ακρίβεια, η συνάρτηση καθορισμού των περιορισμών, έστω $g()$ πρέπει να πληρεί τα κάτωθι προαπαιτούμενα(desiderata) :

Προαπαιτούμενα για την συνάρτηση επιβολής περιοριοριςμών στις αξίες διαδοχικών παικτών $g()$:

- $g(0) = 0$, εννοώντας πως 2 παίκτες με την ίδια ακριβώς αξιολόγηση για το αντικείμενο πρέπει να λάβουν την ίδια ακριβώς αξία (εν προκειμένω και διανομή) από τον μηχανισμό.
- $g(\ 1\ ) = 1$, (υποθέτωντας κανονικοποίηση στο διάνυσμα αξιών, δηλαδή v_i 2 $[0, 1]$, 8 agent i 2 $[n]$) εννοώντας ότι παίκτες με αυθαίρετα διαφορετικές αξιολογήσεις για το αντικείμενο, θα πρέπει να μπορούν να λάβουν αυθαίρετα διαφορετικές αξίες από τον μηχανισμό.
- g '', εννοώντας ότι όσο αυξάνεται η διαφορά μεταξύ των αξιολογήσεων 2 (διαδοχικών) παικτών για το αντικείμενο, τόσο πιο πολύ θα πρέπει να "χαλαρώνει" ο περιορισμός σχετικά με την διαφορά στις αξίες που θα πρέπει να λάβουν από τον μηχανισμό.

Ένα γραμμικό πρόγραμμα εφοδιασμένο με μια συνάρτηση "ελέγχου" των περιορισμών $g()$ με τις παραπάνω ιδιότητες καταφέρνει να επιτύχει κάποιου ίδιου "δικαιοσύνη" εντός του ίδιου στιγμιοτύπου (ας την ονομάσουμε αυθαίρετα τοπική, επειδή ακριβώς αφορά το ίδιο στιγμιότυπο). Επίσης, θα μπορούσαμε να την χαρακτηρίσουμε και ως μη-αυτοαναφορική, διότι αφορά συγκρίσεις μεταξύ 2 (διαφορετικών) παικτών. Γενικά, ευελπιστούμε πως μέσα από αυτήν την μη-αυταναφορική, τοπική ευστάθεια που επιτυγχάνεται μέσα από εφικτές λύσεις του Γραμμκού Προγράμματος που μόλις

περιγράψαμε, καταφέρνουμε να πετύχουμε αυτοαναφορική, ολική ευστάθεια (δηλαδή αν μόνο ένας παίκτης αποκλίνει μοναδιαία από την αξιολόγησή του για το αντικείμενο, τα 2 στιγμιότυπα - το αρχικό και εκείνο που προέκυψε από την μοναδιαία απόκλιση από το αρχικό - θα καταλήξουν σε πολύ παρεμφερείς διανομές του αγαθού). Το πως διαπλέκονται ακριβώς αυτές οι 2 έννοιες ευστάθειας (κατά βάση το αν και πως η μία συνεπάγεται την άλλη) θα μπορούσε να είναι αντικείμενο μελλοντικής έρευνας, αφού δεν μελετήθηκε στα πλαίσια της παρούσας διπλωματικής.
Αυτό, όμως, που μελετήθηκε στα πλαίσια του παρόντος και που αποτελεί - μάλλον - τον βασικότερο κορμό της δουλειάς μας, είναι η βέλτιστη επίλυση τέτοιου είδους Γραμμικών Προγραμμάτων μέσω ενός αλγορίθμου που ονομάζεται "Διατήρησε Διαδοχικά Σκαλάκια κατά το δυνατόν Ψηλότερα"(ΔΔΣΨ). Αυτός ο αλγόριθμος - υποθέτωντας τους παίκτες ταξινομημένους σε φθίνουσα σειρά αξιολόγησης για το αντικείμενο - αρχικά μοιράζει κάποια τυχαία μάζα στον υψηλότερης αξιολόγησης παίκτη. Στην συνέχεια, "κατεβαίνει" προς τους αμέσως χαμηλότερης αξιολόγησης παίκτες διατηρώντας την ισότητα στους περιορισμούς στις αξίες (δηλαδή διαμοιράζει όσο το δυνατόν λιγότερη αξία στον διαδοχικά χαμηλότερης αξίας παίκτη ώστε να μην παραβιάζεται η ευστάθεια). Συνεχίζει την ίδια διαδικασίας διατήρησης της ισότητας στους περιορισμούς μέχρι να φτάσει στον πρώτο παίκτη που προκύπτει ότι πρέπει να λάβει αρνητική αξία - κάτι που φυσικά δεν επιτρέπεται. Σε αυτόν και στους χαμηλότερης αξιολόγησης παίκτες από αυτόν παίκτες, ο μηχανισμός μοιράζει 0. Εν συνεχεία, προσθέτει τα διαμειρασθέντα κλασματικά μέρη του αντικειμένου, και αν αθροίζονται στο 1, ο μηχανισμός βρήκε την βέλτιστη διανομή. Αν αθροίζονται σε ποσότητα μικρότερη του 1, ο μηχανισμός αυξάνει την διανεμηθείσα ποσότητα στον πρώτο παίκτη (με την λογική της δυαδικής αναζήτησης) και επαναλαμβάνει την διαδικασία. Παρόμοια λογική αν βρει ποσότητα μεγαλύτερη του 1. Κατά μία έννοια, αυτός ο μηχανισμός θυμίζει πολύ την λογική του μηχανισμού διανομής αντίστροφης αναλογικότητας, δεδομένου ότι πετυχαίνει υψηλό κοινωνικό όφελος - και ευρωστία ως προς το πλήθος των συμμετεχόντων - αφήνοντας τους χαμηλότερης αξιολόγησης παίκτες με μηδενική διανομή. Αναλυτική περιγραφή του μηχανισμού μπορεί κανείς να βρει στον Αλγόριθμο 5.

β.2.3 Αρχή "Τροποποίησε ή αντικατέστησε την αντικειμενική συνάρτηση κοινωνικού οφέλους"

Η τρίτη προσέγγιση του προβλήματος είναι η άρση των περιορισμών ευστάθειας και ταυτόχρονα η τροποποίηση ή πλήρης αντικατάσταση της αντικειμενικής συνάρτησης του κοινωνικού οφέλους. Ειδικότερα, για το πρώτο σκέλος της τρίτης προσέγγισης, δηλαδή αυτό της τροποποίησης της συνάρτησης κοινωνικού οφέλους, απλώς προσθέτουμε ως έξτρα όρους στην αντικειμενική συνάρτηση την αρνητική εντροπία των πιθανοτήτων προς διανομή (αν δούμε τα προς διανομή κλασματικά μέρη του - μοναδικού - αντικειμένου προς διανομή ως πιθανότητες). Στην γλώσσα των μαθηματικών, ο μετασχηματισμός που επιβάλλουμε παρατίθεται παρακάτω :

$$\sum_{i=1} x_i \cdot v_i \rightarrow \sum_{i=1} x_i \cdot v_i + \eta \sum_{i=1} x_i \log x_i$$

Αυτός ο μετασχηματισμός, έχει διττό ρόλο. Καταρχάς το έξτρα άθροισμα που επαυξάνει την αντικειμενική συνάρτηση την καθιστά ισχυρά κυρτή. ++++ Επίσης, η συγκεκριμένη επιλογή της εντροπίας ως κανονικοποιητή δεν είναι τυχαία. Η εντροπία είναι μια συνάρτηση που "ευνοεί" τις ομοιόμορφες κατανομές και "τιμωρεί" τις υπερσυγκεντρωτικές. Επομένως, αυτός ο κανονικοποιητής επιβάλλει μια αποκεντροποίηση των μαζών από τον υψηλότερης αξιολόγησης παίκτη (όπως "επιθυμεί" η συνάρτηση κοινωνικού οφέλους), καταφέρνοντας έτσι να επιφέρει κάποια δικαιοσύνη στην κατανομή (φυσικά ο βαθμός "δικαιοσύνης"/"αποσυγκέντρωσης" εξαρτάται από το διάνυσμα των αξιών και από την τιμή της παραμέτρου η).

Για το δεύτερο σκέλος της τρίτης προσέγγισης, αυτό της αντικατάστασης της συνάρτησης του κοινωνικού οφέλους, η επιλεχθείσα αντικειμενική συνάρτηση που παίρνει την θέση του είναι το κοινωνικό όφελος Nash. Το κοινωνικό όφελος Nash είναι ο γεωμετρικός μέσος όρος των προς διανομή αξιών, δηλαδή ο μετασχηματισμός της αντικειμενικής συνάρτησης που λαμβάνει χώρα είναι ο εξής :

$$\sum_{i=1} x_i \, v_i \rightarrow \left(\prod_{i=1} x_i \, v_i\right)^{\frac{1}{n}}$$

Η συγκεκριμένη αντικειμενική συνάρτηση μεγιστοποιείται σε κατανομές οι οποίες είναι αναλογικές και συχνά ομοιάζουν προς την ομοιόμορφη (φυσικά αυτό εξαρτάται από την μορφή του διανύσματος αξιών v). Απλώς ένα σύντομο "μπουστάρισμα διαίσθησης". Το κοίλο πρόγραμμα με αντικειμενική συνάρτηση το κοινωνικό όφελος Nash, μεταβλητές απόφασης πιθανότητες και ίσες αξίες/βάρη για όλους τους παίκτες μεγιστοποίεται όταν όλες οι πιθανότητες γίνουν ίσες (δηλ.$x_i = \frac{1}{n}, \forall i \in [n]$). Για - ελαφρώς - αναλυτικότερη επεξήγηση της έννοιας του κοινωνικού οφέλους Nash, παραπέμπουμε τον αναγνώστη στην υποενότητα β.6 της ελληνόγλωσσης περίληψης. Για κάτι πιο αναλυτικό και συγκεκριμένο, παραπέμπουμε στην υποενότητα 5.2 του αγγλικού κειμένου.

β.3 Εκθετικός Μηχανισμός

β.3.1 Εκθετικός Μηχανισμός από την σκοπιά της Διαφορικής Ιδιωτικότητας

Στην παρούσα υποενότητα, εξετάζουμε ένα ιδιαίτερα σημαντικό μηχανισμό διανομής αγαθών που συνδυάζει την επίτευξη ευστάθειας με αυτήν του υψηλού κοινωνικού οφέλους, τον Εκθετικό Μηχανισμό. Η σκοπιά από την οποία θα τον εξετάσουμε δεν είναι η μοναδική που θα παρουσιαστεί στο παρόν κείμενο και - μάλιστα - θα μπορούσαμε να πούμε πως είναι κάπως ιδιαίτερη, αφού δεν συνηθίζεται να "διαπλέκεται" στα ύδατα των διανομών αγαθών και - εν γένει - του σχεδιασμού μηχανισμών. Αυτήν της διαφορικής ιδωτικότητας. Πριν προχωρήσουμε στον φορμαλισμό, απλώς να δώσουμε πολύ συνοπτικά την διαίσθηση σχετικά με το τι είναι Διαφορική Ιδιωτικότητα και πως συσχετίζεται με την σχεδίαση μηχανισμών διανομής αγαθών. Η Διαφορική Ιδιωτικότητα, μέσα σε ένα "σκηνικό" ύπαρξης n παικτών με κάποια αξιολόγηση για ένα -προς κλασματική διανομή- αντικείμενο, αποτελεί μια ιδιότητα περιορισμού της επίδρασης που έχει ένας παίκτης στο τελικό αποτέλεσμα της διανομής. Πιο συγκεκριμένα, αλλά όχι τεχνικά, κρατώντας σταθερές τις αξιολογήσεις $n - 1$ παικτών, αν αλλάξουμε - αυθαίρετα πολύ - την αξιολόγηση ενός μόνο παίκτη, το τελικό αποτέλεσμα της διανομής δεν θα αλλάξει σημαντικά. Γιατί, όμως, αυτή η ιδιότητα να έχει κάποια αξία στην πραγματική ζωή, σε μια κατάσταση διανομής αγαθών ή/και δημοπρασίας με πληρωμές; Ένα πολύ καλό κίνητρο για την ένταξη (και) αυτής της έννοιας στο πλαίσιο μελέτης του σχεδιασμού μηχανισμών είναι σε δημοπρασίες που διεξάγονται σε πολλαπλά (χρονικά) στάδια και τα πονταρίσματα του ενός σταδίου, είναι συσχετισμένα με τα πονταρίσματα σε επόμενα στάδια. Τότε, προφανώς κάθε παίκτης επιθυμεί να διατηρηθούν ιδιωτικά τα πονταρίσματά του στον εκάστοτε γύρο και η Διαφορική Ιδιωτικότητα αποτελεί μια ιδιότητα που κινείται προς αυτήν την κατεύθυνση. Λίγο πιο φορμαλιστικά:

Definition β.3.1. (Πιθανοτικός Αλγόριθμος). Ένας πιθανοτικός αλγόριθμος A δημιουργεί μια αντιστοίχιση $A : D \rightarrow \Delta(R)$. Λαμβάνοντας ως είσοδο $d \in D$, A δίνει ως έξοδο με πιθανότητα $(A(d))_r$ το εξής : $A(d) = r$, για κάθε $r \in R$. Επίσης, ο χώρος πιθανοτήτων ορίζεται επί των ρίψεων νομισμάτων του A.

Definition β.3.2. (Διαφορική Ιδιωτικότητα). Ένας πιθανοτικός αλγόριθμος $A : X^n \; ! \; O$ είναι ϵ διαφορικά ιδιωτικός αν για κάθε - παίκτη - i, για κάθε είσοδο $x \; 2 \; X^n$, για κάθε μοναδιαία απόκλιση $x_i^0 \; 2 \; X$, και για όλα τα ενδεχόμενα εξόδου $E \quad O$:

$$\Pr[A(x) \; 2 \; E] \quad e \quad \Pr[A(x \;\; _i, x_i^0) \; 2 \; E]$$

Ας περιγράψουμε - πολύ αδρομερώς - ένα είδος δημοπρασίας, μόνο και μόνο για να δούμε τον Εκθετικό Μηχανισμό μέσα από ένα παράδειγμα. Η δημοπρασία αυτή, ονομάζεται Δημοπρασία Ψηφιακών Αγαθών. Σε αυτήν την δημοπρασία, έχουμε θεωρητικά άπειρο απόθεμα από ένα (ψηφιακό) αντικείμενο, και n υποψήφιους αγοραστές, καθένας εκ των οποίων επιθυμεί μόνο ένα αντίγραφο του αντικειμένου. Προφανώς, σε αυτήν την κατάσταση έχει νόημα να αναζητήσουμε την τιμή που μεγιστοποιεί το κέρδος, αφού η μεγιστοποίηση του κοινωνικού οφέλους γίνεται τετριμμένα (για παράδειγμα μοιράζοντας δωρεάν ένα αντίγραφο του αντικειμένου σε κάθε παίκτη). Επίσης, ας θεωρήσουμε πως επιθυμούμε να διατηρήσουμε και την ιδιότητα της Διαφορικής Ιδιωτικότητας. Εδώ, λοιπόν, εμπλέκεται ο Εκθετικός Μηχανισμός για να κάνει την επιλογή της τιμής με διαφορικά ιδιωτικό τρόπο που αποφέρει, ταυτόχρονα, και υψηλό κέρδος. Μια σημαντική και υψηλά συσχετισμένη με την Διαφορική Ιδιωτικότητα έννοια του Εκθετικού Μηχανισμού, είναι αυτή της ευαισθησίας του σε μοναδιαίες αποκλίσεις παικτών. Πιο συγκεκριμένα:

Definition β.3.3. (Ευαισθησία του Σκορ Απόδοσης του Εκθετικού Μηχανισμού). Η ευαισθησία του σκορ απόδοσης $s : X^n \quad O \; ! \; \mathrm{R}$ είναι η μέγιστη (απόλυτη) αλλαγή στην τιμή του σκορ που μπορεί να επιφέρει κάποια μοναδιαία απόκλιση. Σε πιο μαθηματική γλώσσα:

$$\Delta s = \max_{i\,2\,[n];\, x\,2\,X^n;\, x_i^0\,2\,X;\, o\,2\,O} j s(x, o) \quad s((x \;\; _i, x_i^0), o) j$$

Ο Εκθετικός Μηχανισμός, στην γενικότερη μορφή του, φαίνεται στον κάτωθι αλγόριθμο.

Algorithm 1 Εκθετικός Μηχανισμός

Input: Εύρος των αντικειμένων O, συνάρτηση σκορ αξίας $s : X^n \quad O \; ! \; \mathrm{R}$, παράμετρος ϵ.

1: **return** $o \; 2 \; O$ με πιθανότητα : $\dfrac{\exp(\frac{s(x;o)}{2 \quad (s)})}{\sum_{o\,2\,O} \exp(\frac{s(x;o)}{2 \quad (s)})}$

Ουσιαστικά, αυτό που επιτυγχάνει ο Εκθετικός Μηχανισμός είναι να δημιουργήσει "γειτονιές" τιμών που επιφέρουν πολύ υψηλό κέρδος και θα επιλεγούν όντως ως τελικές τιμές με πολύ υψηλή πιθανότητα. Επομένως, μέσω της τυχαιοποίησης επιτυγχάνει διαφορική ιδιωτικότητα και μέσω της δημιουργίας αυτών των "γειτονιών" επιτυγχάνει ευρωστία.
Η σημασία του Εκθετικού Μηχανισμού αποτυπώνεται γλαφυρά από τα κάτωθι 2 θεωρήματα.

Το πρώτο θεώρημα δηλώνει ότι ο Εκθετικό Μηχανισμός επιτυγχάνει την - πολυπόθητη - (ϵ)διαφορική ιδιωτικότητα.

Theorem β.3.1. *Ο Εκθετικός Μηχανισμός $M_E(x; s, O, \epsilon)$ είναι ϵ-διαφορικά ιδιωτικός, δηλαδή ισχύει:*

$$\frac{Pr[M_E(x; s, O, \epsilon) = o]}{Pr[M_E(x \;\; _i, x_i^0); s, O, \epsilon]} \quad e$$

Το δεύτερο θεώρημα επισημαίνει το γεγονός ότι ο Εκθετικός Μηχανισμός είναι αποδοτικός, υπό την έννοια ότι με υψηλή πιθανότητα δίνει στην έξοδο ένα υψηλού σκορ αξίας αντικείμενο.

Theorem β.3.2. *"Σταθεροποίησε" ένα εύρος αντικειμένων O. Έστω $OPT_{s;O} = \max_{o;O} s(o, O)$ το υψηλότερο σκορ που δίνει οποιοδήποτε αντικείμενο εντός του εύρους O. Επίσης, έστω $O = fo^{0}$ 2 $O : s(o^{0}, O) = OPT_{s;O}g$ το σύνολο των αντικειμένων εντός του εύρους O που πετυχαίνουν αυτό το σκορ. Τότε, ισχύει ότι :*

$$Pr[s(M_E(O)) \quad OPT_{s;O} \quad \frac{2\Delta}{\epsilon} \quad (\ln(\frac{jOj}{jO\ j}) + t)] \quad e^{\ t}$$

Έτσι, όπως φαίνεται στο τελευταίο θέωρημα αυτής της υποενότητας, η χρήση του Εκθετικού Μηχανισμού ως "επιλογέα τιμής" οδηγεί σε ένα μηχανισμό για την δημοπρασία ψηφιακών αγαθών που - πέραν της διαφορικής ιδιωτικότητας - εξασφαλίζει και πολύ υψηλό κέρδος:

Theorem β.3.3. *Για την δημοπρασία ψηφιακών αγαθών, υπάρχει μία ϵ-προσεγγιστική κυρίαρχης στρατηγικής φιλαλήθης δημοπρασία, που για το διάνυσμα αξιολογήσεων χειρότερης περίπτωσης v, δίνει το κάτωθι για τα έσοδα:*

$$Pr[\text{Έσοδα} \quad OPT(v) \quad O(\frac{\log n}{\epsilon})] \quad 0.99$$

β.3.2 Συνάρτηση κοινωνικού οφέλους με Εντροπική Κανονικοποίηση

Ένας διαφορετικός - πιο έμμεσος - τρόπος για να πάρουμε τον εκθετικό μηχανισμό, είναι επαυξάνοντας/κανονικοποιόντας την συνάρτηση κοινωνικού οφέλους ($\sum_{i=1}^{n} x_i \ v_i$) με την εντροπία της κατανομής πιθανότητας των μεταβλητών απόφασης x_i ($\sum_{i=1}^{n} x_i \ \log x_i$) πολλαπλασιασμένη με μια παράμετρο "ελέγχου" η, δημιουργώντας έτσι μια νέα αντικειμενική συνάρτηση προς μεγιστοποίηση και άροντας οποιονδήποτε άλλο περιορισμό ευστάθειας/"δικαιοσύνης". Ουσιαστικά, το πρώτο άθροισμα (κοινωνικό όφελος) τείνει να δημιουργήσει υπερσυγκεντρωτικές κατανομές, συσσωρεύοντας την προς διανομή μάζα στον υψηλότερης αξιολόγησης παίκτη. Από την άλλη, το δεύτερο άθροισμα (εντροπία) μεγιστοποιείται στην ομοιόμορφη κατανομή, επομένως τείνει να δημιουργήσει ισομοιρασμό στην προς διανομή μάζα. Συνεπώς, ανάλογα την τιμή του συντελεστή της εντροπίας (η), μπορούμε δυνητικά να παράγουμε κατανομές υψηλού κοινωνικού οφέλους, οι οποίες παράλληλα ικανοποιούν κάποια έννοια διακαιοσύνης/ευστάθειας.
Ένα αξιοσημείωτο γεγονός, είναι πως η βέλτιστη (μέγιστη) λύση για αυτήν την αντικειμενική συνάρτηση, είναι - πάλι - ο Εκθετικός Μηχανισμός, όπως φαίνεται και στο κάτωθι λήμμα:

Lemma β.3.4. *Η βέλτιστη λύση του προβλήματος μεγιστοποίησης:* $\max \sum_{i=1} x_i \ v_i \quad \eta \ (\sum x_i \log x_i)$ *είναι ο Εκθετικός Μηχανισμός, δηλαδή:*

$$x_i = \frac{e^{\frac{v_i}{}}}{\sum_{j=1} e^{\frac{v_j}{}}}$$

β.3.3 Αλγόριθμος ανανεώσεων πολλαπλασιαστικού βάρους

Απλώς μια σύντομη αναφορά, λόγω της μεγάλης χρησιμότητάς του, στον Αλγόριθμο ανανεώσεων πολλαπλασιαστικού βάρους. Αφορά online "σκηνικά", δηλαδή καταστάσεις όπου η είσοδος έρχεται τμηματικά σε διαδοχικές χρονικές στιγμές και πρέπει να ληφθεί μια -μη αναστρέψιμη- απόφαση σχετικά με την επιλεχθείσα ενέργεια την τρέχουσα χρονική στιγμή με γνώση μόνο του παρελθόντος και του παρόντος και όχι του μέλλοντος. Ο Αλγόριθμος αυτό που κάνει, εν συντομία, είναι το εξής: Αρχικά επιλέγει ομοιόμορφα μια ενέργεια. Σε κάθε χρονική στιγμή, μειώνει την πιθανότητα επιλογής των "κακών" ενεργειών εκθετικά, οπότε από ένα σημείο και μετά θα επιλέγει με υψηλή πιθανότητα σχεδόν αποκλειστικά "καλές" ενέργειες. Αποδεικνύεται πως η τελευταία κατανομή στην οποία συγκλίνει αυτός ο αλγόριθμος, είναι πάλι ο Εκθετικός Μηχανισμός. Για λόγους πληρότητας, παραθέτουμε σε ψευδοκώδικα στα Ελληνικά τον Αλγόριθμο παρακάτω :

Algorithm 2 Αλγόριθμος ανανεώσεων πολλαπλασιαστικού βάρους

Input: σύνολο δυνατών δράσεων $A = \{A_1, A_2, \ldots, A_n\}$, παράμετρος "εξερεύνησης-στόχευσης" $\eta \in (0,1)$.

Αρχικοποίησε το βάρος κάθε δράσης A_i ως εξής : $w_i^{(1)} \leftarrow 1$

for κάθε χρονικό βήμα $t = 1, 2, \ldots, T$ **do**

  Επίλεξε δράσεις με πιθανότητα ανάλογη των βαρών $w_i^{(t)}$,δηλ. μέσω της κατανομής:

$$p^{(t)} = \left(\frac{w_1^{(t)}}{\sum_{i=1} w_i^t}, \frac{w_2^{(t)}}{\sum_{i=1} w_i^t}, \ldots, \frac{w_n^{(t)}}{\sum_{i=1} w_i^t}\right)$$

  Παρατήρησε το διάνυσμα απώλειας ℓ_t για κάθε δράση

  Ανανέωσε το βάρος κάθε δράσης A_i ως εξής : $w_i^{(t+1)} \leftarrow w_i^{(t)} \cdot e^{\eta \ell_t^{A_i}}$

end for

β.4 Μηχανισμοί διανομής ευθείας και αντίστροφης αναλογικότητας

Στην παρούσα υποενότητα, περιγράφουμε εννοιολογικά και με ψευδοκώδικα τους 2 αλγορίθμους που συναποτελούν την υλοποίηση της αρχής "Αναλογικού Κανόνα Διανομής". Αυτοί οι αλγόριθμοι είναι ο μηχανισμός διανομής ευθείας αναλογικότητας και ο μηχανισμός διανομής αντίστροφης αναλογικότητας.

Αναφορικά με τον μηχανισμό ευθείας αναλογικότητας, η λογική όπως περιγράφηκε στην υπουποενότητα β.2.1 και επαναλαμβάνεται εδώ για λόγους πληρότητας είναι η κάτωθι: Διαμοίρασε σε κάθε παίκτη μέρος του αντικειμένου ανάλογο (μιας αύξουσας συνάρτησης) της αξίας του. Ο - πολύ απλός αυτός - μηχανισμός, περιγράφεται στον Αλγόριθμο 3 ακριβώς από κάτω.

Algorithm 3 Μηχανισμός Διανομής Ευθείας Αναλογικότητας παραμετροποιημένος από την συνάρτηση $g()$

Input: Συνάρτηση $g() : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ συνεχής, υπερπροσθετική (δηλ. $g(x) + g(y) \leq g(x+y)$), αύξουσα συνάρτηση. Αξίες $v_1, \ldots, v_n$.

for $i \in [n]$ **do**

  Θέσε $a_i = \frac{g(v_i)}{\sum_{j=1} g(v_j)}$

end for

return a.

Το πρόβλημα με τον παραπάνω αλγόριθμο είναι ότι η απόδοσή του εξαρτάται από το πλήθος των συμμετεχόντων στην διαδικασία διανομής, έστω n. Όταν $n \rightarrow \infty$, τότε το κοινωνικό όφελος $\rightarrow 0$, αφού κάθε (μη-μηδενικής αξιολόγησης) παίκτης θα λάβει κάποιο κλασματικό μέρος του αντικειμένου, οπότε ουσιαστικά κάθε παίκτης (ακόμα και ο υψηλότερης αξιολόγησης) θα λάβει ελάχιστο μέρος του αντικειμένου και η συνεισφορά του στο κοινωνικό όφελος θα τείνει στο 0. Το πραγματικό πρόβλημα αυτού του μηχανισμού που "κονιορτοποιεί" την ευρωστία του είναι το γεγονός ότι δίνει μη-μηδενικό κλάσμα του αντικειμένου σε πολύ χαμηλής αξιολόγησης παίκτες, ενώ θα μπορούσε να μεταφέρει αυτές τις μάζες σε υψηλότερης αξίας παίκτες, χωρίς να παραβιάζει κάποια - εύλογα - κριτήρια "δικαιοσύνης" (αφού είναι σαφές ότι όσο χαμηλότερης αξιολόγησης είναι ο παίκτης, τόσο λιγότερη αξία χρειάζεται για να ικανοποιηθεί). Επομένως, η βασική ιδέα που οδηγεί στην μετάβαση στον πιο "έξυπνο" και κατ'επέκταση εύρωστο αλγόριθμο είναι να αφήνουμε - δυνητικά - τους χαμηλότερης αξιολόγησης παίκτες με μηδενική διανομή αγαθού.

Ας το δούμε πιο αναλυτικά. Σε "υψηλό επίπεδο" η λογική του νέου αλγορίθμου - αντίστροφης αναλογικότητας - είναι αντί να μοιράζουμε το αντικείμενο ανάλογα με μια αύξουσα συνάρτηση των αξιών των παικτών, να "κόβουμε" ανάλογα μιας φθίνουσας συνάρτησης των αξιών τους. Πιο συγκεκριμένα, όπως περιγράφεται πιο τεχνικά στον Αλγόριθμο 4, ο μηχανισμός αρχικά μοιράζει σε κάθε παίκτη ένα "εικονικό" αντίγραφο ολόκληρου του - προς διανομή - αντικειμένου και ξεκινάει να "κόβει" από καθένα ένα μέρος του, με ρυθμό ανάλογο κάποιας φθίνουσας συνάρτησης της αξιολόγησής του. Όταν κάποιος παίκτης κατά την διαδικασία των συνεχών "κοψιμάτων" βρεθεί με αρνητικό μερίδιο στην διανομή, τότε χαρακτηρίζεται ως μη ενεργός, λαμβάνει - οριστικά - μηδενικό μερίδιο του αντικειμένου και τα "κοψίματα" συνεχίζονται κατά τον ίδιο τρόπο για όλους τους υπόλοιπους. Η διαδικασία σταματά όταν το συνολικό άθροισμα των διαμοιραζόμενων μεριδίων του αντικειμένου αθροίζεται στο 1 (δηλαδή όταν φτάσουμε για πρώτη φορά σε εφικτή λύση/διαμοιρασμό).

Algorithm 4 Μηχανισμός Διανομής Αντίστροφης Αναλογικότητας παραμετροποιημένος από την συνάρτηση $g()$

Input: Συνάρτηση $g() : \mathbb{R}_{\geq 0} \to (0, 1]$ με $g(0) = 1$ και $\lim_{x \to \infty} g(x) = 0$, φθίνουσα συνάρτηση. Αξίες $v_1, \ldots, v_n$.

Αρχικιποίησε $a_i = 0$ για $1 \leq i \leq n$

Ταξινόμησε τις αξίες ώστε $v_1 \leq v_2 \leq \ldots \leq v_n$.

if $v_n = 0$ **then**

  Θέσε $a_i = \frac{1}{n}$ για όλους τους δείκτες $1 \leq i \leq n$.

  return a.

end if

Αρχικοποίησε $s = \min(\{i \in [n] | v_i > 0\})$.

while $1 - \frac{(n-s)\, g(v_s)}{\sum_{j=s} g(v_j)} \leq 0$ **do**

  $s{+}{+}$

end while

for $i \geq s$ **do**

  Θέσε $a_i = 1 - (n - s) \cdot \frac{g(v_i)}{\sum_{j=s} g(v_j)}$

end for

return a.

β.5 Μηχανισμός διανομής "Διατήρησε Διαδοχικά Σκαλάκια κατά το δυνατόν Ψηλότερα"

Όπως αναφέραμε και στην αδρομερή περιγραφή των αρχών για την προσέγγιση του προβλήματος και συγκεκριμένα στην Αρχή "Επίλυσης του Γραμμικού Προγράμματος με επιβολή περιορισμών δικαιοσύνης", το βασικότερο μέρος της δουλειάς μας στα πλαίσια της παρούσας διπλωματικής εργασίας συνίσταται στον σχεδιασμού ενός νέου μηχανισμού διανομής αγαθών (για την περίπτωση διαμοιρασμού ενός αγαθού - Single-Item case), που τον ονομάζουμε "Διατήρησε Διαδοχικά Σκαλάκια κατά το δυνατόν Ψηλότερα"(ΔΔΣΨ). Η αφορμή δημιουργίας και ο λόγος ύπαρξης του συγκεκριμένου μηχανισμού διανομής εντάσσονται στα πλαίσια της προσπάθειάς μας να επιλύσουμε με βέλτιστο τρόπο το κάτωθι γραμμικό πρόγραμμα:

LP1(Single-Item)

$$\max \sum_{i\,2\,N} x_i \quad v_i$$

$$\text{s.t.} \sum_{u\,2\,N} x_i = 1$$

$$x_i \quad v_i \qquad x_{i+1} \quad v_{i+1} \qquad g(v_i, v_{i+1}), 8 \text{ agent } i\ 2\ [0, \ldots, N \qquad 1],$$

$$x_i\ 2\ [0, 1], 8 \text{ agent i } 2\ N$$

Σε "υψηλό-επίπεδο" η λογική του αλγορίθμου που προτείνουμε είναι η εξής: Προσπάθησε να ευνοήσεις όσο το δυνατόν περισσότερο τους υψηλότερης αξιολόγησης παίκτες, μην παραβιάζοντας τους περιορισμούς δικαιοσύνης και μην αφήνοντας μάζα χωρίς διανομή. Για λόγους πληρότητας θα υποπέσουμε στο ολίσθημα της επαναληψιμότητας και θα διατυπώσουμε σε πιο τεχνικό επίπεδο πως δουλεύει ο μηχανισμός μας πριν δώσουμε τον ψευδοκώδικα, κάτι που κάναμε και στην υποενότητα β.2.2. Αυτό που κάνει, λοιπόν, ο μηχανισμός μας είναι το εξής: Διατάσσει τους παίκτες σε φθίνουσα σειρά (ως προς την αξιολόγηση τους για το αντικείμενο). Αρχικά, μοιράζει κάποια τυχαία μάζα του - προς διανομή - αντικειμένου στον υψηλότερης αξιολόγησης παίκτη. Στην συνέχεια, δίνει στον αμέσως χαμηλότερης αξιολόγησης παίκτη μάζα τόση ώστε να ικανοποιηθεί ο μεταξύ τους περιορισμός με ισότητα (δηλαδή την ελάχιστη δυνατή). Συνεχίζει κατά τον ίδιο τρόπο στους επόμενους παίκτες, μέχρι να φτάσει σε κάποιον που "θα έπρεπε" να λάβει αρνητικό μέρος του αντικειμένου ή να τελειώσουν οι παίκτες. Στο πρώτο ενδεχόμενο, κρατάει αυτόν τον παίκτη όπως και όλους τους επόμενούς του με μηδενική διανομή του αγαθού. Αφού ολοκληρωθεί η φάση της διανομής, ο αλγόριθμος προσθέτει όλες τις διανεμηθέντες μάζες και έστω ότι αθροίζονται σε ποσότητα t. Αν $t < 1$, τότε μοιράζει ποσότητα $\frac{t+1}{2}$ στον πρώτο παίκτη, ενώ αν $t > 1$, τότε του μοιράζει ποσότητα $\frac{t}{2}$. Έπειτα επαναλαμβάνει την διαδικασία όπως αναφέρθηκε παραπάνω. Ο αλγόριθμος σταματά όταν $t = 1$. Η παραπάνω περιγραφή φαίνεται πιο αναλυτικά και τεχνικά στον Αλγόριθμο 5.

Algorithm 5 Μηχανισμός διανομής "Διατήρησε Διαδοχικά Σκαλάκια κατά το δυνατόν Ψηλότερα" (ΔΔΣΨ) παραμετροποιημένος από την συνάρτηση $g()$

Input: Πίνακας 1 n από n μη-αρνητικές αξίες παικτών για το αντικείμενο (χβτγ "επαναετικετοποιούμε" τις αξίες, έτσι ώστε : $v_1 \quad v_2 \quad \ldots \quad v_n$)
Output: Διανομή $x(v)$

Έστω α ένας τυχαίος πραγματικός αριθμός που επιλέγεται ομοιόμορφα από το διάστημα $[0, 1]$.
κατωΦραγμα 0;
ανωΦραγμα 1;
while (αληθές) **do**
  Θέσε $x_1 = \alpha$;
  $s \quad 2$;
  $x_s \quad \frac{x_1 \, v_1 \quad g(v_1, v_s)}{v_s}$;
  while $x_s \quad 0$ **do**
    $s{+}{+}$;
    $x_s \quad \frac{x_{s \quad 1} \, v_{s \quad 1} \quad g(v_{s \quad 1}, v_s)}{v_s}$;
  end while
  for i 2 $[s, n]$ **do**
    $x_i \quad 0$;
  end for
  if $\sum_{i=1}^{} x_i > 1$ **then**
    $a \quad \frac{\text{κατωΦραγμα} + x_1}{2}$;
    ανωΦραγμα x_1;
  else if $\sum_{i=1}^{} x_i < 1$ **then**
    $a \quad \frac{x_1 + \text{ανωΦραγμα}}{2}$;
    κατωΦραγμα x_1;
  else
    return $x(v)$
  end if
end while

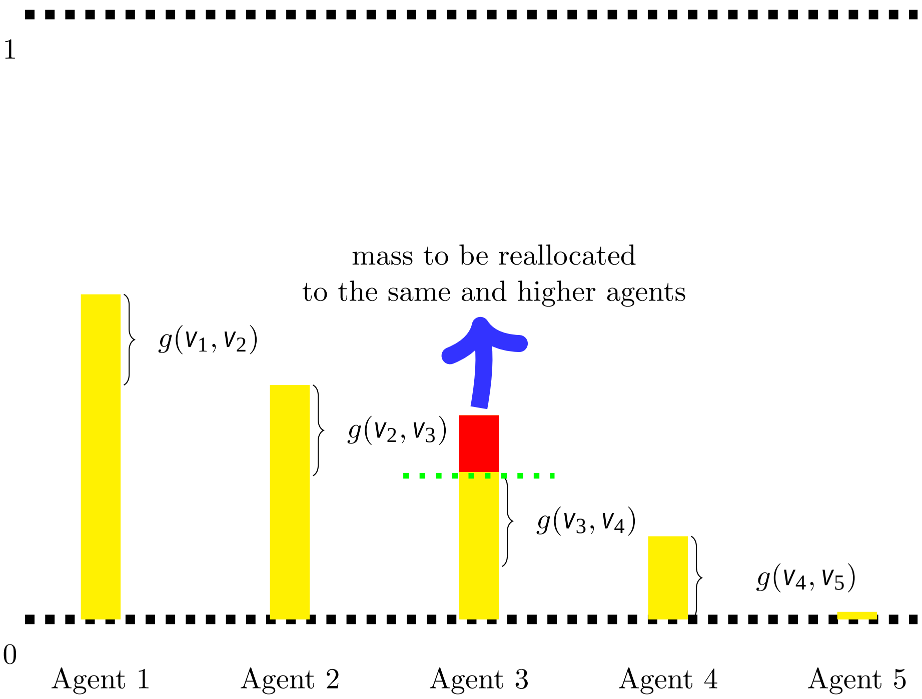
Στην συνέχεια, αποδεικνύουμε ότι ο παραπάνω μηχανισμός διανομής αγαθών λύνει βέλτιστα το γραμμικό πρόγραμμα ΓΠ1, για κάθε συνάρτηση απόστασης $g()$. Στην γλώσσα των μαθηματικών, έχουμε:

Theorem β.5.1. *Ο μηχανισμός διανομής "Διατήρησε Διαδοχικά Σκαλάκια κατά το δυνατόν Ψηλότερα" διανέμει το αντικείμενο βέλτιστα για το ΓΠ1, για κάθε συνάρτηση απόστασης* $g()$.

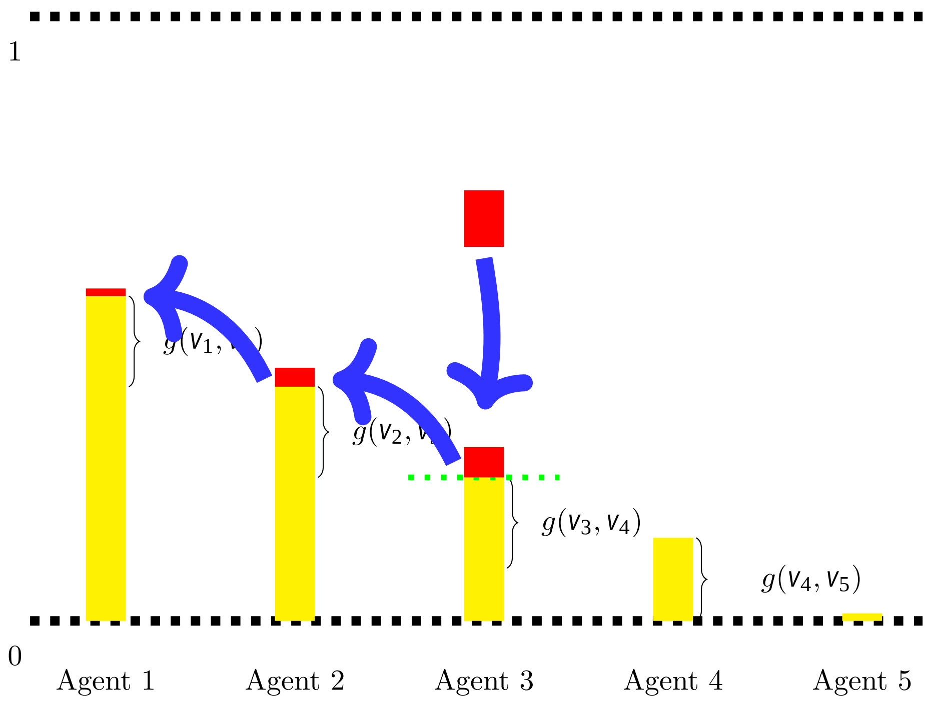
Επειδή πρόκειται για ένα από τα κυριότερα αποτελέσματα της διπλωματικής μας θα δώσουμε μια αδρομερή περιγραφή της απόδειξης.

Σκαρίφημα της Απόδειξης :

Η απόδειξη είναι εύκολη και μάλλον είναι καλύτερα να την αποδώσουμε γραφικά. Μόνο μερικά λόγια που "βοηθούν" τις εικόνες να εξηγήσουν την λογική. Ο αλγόριθμός μας είναι άπληστος και για αυτό επιχειρούμε να αποδείξουμε την βελτιστότητά του μέσα από ένα επιχείρημα ανταλλαγής. Συγκεκριμένα, υποθέτουμε ότι η λύση μας δεν είναι βέλτιστη και διαφοροποιείται από την βέλτιστη για πρώτη φορά στην αξία που παίρνει από τον μηχανισμό ο παίκτης i. Συνεπώς, βάσει της λογικής του αλγορίθμου μας, ο βέλτιστος μηχανισμός έχει δώσει παραπάνω αξία στον i στο παίκτη από την ελάχιστη που χρειάζεται ώστε να παραβιαστούν οι περιορισμοί "δικαιοσύνης" που έχουμε ορίσει. Την επιπρόσθετη μάζα που δημιούργησε αυτήν την επιπρόσθετη αξία, την αφαιρούμε από τον i στο παίκτη και την διαμοιράζουμε στους παίκτες 1 έως i με τρόπο ώστε να διατηρούνται οι περιοριμοί δικαιοσύνης με "ισότητα" (προφανώς και αυτός μεταξύ $(i \quad 1)$ στου και i στου παίκτη) μέχρις ότου να μην υπάρχει άλλη μάζα προς διανομή. Με αυτόν τον τρόπο, σταδιακά τροποποιούμε την βέλτιστη λύση με τρόπο ώστε να ταυτιστεί με την άπληστη λύση μας και μάλιστα αυξάνοντας το συνολικό κοινωνικός όφελος (αφού η μάζα αναδιανέμεται στον i στό παίκτη και σε παίκτες υψηλότερης αξίας από αυτόν). Επομένως, η βέλτιστη λύση πετυχαίνει μικρότερο κοινωνικό όφελος από την δική μας, άτοπο. Τα παραπάνω, αποτυπώνονται πιο γλαφυρά στις Εικόνες β.1 και β.2.



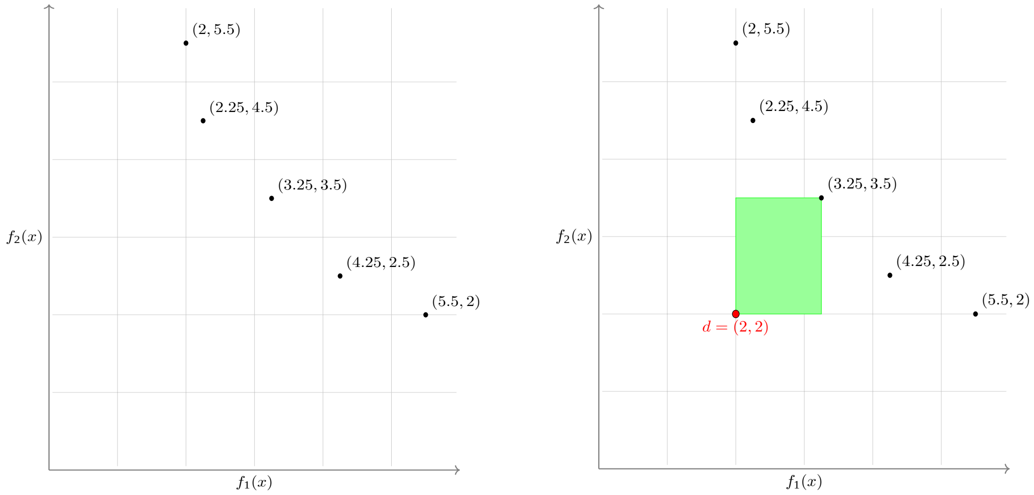
Σχήμα β.1: An illustration of the initial allocation, assuming $k-$th constraint is not tight



Σχήμα β.2: An illustration of our many-to-one greedy exchange argument

β.6 Συνάρτηση κοινωνικού οφέλους Nash: Εναλλακτικός τρόπος μηχανισμού διανομής ευθείας αναλογικότητας

Ένας εναλλακτικός τρόπος για επίτευξη αναλογικής διανομής αγαθών, και πιο συγκεκριμένα στην πλέον "πρωτόλεια" μορφή της (δηλαδή όταν ο i στος παίκτης παίρνει $\frac{v_i}{\sum_{j=1}^{n} v_j}$) προκύπτει μέσα από την Αρχή "Τροποποίησε ή αντικατέστησε την αντικειμενική συνάρτηση κοινωνικού οφέλους" και πιο συγκεκριμένα από την αντικατάσταση της συνάρτησης κοινωνικού οφέλους με τον *γεωμετρικό μέσο όρο των αξιών των παικτών για το προς διανομή αντικείμενο* (ας επικεντρωθούμε στην περίπτωση διανομής ενός αγαθού). Αυτή η συνάρτηση ονομάζεται συνάρτηση κοινωνικού οφέλους Nash - ναι και εδώ είναι χωμένος αυτός ο άνθρωπος - και για να μεγιστοποιηθεί απαιτεί κάποιου είδους "αναλογικότητα"/"δικαιοσύνη" στις διανομές (ας το αφήσουμε κάπως αφηρημένο - τουλάχιστον στα πλαίσια της περίληψης). Στο Σχήμα β.3 παρατηρούμε πως το σημείο του μετώπου Pareto που μεγιστοποιεί το εμβαδόν του ορθογωνίου παραλληλογράμμου μεταξύ αυτού και του σημείου αναφοράς, είναι ένα ενδιάμεσο σημείο (εν προκειμένω το $(3.25, 4.5)$) και όχι ένα ακρειανό.



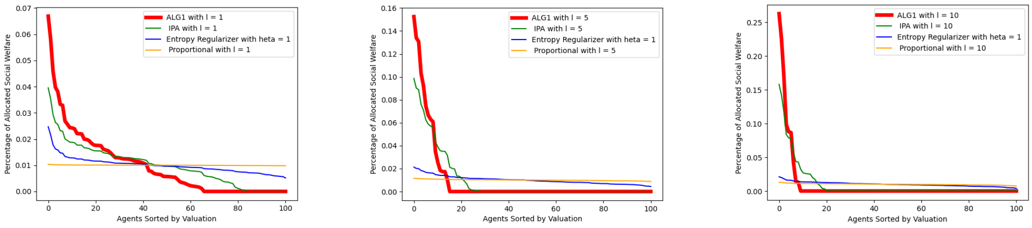
Σχήμα β.3: Nash social welfare maximizing point is the "fairest" point of the Pareto frontier

Αυτό σημαίνει ότι υπάρχει κάποια ισορροπία στο σημείο επιλογής, υπό την έννοια ότι και οι 2 συναρτήσεις λαμβάνουν μεγάλες τιμές και δεν μεγιστοποιείται μόνο μία από τις 2 με την άλλη να παίρνει μικρή τιμή (όπως συμβαίνει στα ακρειανά σημεία του μετώπου Pareto). Αυτή η ιδιότητα έχει γενικότερη ισχύ για την συγκεκριμένη συνάρτηση. Συγκεκριμένα, αποδεικνύουμε πολύ εύκολα με χρήση πολλαπλασιαστών Lagrange πως σε συνθήκη διανομής κλασματικών μερών ενός αντικειμένου στους παίκτες (ισοδύναμα διαμοιρασμού πιθανοτήτων από κατανομή πιθανότητας). η μεγιστοποίηση του κοινωνικού οφέλους Nash συμβαίνει όταν κάθε παίκτης λάβει την κανονικοποιημένη αξία του για το αντικείμενο, ή - ισοδύναμα - λάβει την αναλογική διανομή του αγαθού στην "πρωτόλεια" μορφή της. Η παραπάνω πρόταση, σε μορφή θεωρήματος μεταφράζεται ως εξής:

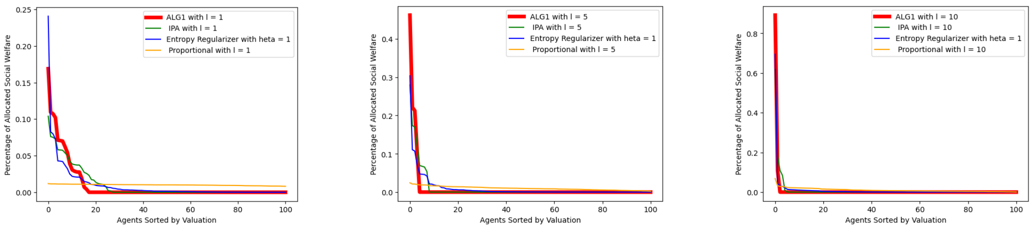
Theorem β.6.1. *Το κοίλο πρόγραμμα διαμοιρασμού πιθανοτήτων με αντικειμενική συνάρτηση το κοινωνικό όφελος Nash (με μεταβλητές απόφασης τις πιθανότητες και παραμέτρους τις αξίες των παικτών), έχει ως βέλτιστη λύση την αναλογική κατανομή.*

β.7 Πειραματική σύγκριση μεταξύ μηχανισμού διανομής ΔΔΣΨ, μηχανισμού διανομής αντίστροφης αναλογικότητας, μεγιστοποιητή κοινωνικού οφέλους με αρνητική εντροπική κανονικοποίηση και μηχανισμού διανομής ευθείας αναλογικότητας

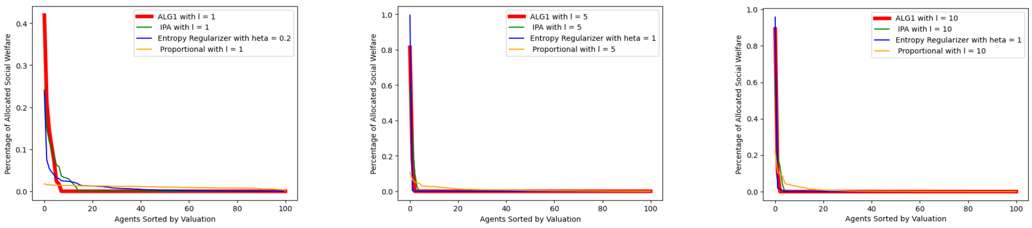
Στην παρούσα υποενότητα, παραθέτουμε τα πειραματικά αποτελέσματα από διάφορα "τρεξίματα" που επιχειρήσαμε για την σύγκριση των κατανομών που παράγουν ως προς τις διανομές οι 4 μηχανισμοί του ενδιαφέροντός μας καθώς και την επίδοσή τους (κοινωνικό όφελος) εν συγκρίσει και με την μέγιστη δυνατή, δηλαδή εκείνη που επιτυγχάνει ο "Υψηλότερου Πονταρίσματος Παίκτης Κερδίζει" μηχανισμός.



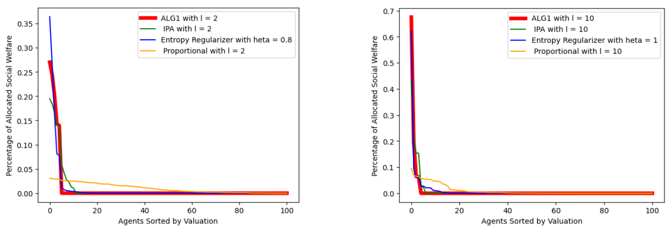
Σχήμα β.4: IPA vs KCST vs PA vs Entropy Regularizer for Single Gaussian($\mu = 30, \sigma = 0.3$)



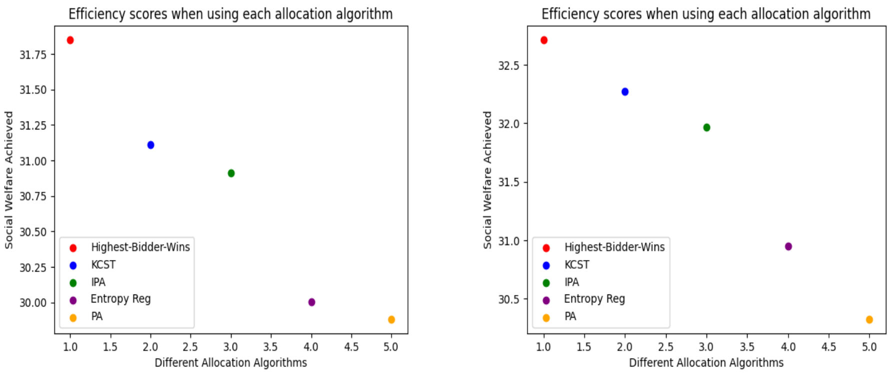
Σχήμα β.5: IPA vs KCST vs PA vs Entropy Regularizer for Single Gaussian($\mu = 30, \sigma = 3$)



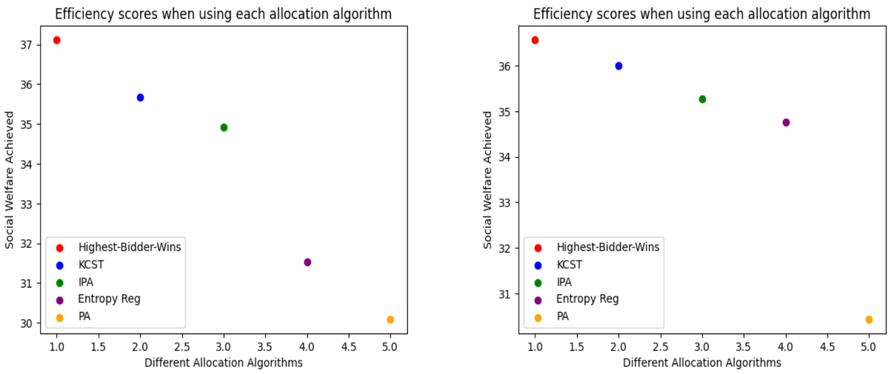
Σχήμα β.6: IPA vs KCST vs PA vs Entropy Regularizer for Single Gaussian($\mu = 30, \sigma = 9$)



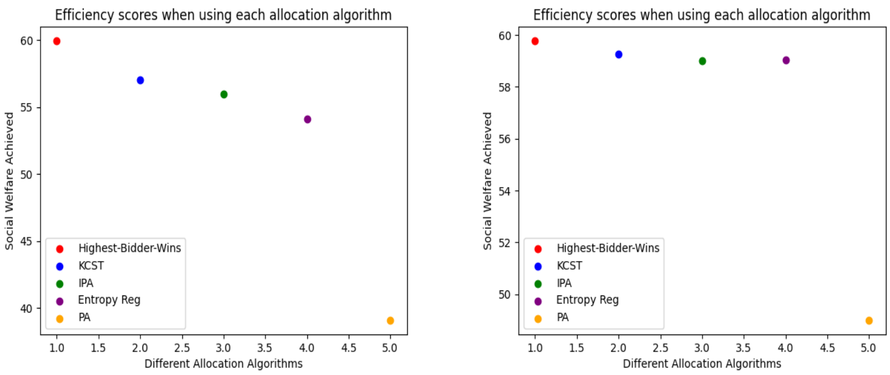
Σχήμα β.7: IPA vs KCST vs PA vs Entropy Regularizer for Uniform($\mu = 30$)



Σχήμα β.8: Efficiency of IPA vs KCST vs PA vs Entropy Regularizer vs HBW for Gaussian($\mu = 30, \sigma = 0.9$)



Σχήμα β.9: Efficiency of IPA vs KCST vs PA vs Entropy Regularizer vs HBW for Gaussian($\mu = 30, \sigma = 3$)



Σχήμα β.10: Efficiency of IPA vs KCST vs PA vs Entropy Regularizer vs HBW for Uniform($\mu = 30$)

Quicksand

"The only way to get rid of a temptation is to yield to it"
"Consistency is the last refuge of the unimaginative"

Mqa_p Ugjbc

Chapter 1

Ariadne's Thread

1.1 Introducing the Introduction

First things first, we have to define the land through which we will walk in the entire diploma thesis. This is, actually, a complex task to achieve since we stepped in different fields and adopted different principles, in order to achieve the same thing. I mean it's not so easy to explain both conceptually and technically what we are trying to do here without sticking into details, that affects negatively the intuition pump that we wish to accomplish through the introductory material. Anyway, we will give it a try!
Briefly, our goal is to achieve some sort of "individual fairness" for every single participant in a (single) item allocation process, but in a way that it scores consistently high social welfare. By the term "individual fairness", we mean something really simple that includes an interior "load" of fairness in a natural and straightforward way: Agents with similar valuations should receive similar proportions of the item. In a more technical level, the basis of our work was the work of Chawla et al.([5]), where they try to design an allocation mechanism that achieves this desideratum indirectly. More specifically, they come up with an indirect definition of individual fairness, that - if and when satisfied - implies the straightforward definition that we gave above. In a natural language, this definition can be described like this: for an arbitrary agent, say i, define a radius for his value, say λ. Keeping all the other agents' valuations fixed and agent i's valuation multiplicatively close to his original valuation with respect to his radius (i.e. $v_i' \in [\frac{1}{\lambda}, \lambda] \cdot v_i$), agent i should receive similar allocation in the 2 instances. In our line of work, in order to achieve individual fairness, we come up with a more direct definition that we try to satisfy than the one in Chawla's paper. More precisely (but not ultimately precisely!), considering that the agents are sorted based on their valuation for the -soon to be allocated-item, we enforce that agents of consecutive valuations (hence somehow similar valuations) will receive similar ***values*** (and not just allocations) from the mechanism. In order to achieve this, we have experimented - both theoretically and programmatically - with different approaches (let's call them principles to emphasize on the generalizing power and value of these though processes), that we briefly summarize below:

- **The "Proportional Allocation Rule" Principle:** Allocate proportionally to (some increasing function of) the values of agents or deduct proportionally to (some increasing function of) the inverse of values of agents.
- **The "Linear Programming" Principle:** Solve optimally a Linear Program with "fairness" constraints of our choice, where the decision variables represent the fraction of

the item to be allocated to each agent.

- **The "Slightly or completely modify the objective function" Principle:** Remove the "fairness" constraints and instead "enhance" the (social welfare) objective function with some term that enforces decentralization of the mass **or** completely change the objective function with a new one who is maximized in "fair" allocations.

Before jumping into the main analysis and technical details of each approach, we consider our duty in a scientific but also ethical way, to mention some worth mentioning works that lie in the - really vast - algorithmic fairness literature, since they inspired and educated either us or the researchers that inspired and educated us (and therefore us after 2 hops in the social graph!). The definition of individual fairness that created the basis of the works that created the basis of our work (:)) can be found in the paper of Dwork et al.([15]). In a very high-level and abstract way, some illustrious works on the algorithmic fairness from the prism of classifying - under certain criteria - the agents into several groups and trying to achieve some sort of (well-defined) balance between and/or inside the different groups can be found in the following papers: [10], [20], [22],[23], [46]. Staying - for a little bit more - in the waters of group fairness, it can be smoothly established a logical bridge and (hence) a scientific interconnection with the machine learning/statistical group fairness - meaning having people being categorized into groups, we try to minimize the error of a (say) binary decision that concers each individual that is included in one of these groups (for example people with higher market force and rich medical data history and people with lower market force and poorer medical data history, when the decision is to be hospitalized or not). Epigrammatically, some works that lie on this direction can be found here: ([25],[45],[44]).
In another research direction, that lies under the same Algorithmic Fairness umbrella, allocating indivisible items between individuals (and/or groups of individuals) was the starting point of our conceptual trip to this field. In this topic, Markakis has some very important contributions in both research and educational level, with works (and joint works) like [26], [27] and [28].
A research direction very similar (and in some cases identical) to the one followed in our thesis is allocating divisible item(s) to the agents under a proportional logic. We chose to mention the following works: [35], [3],[6] and [8].
Finally, from a clearly more technical standpoint, the main algorithmic topics that enhanced our tool-set throughout the whole diploma thesis are Randomized Algorithms ([31], [30]), Online Algorithms([16]) and Differential Privacy([37]) (and many more that we mention below and find tiring to mention here:)).

Alright, now that we have established what we are trying to do in this thesis and how (or at least I hope so!), let's jump into the details.

1.2 Motivation

Before giving a high-level explanation of our technical trigger that initiated our research attempts, maybe it is better to talk a little bit about the real-life perspective that could motivate theoretical research. Why should we care about stability in allocating things at auction settings in the first place? As we will see later on (in the Differential Privacy subsection of the Preliminaries section), privacy can be a major concern for the mechanism designer if it is a major concern for the participants (incentivizing their behavior). And when do privacy issues concern the participants? For example, let's think of a multi-stage auction. If the valuation

and, thus, the bidding strategy of an agent in the current stage is somehow correlated with his bidding behavior in previous stages, then -definitely - she should be aware of any potential information leakage.
In a more technical level, the onset of our own work was the work of Chawla and Jagadeesan ([5]) and a joint work of them with Ilvento ([21]). In fact, the first paper was the one that mostly triggered our research interest, where we first got in touch with the idea of the Inverse Proportional Allocation Mechanism and, in a second level, with the idea of "keeping" low valuation agents with zero-allocation without hurting "fairness", in order to achieve higher and robust social welfare guarantees. More concretely, our initial intuition was that this idea was generated from a principle quite different from the proportional allocation one. Hence, we developed a linear program with our own defined stability constraints and constructed an algorithm that solves it optimally. If the principle behind this Inverse Proportional Allocation rule is the optimal solution of our program, then obviously this mechanism should be equivalent with our algorithm (short and honest disclaimer, we aren't quite sure yet :)). During the development of our diploma thesis we experimented with other "auction-stabilizing" techniques too, such as Negative Entropy Regularization and changing the objective function from social welfare to Nash social welfare. For a more intuitive and mathematical explanation of all the above, we refer the reader to the corresponding sections.

1.3 Related Work

In this section, we chose to mention only the absolutely essential papers for our own line of work, along with their corresponding - twofold - contribution, both to our theoretical background and ideas, for generating our own theoretical background and ideas. As we mentioned earlier and we will mention a lot of times in our diploma thesis, our starting point was the work of Chawla and Jagadeesan ([5]). In fact, the whole "Linear Programming" and "Proportional Allocation Rule" part was generated - mostly - based on the above paper. Obviously, our first introduction to the differential privacy notion/solution concept came through the fundamental work of McSherry & Talwar ([29]). Now, for the "Entropy Regularizer"-related sections, we should mention the amazing lecture notes of Syrgkanis ([42]) as the main contributing factor of our understanding and explanations. Finally, we got a glimpse of the Nash Social Welfare objective function through the work of Charkhgard et al. ([4]) and the classic paper of Nakamura and Kaneko ([32]).

1.4 Overview of Our Contribution

In our diploma thesis, our main contribution is the introduction of a new "family" of allocation algorithms, the "Keep Consecutive Scales Tight" allocation mechanisms. Also, we gave a theoretical proof of an alternative way to achieve the "vanilla" proportional allocation for the single-item case, by simply solving optimally (with the method of Lagrange multiplier) an unconstrained concave program with the Nash Social Welfare Objective function with a probability distribution in the role of the decision variables. Besides the above, we have conducted some experimental analysis on several allocation techniques like Proportional & Inverse Proportional, maximizing Social Welfare with an Entropy Regularized term and our "Keep Consecutive Scales Tight" allocation mechanisms. Obviously, my apologies for the density of the terminology, it is a necessary evil for the provided facilitation of the "obligatory" structure of the diploma thesis. I promise that everything will make sense in a while, given one constraint: the reader's little patience!

Look in the mirror, what do you see ?

"Adversity makes men, and prosperity makes monsters"
"The greatest happiness you can have, is knowing that you do not necessarily need happiness"

Tgarmp F sem

Chapter 2

Preliminaries

2.1 Convex Analysis

Before jumping to the more algorithmic aspect of our Preliminaries section, it will be useful to introduce and formally define some very basic notions that will enhance our mathematical tool-set and accompany us until the end. Our main source is the fundamental book of Shalev-Shwartz and Ben-David [39].

2.1.1 Convex Sets

Definition 2.1.1. (Convex Set). A set S in a vector space is convex if for every pair of value vectors $\boldsymbol{x}, \boldsymbol{y} \in S$, S contains the line segment that lies between $\boldsymbol{x}$ and $\boldsymbol{y}$. In a more mathematical expression :

$$\lambda \boldsymbol{x} + (1 - \lambda)\boldsymbol{y} \in S, \text{ for all } \lambda \in [0, 1]$$

For better understanding, in the figure below we illustrate some examples of convex and non-convex sets.

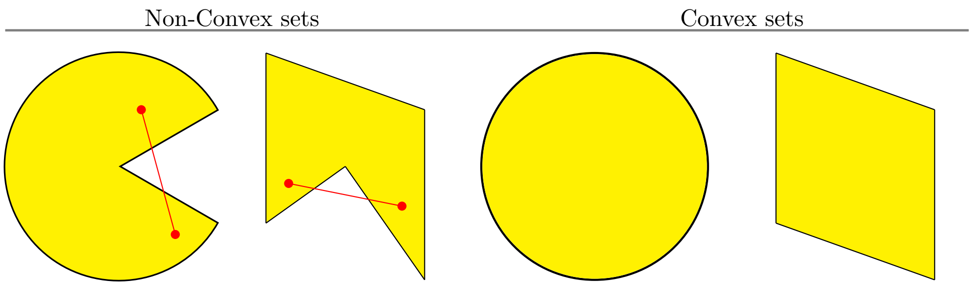


Figure 2.1: Illustration of convex and non-convex sets

2.1.2 Convex Functions

Definition 2.1.2. (Convex Function). Let S be a convex set. A function $f : S \to \mathbb{R}$ is convex if for every pair of vectors $\boldsymbol{x}, \boldsymbol{y} \in S$ and $\lambda \in [0, 1]$,

$$f(\lambda \boldsymbol{x} + (1 - \lambda)\boldsymbol{y}) \leq \lambda f(\boldsymbol{x}) + (1 - \lambda) f(\boldsymbol{y})$$

In English, f is convex if for any x, y, the graph of f between x and y lies below the line segment that joins $f(x)$ and $f(y)$. For an illustration of a convex function, you can see the figure below.

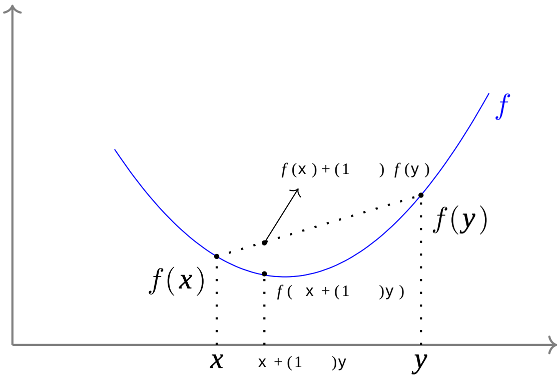


Figure 2.2: An illustration of Convexity

2.1.3 Strong Convexity

Definition 2.1.3. (Strongly Convex Function). A function f is λ-strongly convex if for all x, y and $\mu \in (0, 1)$ it holds that :

$$f(\mu x + (1 - \mu)y) \leq \mu f(x) + (1 - \mu)f(y) - \frac{\lambda}{2}\mu(1 - \mu)||x - y||^2$$

Obviously, every convex function is 0-strongly convex. Just to gain some intuition, strong convexity implies a quadratic lower bound on the growth of the function. An illustration of strong convexity can be found in the figure below.

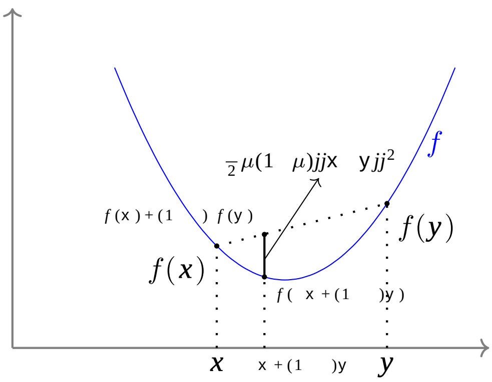


Figure 2.3: An illustration of Strong Convexity

2.1.4 Lipschitz Condition

Definition 2.1.4. (Lipschitzness with respect to the Euclidean norm). Let $S \subseteq \mathbb{R}^d$. A function $f : \mathbb{R}^d \to \mathbb{R}^k$ is ρ-Lipschitz over S if for every $x_1, x_2 \in S$ it holds that :

$$||f(x_1) - f(x_2)|| \leq \rho \; ||x_1 - x_2||$$

2.1.5 Jensen's Inequality

Definition 2.1.5. For any real concave function ϕ, numbers $x_1, x_2, \ldots, x_n$ in its domain, and positive real numbers $a_1, a_2, \ldots, a_n$:

$$\frac{\sum_{i=1}^{n} a_i \phi(x_i)}{\sum_{i=1}^{n} a_i} \leq \phi\left(\frac{\sum_{i=1}^{n} a_i x_i}{\sum_{i=1}^{n} a_i}\right)$$

Equality holds if and only if $x_1 = x_2 = \ldots = x_n$ or ϕ is linear on a domain containing $x_1, \ldots, x_n$.

2.2 Algorithmic Mechanism Design Basics

2.2.1 Let's talk about the "science of decision-making"

So, now let's move to the main "wrapper" of all the work produced in the limits of our diploma thesis. As we mentioned earlier, we are focused on "stabilizing" (let's leave this desideratum somehow abstract until we will make it more concrete on the fly) allocations rules. This means that we wish to concentrate the masses for allocation to the highest valuation players in order to maximize the social welfare (let's leave this desideratum somehow abstract until we will make it more concrete on the fly :)), but - at the same time - we have some "interior" sense of fairness and, hence, we impose a "centrifugal force" in order to decentralize the allocation to some degree. But, probably this is too much information for the start. Let's see the bigger picture.
Continuing with the same modus operandi as in the previous subsections, let's start with the name. Why "science of decision making"? High-level, mechanism design focuses on the design of systems with strategic participants. This means that the players that "constitute" the system can be considered as autonomous decision-makers whose objectives usually isn't aligned with the designer's one. In fact, the mechanism designer wants to "serve a greater good" by maximizing the total social welfare or raise significant revenue for his "employer". On the other hand, system participants are generally considered as selfish - thus caring only about maximizing their own utility. We will go back on "redefining" the notion of mechanism design in a more concrete way in a moment. But before that, there is one promise from the previous subsection that we have to fulfill. Why so much noise around truthfulness/incentive-compatibility? For these types of questions there is a very specific oracle that always returns satisfying answers. Life!
In the box below, we mention some real-life examples of agents' misreport because of bad mechanism design. These examples are taken from the tutorial on Incentive-Compatible and Incentive-Aware Learning at EC'20 from Nika Haghtalab & Chara Podimata ([19]).

Truly incentivizing theorists to be truly occupied with incentivizing people to tell the truth :

First the principle and afterwards the motivation.

Goodhart's Law

When a measure becomes a target, it ceases to be a good measure.

In other words, when we use a measure to reward performance, we provide an incentive to manipulate the measure in order to receive the reward. This can sometimes result in actions that actually reduce the effectiveness of the measured system while paradoxically improving the measurement of system performance.

So, let's get a taste of why and how things can go bad when paying for a good mechanism designer isn't high on our agenda, from real-life examples :

1. **Zara Restocking Process :** Back in 2010, Zara decided to reengineer its global distribution process, with the help of Operations Research. In order to do so, it needed some data, so it asked the managers of each store to report the expected amount of items that they will sell form different articles. From the managers' point of view, they were strongly incentivized to misreport, overestimating the expected amount of the top-selling items and underestimating the rest, since acquiring more top sellers could lead to increased profit for their local store and bigger bonuses for themselves. And - obviously - they did so!
2. **School Admissions :** A very classical example of data manipulation. In high school, a lot of students have good motivation to misreport their school preferences, to move to a lower-ranking school in order to achieve higher class ranking for themselves. Also, before college, candidates may take SAT multiple times or pay extra to take SAT preparation classes, thus achieving higher scores by simply "overfitting" to the examined material, without actually acquiring a better understanding.

2.2.2 Basic Definitions

In this subsection, we chose to deal with some (semi-)"formalism". More precisely, the goal is to concretely define some very basic notions of mechanism design that will follow us from the beginning to the end, mostly addressed to the algorithmically educated but without "mechanism designing theoretical background" reader.

First and foremost, let's define the objective function that will lie on the center of our attention through the entire document (along with some "tweaked" versions) : the Social Welfare Function.

Definition 2.2.1. (Social Welfare Function from the Mechanism Designer's point of view). In mechanism design, social welfare function is a function that ranks items' allocations as less desirable, desirable, or indifferent for every possible pair of allocations.

Alright. So, speaking about mechanism design, everyone understands the term design, but what about the mechanism part? From the standpoint that we examine here, a mechanism is a set of rules that orchestrates how agents with selfish motives should behave in a strategic interaction to achieve a desirable (for the designer and/or for the society) collective outcome. More concretely, this set of rules consists of 2 parts : an allocation rule (how recourses should be distributed among the agents) - which is our major concern in this thesis - and a payment rule (how much each agent should pay or receives a compensation given the result of the allocation rule). Both the challenge and the goal is to create rules that align individual incentives with the overall goal, ensuring that the outcome is not only efficient but satisfies some notion of fairness. Ok, so now, let's try to define formally our fundamental notion.

Definition 2.2.2. (Mechanism). A mechanism M can be represented as a tuple $M = (G, A, T, O, R, P)$, where:

- G is the set of agents.
- $A = \prod_{i \in G} A_i$ is the set of possible joint action profiles, where A_i is the set of possible actions for agent i.
- $T = \prod_{i \in G} T_i$ is the set of possible types for each agent, where T_i is the set of possible types for agent i.
- O is the set of possible outcomes.
- $R : A \times T \to O$ is the allocation rule that maps joint actions and types to outcomes.
- $P : A \times T \to \mathbb{R}$ is the payment rule that maps joint actions and types to allocation profiles.

For reasons of brevity, we will mostly refer to a mechanism M as $M : T^n \to O$. Next, we will define the notion of prior-free mechanism, which will be useful in the "Proportional Allocation Rule" principle section, since the designed allocation mechanisms there display that property.

Definition 2.2.3. (Prior-free Mechanism). Let a mechanism $M : T^n \to O$, where agents' types T are drawn from some probability distribution. The mechanism M is prior-free if it does not require any prior knowledge or assumptions about the distribution that generated the agents' types.

Finally, let's define formally 2 notions (in fact there are the 2 sides of the same coin) of (approximate) truthfulness, that we will need in just 2 subsections below (in the Differential Privacy part) :

Definition 2.2.4. ($\epsilon-$ approximate Dominant Strategy Truthfulness for Deterministic Mechanisms). A deterministic mechanism $M : T^n \to O$ is $\epsilon-$ approximately strategy truthful if for all $t \in T^n$ and for all i and $t_i' \in T$:

$$u_i(M(t)) \geq u_i(M(t_{-i}, t_i')) - \epsilon$$

Intuitively, this means that no agent can gain more than ϵ utility by misreporting their type.

Definition 2.2.5. ($\epsilon-$ approximate Dominant Strategy Truthfulness for Randomized Mechanisms). Assuming risk neutral players, a randomized mechanism $M : T^n \to O$ is $\epsilon-$ approximately strategy truthful if for all $t \in T^n$ and for all i and $t_i' \in T$:

$$E_{o \sim M(t)}[u_i(M(t))] \geq E_{o \sim M(t)}[u_i(M(t_{-i}, t_i'))] - \epsilon$$

The interpretation is pretty much the same with the one in the deterministic version.

2.3 Solution Concepts of Stability with motivation in Auction Settings

2.3.1 Self-Referential Global Stability a.k.a. Differential Privacy for Mechanism Designers

2.3.1.1 Basic Definitions

In this section, we will be concerned with a very important privacy/stability solution concept, Differential Privacy, introduced by McSherry & Talwar in [29]. Since we are interested in auction settings, we will examine this kind of privacy as a gametheoretic desideratum. But, what does that even mean?
Before establishing the logical bridge between Differential Privacy and Mechanism Design, we have to define the notion of Differential Privacy in the first place. The setting where this concept lives in is the so called "trusted curator model".

Trusted Curator Model :

- n individuals, who each have their own datapoint
- each individual trusts the curator with their datapoint in raw form, but no one else
- with the individuals' raw datapoints in hand, the curator runs an algorithm A and outputs the result of this computation

Differential Privacy is the property of the algorithm A that there is no individual whose datapoint has a large impact on the algorithm's output. In order to jump into our main area of concern, meaning a mechanism that ensures this property in a prevalent auction setting, some necessary evil first : formalism!
Already from the definition of Differential Privacy, it will become obvious that there exists a strong interconnection between this idea and randomness. So, let's define what a Randomized Algorithm is and afterwards proceed to our main definition.

Definition 2.3.1. (Randomized Algorithm). A randomized algorithm A creates a mapping $A : D \to \Delta(R)$. When receiving an input $d \in D$, A outputs with probability $(A(d))_r$ the following : $A(d) = r$, for each $r \in R$. Also, the probability space is defined over the coin flips of A.

Definition 2.3.2. (Differential Privacy). A randomized algorithm $A : X^n \to O$ is ϵ differentially private if for all i, for all $x \in X^n$, for all $x_i' \in X$, and for all outcome events $E \subseteq O$:

$$\Pr[A(x) \in E] \leq e^{\epsilon} \Pr[A(x_{-i}, x_i') \in E]$$

Now that we have defined Differential Privacy, it is easy to see that this notion is quite similar (at least from the mathematical standpoint) with the truthfulness of a mechanism, since they both require "similarity guarantees" when unilateral deviations take place. Actually, (ϵ-) differential privacy implies (ϵ-approximate) truthfulness. More formally, the following theorem holds :

Theorem 2.3.1. *If M is ϵ-diferentially private for $\epsilon \leq 1$, then it is also ϵ-approximately dominant strategy truthful.*

Ok, so we have established a mathematical interconnection between mechanism design and differential privacy. But is there any real-life motivation behind this interconnection? Why should a mechanism designer should be concerned with privacy issues, anyway? I mean, obviously, a mechanism designer cares only about the incentives and the motivation of the agents to participate in the auction. So, if serious "privacy-preserving issues" are a disincentive for an agent's participation, then yes, the auction designer should be concerned a lot. But is this the case here ?
In fact, it could be. Perhaps an agent might unbeknownst to the seller be participating in another auction tomorrow. If the agent's value for the item he might be purchasing tomorrow is correlated with her value for the current item, then he should care a lot about any possible information leakage.

2.3.1.2 Motivating the Transition to the Exponential Mechanism

After getting a glimpse of the Differential Privacy notion and establishing a (high-level) framework - consisting mostly of definitions -, now it's time to introduce one of the most fundamental, yet powerful, mechanisms in the corresponding literature, the exponential mechanism. To illustrate the type of problems this mechanism tries to address, let's take a look at a very simple auction setting, the ***digital goods auction (DGA)***. So, the scenario is the following :

Digital Goods Auction Scenario :
• n unit-demand agents
• infinite supply of identical items

The name derives from the fact that the motivation of the above auction format comes from selling items with zero marginal cost of production, meaning digital items like software. Obviously, in this infinite supply setting it makes no sense to seek for welfare maximization, since we can easily achieve $\sum_{i=1} v_i$ always (just give each agent 1 copy of the item for free). Hence, our main focus here is revenue maximization. The logical bridge with the stability framework that we are interested in this section is created easily, since we wish to "run" this auction in a differentially private way. The revenue, obviously, has to depend from the chosen price, and more specifically it holds that : Revenue = Price $|\{i : \text{Price} \leq v_i\}|$. Before diving into the main idea, let's take a look at some other - more naive - approaches first :
First Idea : As always, the simplest idea is to "brute force". That means, simply to set repeatedly the threshold/price equal to every agent's valuation and keep the revenue maximizing one. This idea, obviously, leads to efficiency, since we will always achieve maximum revenue, but completely violates the differential privacy restriction, in the "plausible deniability property" sense (just think of the case where some agent has very high violation for the digital item compared to the others and raises the price super high).
Second Idea : Ok, the first approach was indeed a bit "vanilla". But what about adding some noise to whatever price we end up with? Besides, a lot of brain power and work have been dedicated to that direction - always inside the privacy context -, like Laplacian ([9]) and Gaussian([11],[12],[13],[14]) mechanisms. It turns out that this idea will also fail because of revenue's sensitivity to price changes, meaning a small "perturbation" in the chosen price may cause huge revenue loss. To see this, let's go through an example. Suppose there are 3 unit-demand agents, with valuations $\{\alpha, \alpha, 3\ \alpha + \epsilon\}$. As we can see graphically in Figure 2.4, there are 2 prices (α and $\alpha + \epsilon$) where small price increases lead to big revenue decreases.

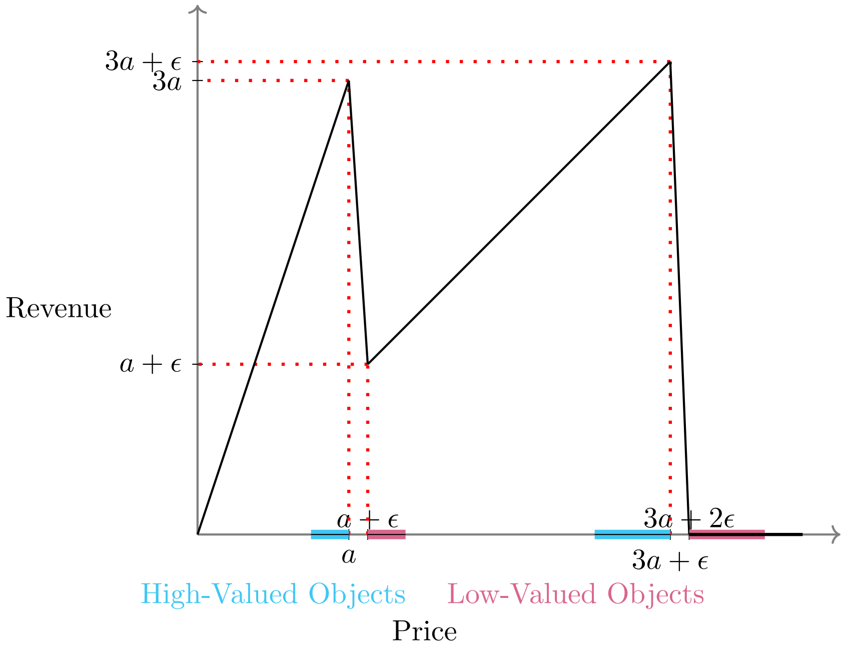


Figure 2.4: Revenue sensitivity to small price changes in Digital Goods Auction without randomization

Ok, so can we do any better? Yes, randomize!
The Key Idea : Use randomization to select prices, giving a significant advantage to those prices who end up in high revenues. More technically, instantiate a mechanism equipped with a quality score function $s : X^n \times O \to \mathbb{R}$, which is an indicator of how good an output (let's call it object) o is, for input x. After the quality score "assignment", the mechanism outputs a random object within the range O, giving exponentially weighted preference to objects with high quality score.

2.3.1.3 The Exponential Mechanism

Besides the previous differential privacy-related works that we mentioned in the previous subsections, we should also mention the work of Balcan et.al ([2]) that - alongside with the others - helped us create a solid theoretical background on the Exponential Mechanism. Before presenting the (super simple) Algorithm of how the Exponential Mechanism produces allocations, it is important to formally define the most important and "stability related" aspect of the mechanism, meaning the sensitivity of the quality score to unilateral deviations. More precisely :

Definition 2.3.3. (Sensitivity of the Quality Score of the Exponential Mechanism). The sensitivity of the quality score $s : X^n \times O \to \mathbb{R}$ is the maximum absolute value change a unilateral deviation of an agent can affect on the quality score. More formally :

$$\Delta s = \max_{i \in [n]; x \in X^n; x_i' \in X; o \in O} |s(x,o) - s((x_{-i}, x_i'), o)|$$

So, now, after having established the basic Differential Privacy framework, it is time to present the algorithm that realizes the Key Idea presented above, the Exponential Mechanism.

Algorithm 6 The Exponential Mechanism

Input: Range of objects O, quality score function $s : X^n \times O \to \mathbb{R}$, value ϵ.

1: **return** $o \in O$ with probability : $\frac{\exp(\frac{\epsilon s(x;o)}{2\Delta(s)})}{\sum_{o\in O}\exp(\frac{\epsilon s(x;o)}{2\Delta(s)})}$

The importance of the Exponential Mechanism can be confirmed by two fundamental theorems.

The first theorem states that the mechanism produced by the above algorithm is ϵ-differentially private. Aside from simply stating the theorem, we also proceed to prove it. There is a twofold reason for this. Besides the importance of this result, all the "different versions" of this proof that we are aware of, skip some non-trivial parts, rendering her obscure in a certain degree. Let's become more formal :

Theorem 2.3.2. *The exponential mechanism $M_E(x; s, O, \epsilon)$ is ϵ-diferentially private.*

Proof. From the definition of differential privacy, it is enough to show that :

$$\frac{\Pr[M_E(x; s, O, \epsilon) = o]}{\Pr[M_E((x_{-i}, x'_i); s, O, \epsilon]} \leq e^{\epsilon}$$

$$\begin{aligned}
\frac{\Pr[M_E(x; s, O, \epsilon) = o]}{\Pr[M_E((x_{-i}, x'_i); s, O, \epsilon]} &= \frac{\dfrac{e^{(\frac{\epsilon s(x;o)}{2})}}{\sum_{o\in O} e^{(\frac{\epsilon s(x;o)}{2})}}}{\dfrac{e^{(\frac{\epsilon s((x_{-i};x'_i);o)}{2})}}{\sum_{o\in O} e^{(\frac{\epsilon s((x_{-i};x'_i);o)}{2})}}} = \\
&= \frac{\sum_{o\in O} e^{(\frac{\epsilon s((x_{-i};x'_i);o)}{2})}}{\sum_{o\in O} e^{(\frac{\epsilon s(x;o)}{2})}} \cdot e^{\frac{\epsilon(s(x;o)-s((x_{-i};x'_i);o))}{2}} \leq \\
&\leq \frac{\sum_{o\in O} e^{(\frac{\epsilon s((x_{-i};x'_i);o)}{2})}}{\sum_{o\in O} e^{(\frac{\epsilon s(x;o)}{2})}} \cdot e^{\frac{\epsilon \overbrace{|s(x,o) - s((x_{-i},x'_i),o)|}^{\leq \Delta}}{2}} = \\
&= \frac{\sum_{o\in O} e^{(\frac{\epsilon s((x_{-i};x'_i);o)}{2})}}{\sum_{o\in O} e^{(\frac{\epsilon s(x;o)}{2})}} \cdot e^{\frac{\epsilon}{2}}
\end{aligned} \tag{2.1}$$

Now, let's fix an arbitrary object, say $o \in O$. It holds that :

$$\begin{aligned}
&e^{\epsilon}\uparrow \; e^{s((x'_i;x_{-i});o) - s(x;o) - \max_{i\in[n];x\in X^n;x'_i\in X;o\in O}|s((x'_i;x_{-i});o)-s(x;o)|} \leq e^0 \Rightarrow \\
&\Rightarrow e^{s((x'_i;x_{-i});o) - s(x;o)} \leq e^{\max_{i\in[n];x\in X^n;x'_i\in X;o\in O}|s((x'_i;x_{-i});o)-s(x;o)|} \Rightarrow \\
&\Rightarrow e^{\frac{\epsilon}{2\Delta}(s((x'_i;x_{-i});o) - s(x;o))} \leq e^{\frac{\epsilon}{2\Delta}\max_{i\in[n];x\in X^n;x'_i\in X;o\in O}|s((x'_i;x_{-i});o)-s(x;o)|} \Rightarrow \\
&\overset{def}{\to} e^{\frac{\epsilon}{2\Delta}(s((x'_i;x_{-i});o) - s(x;o))} \leq e^{\frac{\epsilon\Delta}{2\Delta}} \Rightarrow \\
&\Rightarrow e^{\frac{\epsilon s((x'_i;x_{-i});o)}{2}} \leq e^{\frac{\epsilon}{2}} \cdot e^{\frac{\epsilon s(x;o)}{2}}
\end{aligned} \tag{2.2}$$

By applying 2.2 for every object within the range O and summing, we end up with the following inequality :

$$\frac{\sum_{o \in O} e^{\left(\frac{\epsilon s((x_{-i};x_i');o)}{2\Delta}\right)}}{\sum_{o \in O} e^{\left(\frac{\epsilon s(x;o)}{2\Delta}\right)}} \leq e^{\frac{\epsilon}{2}} \tag{2.3}$$

By combining 2.2,2.3 we finally prove our desired inequality :

$$\frac{\Pr[M_E(x;s,O,\epsilon)=o]}{\Pr[M_E(x_{-i},x_i');s,O,\epsilon]} \leq e^{\frac{\epsilon}{2}} e^{\frac{\epsilon}{2}} \implies \Pr[M_E(x;s,O,\epsilon)=o] \leq e^{\epsilon} \Pr[M_E(x_{-i},x_i');s,O,\epsilon]$$

The second theorem emphasizes on the fact that the Exponential Mechanism is efficient, meaning that with high probability it outputs a high quality (score) object.

Theorem 2.3.3. *Fix a range of objects O. Let $OPT_{s,O} = \max_{o \in O} s(o,O)$ be the highest score obtained by any object within the range O. Also, let $O^* = \{o' \in O : s(o',O) = OPT_{s,O}\}$ be the set of objects within the range O that achieve this score. Then, it holds that :*

$$Pr[s(M_E(O)) \leq OPT_{s,O} - \frac{2\Delta}{\epsilon}(\ln(\frac{|O|}{|O^*|}) + t)] \leq e^{-t}$$

Alright, so at the moment we have a good algorithm in our hands (in the stability and efficiency sense). But, the original problem still remains. How do we pick a price in a way that leads to a differentially private and at the same time high-revenue mechanism for the digital goods auction? The following algorithm answers that question.

Algorithm 7 Digital Goods Auction with the Exponential Mechanism as Price selector

Input: Discretization parameter $\alpha \in [0,1]$, values $v_1,\ldots,v_n$, value ϵ.

```
Set range of objects O = {α, 2α, ..., 1}
Set score function s(v, p) = p · |{i : v_i ≥ p}|
Pick price p̂ ∈ O with probability : exp(ε s(v, p̂) / 2Δ(s)) / Σ_{p̂ ∈ O} exp(ε s(v, p̂) / 2Δ(s))
for i ∈ [n] do
    if v_i ≥ p̂ then
        Sell item to agent i for a price of p̂
    end if
end for
```

Just a bunch of things to work out a bit. To begin with, why did we chose to set as score function the function $p \cdot |\{i : v_i \geq p\}|$? Actually, it is a quite natural choice, since picking a single price for the digital item is a very simple and "fair" policy (price discrimination not only complexifies the auction format, but a lot of times it may be a forbidden practice). So, by picking a "horizontal" price, the total revenue equals to the price of the item times the number of bidders willing to pay at that price. Hence : revenue $= p \cdot |\{i : v_i \geq p\}|$. Also, by

choosing the fixed price policy, a natural benchmark arises for measuring the performance of our algorithms : the *best* fixed price (assuming normalized valuation vector, meaning $v \in [0,1]^n$) aka $\mathrm{OPT}(v) = \max_{p \in [0,1]} p \cdot |\{i : v_i \geq p\}|$. A worth-mentioning feature of our score function is the fact that its sensitivity equals to 1 (since changing a single agent's valuation can only change the quantity $|\{i : v_i \geq p\}|$ by 1 and also $p \leq 1$).
One more observation (and one more theorem). About the discretization parameter a. In fact, it expresses/generates tradeoffs between "accuracy" and "uncertainty". Wait, a lot of quotation marks. What does that even mean? It means that by assigning to a lower values, that implies a greater value for $\frac{1}{a}$ (i.e. number of objects inside the range $O = \{a, 2a, \ldots, 1\}$), so a finer discretization, meaning a closer approximation to the optimal price/object $\mathrm{OPT}(v)$. On the other hand, this increase in the number of objects decreases the chance for the exponential mechanism to actually pick the price that is closer to the optimal.
Finally, let's close the differential privacy part of this section with some (very) good news. It can be easily proven that with the use of the exponential mechanism price selector, there is a (approximately) truthful mechanism that, with high probability, ends up with high revenue. More concretely:

Theorem 2.3.4. *For the digital goods auction setting, there is an ϵ-approximately dominant strategy truthful auction, that in the worst case value vector v, gives the following guarantee for the revenue:*

$$Pr[Revenue \geq OPT(v) - O(\frac{\log n}{\epsilon})] \geq 0.99$$

2.3.2 Epimythium

What's the Epimythium?
Differential Privacy Version

- Differential Privacy requires "plausible deniability", that is any agent participating in the auction should be able to deny his participation (aka no big changes to the outcome with or without him).
- Strict determinism hits the wall of plausible deniability (just think the case of an agent with huge bid with respect to all the others) $\Longrightarrow$ "Randomization for Robustness!"
- Adding some noise could reassure differential privacy, but for unstable revenue functions, doesn't fix the stability issue. Hence, it doesn't provide good efficiency guarantees (aka small changes could lead to huge revenue loss).
- Core idea of the Exponential Mechanism : Choose an object at random, but giving "exponential preference" to objects with high quality score.

2.4 Some Online Learning Paraphernalia

2.4.1 "Smooth" Introduction to Online Learning & Basic Concepts through the classical Ski Rental example

Online Learning is a sub-field of Algorithms and Algorithmic Design occupied with situations of irrevocable decision-making under partially revealed input. More precisely, it deals with computational problems where : the input arrives piece-by-piece and our algorithm needs to

make an irrevocable decision each time it receives a new piece of the input. Obviously, the term online seems a bit misleading, since it has nothing to do with the Internet. Probably, the main reason of this "bad"/outmoded naming is that it was developed about 40 years ago.
A classic real-life problem with online nature is the Paging Problem. We have a 2-level computer memory (fast memory with capacity K items and slow memory with capacity N items). We need to answer a request sequence $\sigma = \sigma_1\sigma_2 \ldots \sigma_m$. Each time we access page request σ_i, we put it in fast memory with cost 1 if it's not in fast memory, otherwise we do nothing and the cost is 0. Of course, the goal is to minimize the total cost and our Paging Algorithm when receiving a request that is not in fast memory, needs to decide which item to evict and make room for the new item.
Alright, enough with the Introduction to the Introduction. Now, let's jump to the Introduction! As promised, we will try to get a taste of Online Algorithms as well as the importance of randomization in this setting through a very classical online problem, Ski Rental. This warm-up problem will help us to demonstrate basic techniques and to transition into more sophisticated ones in a smooth way. After finishing with the example, we will define all the above (and more) in a more formal way, while still emphasising in giving the intuition behind the maths.
Quick but important disclaimer before directing our attention to the example. For the Ski Rental example, our main source is the classic book of Fiat and Woeginger ([17]). The core of the core of this online learning paraphernalia subsection (aka the introduction to the online-learning setting part) consists of the work of Shalev([38]) and 2 joint works with Singer ([40],[41]). For the brief part of online convex optimization at the end, our basis is the work of Freund and Schapire ([18]).

Ski Rental

Suppose you stay in a ski resort for T days. You can either rent skis (cost 1 per day) or buy them (cost B once). You do not know the number of days you will go skiing in advance. Every morning, you get to know if you will go skiing that day. Then, you can choose if you rent skis for that day or if you buy them, in which case you can use them for the rest of the stay and do not have to rent or buy them again.

First Idea : Ok, here is an idea. Let's try to exhaust the possibilities to avoid the "big purchase", but if the situation starts getting too expensive, just buy to avoid any big damage. In a more algorithmic perspective, rent skis up to B 1 times and in the Bth skiing day, buy the skis. The above algorithm is presented below in pseudocode.

Algorithm 8 Online Deterministic Algorithm for Ski Rental

Initial Input: Cost of buying the skis B, time horizon T.

Initialize skiingDays $= 0$
Initialize cost $= 0$
while skiingDays $\quad B \quad 1$ **do**
    Read skiingDay
    if skiingDay $==$ **True then**
        skiingDays $++$
        cost $++$
    else if skiingDay $==$ **Null then**
        return cost.
    else continue
    end if
end while
cost $= B$
return cost.

We can easily prove the following result :

Theorem 2.4.1. *For any sequence of "skiing"/"not skiing" days, Online Deterministic Algorithm's cost for Ski Rental exceeds the cost of the optimal solution by a factor of at most* $2 \quad \frac{1}{B}$.

Moreover, it can be proved that the above algorithm is optimal (in terms of competitive ratio) among every deterministic online algorithm.

Theorem 2.4.2. *There is no deterministic algorithm for Ski Rental that achieves a competitive ratio strictly less than* $2 \quad \frac{1}{B}$.

To see why we hit the wall of $2 \quad \frac{1}{B}$, let's think about 2 key factors. First, every deterministic algorithm displays the same behavior no matter what the input sequence is (say that it decides to rent skis for ℓ days - $0 \quad \ell \quad T$). Second, let's think about an adversarial sequence. Obviously, this sequence consists of $\ell + 1$ skiing days, since she will "force" the deterministic algorithm to "waste" an extra of $B \quad 1$. Hence, we see that the deterministic approach is limited to competitive ratios no less than $2 \quad \frac{1}{B}$. If we want to do something better - or at least exhaust the possibilities to do something better - we have to shift gears.

Key Idea : As we saw earlier in Differential Privacy, there is an Invariant in terms of principle : "Randomization for Robustness!". So, let's take a look at the following approach :

Algorithm 9 Online Randomized Algorithm for Ski Rental

Initial Input: Cost of buying the skis B, time horizon T.

1: Flip a coin.
2: **if** Heads **then**
3: Run Online Deterministic Algorithm for Ski Rental (B,T).
4: **else**
5: Set $B^{\theta} = \frac{3}{4}\ B$
6: Run Online Deterministic Algorithm for Ski Rental (B^{θ},T).
7: **end if**

So, did we manage anything? In fact yes, by randomizing we managed to free ourselves from the B dependence. More specifically, it holds that :

Theorem 2.4.3. *The Online Randomized Algorithm for Ski Rental is strictly* $\frac{15}{8}$ *competitive.*

Just a brief intuitive explanation about this improvement, since we will rejoin this idea in the Online Leaning setting that comes next. By inserting randomness, we gave ourselves a good chance (i.e. $\frac{1}{2}$) to "escape" from adversarial sequences, which has a direct impact on the algorithm's worst-case performance.

2.4.2 Super High-Level Online Learning Framework

After our first meeting with the Online Learning setting through the Ski Rental Problem, let's direct our attention to the bigger picture. What is an Online Problem? What is an Online Algorithm? Is there any meaningful "metric" in order to measure the efficiency of our algorithms? In this section, we will try to answer briefly to the above questions.

2.4.2.1 Online Problem

Ok, so let's start with the very basics. When we reason about an online problem (at least in the context of algorithms), we refer to problems of the following nature :

Online Problem

- Our input arrives "one piece at a time".
- A decision-maker has to make an irrevocable decision every time she receives a new piece of the input, having only the knowledge of the past and the present (she is oblivious about future pieces of the input).

2.4.2.2 Online Algorithm

Having defined the type of problems that we are interested in in this section, we need to define, also, the behavior of the "guy" who tries to solve them. Maybe I should rephrase a little bit. We should define, the behavior of the "guy" who tries to solve them as well as the context where he lives in.

> **Online Algorithm**
>
> Suppose there is a set A of actions, where $|A| \geq 2$, and a time horizon $T \geq 1$.
> Every Online Algorithm works as follows :
> For each time step $t = 1, \ldots, T$:
>
> - The Algorithm commits to a probability vector over his actions, say $\boldsymbol{p}^t$, based on what it has observed so far.
> - The Adversary picks a loss vector $\ell^t : A \to [0, 1]$ (**<u>after</u>** the Algorithm's commitment)
> - An action $a^t \in A$ is chosen according to the probability vector $\boldsymbol{p}^t$, and the Algorithm incurs a loss $\ell^t(a^t)$
> - The Algorithm learns ℓ^t (meaning the entire loss vector)

2.4.2.3 Regret

Ok ok. The setup seems a little bit puzzling, but after some thought maybe it starts to make sense. Obviously we are not just seeking for algorithms, but for good ones. But what does that mean? We are looking for some sort of metric, that quantifies the distance between the online decision-maker's performance and a well-defined benchmark.
The whole point of this subsection is that we need to lower our expectations it terms of the choice of our benchmark. Why is that? Because the setting is unfair! In fact, the fact that the Adversary gets to choose the loss vector ***after*** the commitment of the Algorithm in each time step, gives him a significant advantage. Let's take a closer look at this "choice of benchmark" necessity:
First Idea : Man vs Prophet (aka compare with the best action sequence in hindsight). Let's think of the following situation: Suppose that $A = \{A_1, A_2\}$ and the Algorithm commits to the probability vector $\boldsymbol{p}^t$, at time step t. Also, suppose that the Adversary is smart enough to do the following : say A_i is the action that the Algorithm chooses to play with probability greater or equal than $\frac{1}{2}$ at time step t (obviously $i \in \{0, 1\}$). The Adversary chooses to "place" a loss of 1 at the action A_i and 0 at the other action (say i'). Hence, the best action sequence in hindsight achieves reward : $T - \sum_{t=1} \ell_t(A_{i'}) = T - T \cdot 0 = T$. The online decision-maker gets : $T - \sum_{t=1} [p_t^{A_i} \cdot \ell_t(A_i) + (1 - p_t^{A_i}) \cdot \ell_t(A_{i'})] \leq T - T \cdot \frac{1}{2} = \frac{T}{2}$. But, the thing is that we wish to design algorithms with performances that grow sub-linearly with T and that's just impossible.
So, what's next? A very reasonable trade-off is to shift benchmarks from the best action sequence in hindsight to the best ***fixed*** action sequence in hindsight. Lucky for us, this idea actually works. Hence :
Key Idea : Man vs Lucky Man (aka compare with the best <u>***fixed***</u> action sequence in hindsight). This approach leads to the main definition of this subsection.

Definition 2.4.1. (Regret). The regret of an online learning algorithm is the difference of its cumulative loss and the cumulative loss of the best fixed action in hindsight, in the worst case over loss sequences. Therefore, the following definition holds :

$$\text{Regret}(T) = \sup_{\ell_1, \ldots, \ell_T} \left(\sum_{t=1} \ell_t(a_t) - \min_{a \in A} \sum_{t=1} \ell_t(a) \right)$$

Just to avoid any possible misunderstandings, let's mention one more highly correlated definition.

Definition 2.4.2. (No-regret algorithm). An algorithm is called no-regret if it's regret grows sub-linearly with T, i.e. if $\text{Regret}(T) = o(T)$

2.4.3 Is Randomization that important? Absolutely Yes!

Here, we will define the most fundamental algorithm for our analysis up until the Entropy Regularization technique in the Online Convex Optimization setting, the Follow-the-Leader(FTL) algorithm. This algorithm will form the baseline for the whole entropy-related "conceptual building" that will be created later on. Besides introducing this algorithm, by emphasizing in one of its caveats we will highlight one more time the - probably - second most important principle principle presented in this diploma thesis : the "Randomization for Robustness" principle. Let's jump into the main analysis.

The setting is super simple : An action set $A = \{A_1, A_2\}$ and a time horizon T. At each time step t ($1 \leq t \leq T$) a decision-maker has to choose between his 2 options (A_1 or A_2) and afterwards, he incurs a loss $\ell_t(a_t) \in [0, 1]$, depending on his choice a_t. Also, he observes the loss that both actions would have incurred had he chosen them. The decision-maker's goal is to achieve sub-linear regret.

Enough with the definitions. Let's try to solve it!
First Idea : Since the decision-maker knows only (but at least completely) the past, just play the "historically best" action. This idea -in a more formal way- is given in pseudocode below :

Algorithm 10 Follow-the-Leader (Deterministic Version)

Input: Action set $A = \{A_1, A_2\}$, loss vector ℓ with all the values/losses up to step $t - 1$ (considering t as the current decision-making time step).

return $a_t = \arg\min_{a \in \{A_1, A_2\}} \sum_{t'=1}^{t-1} \ell_{t'}^a$

The thing is that - as we explained earlier - the above algorithm, given his deterministic nature has a bad worst-case performance and is unable to achieve no-regret. More precisely, no matter what action the deterministic FTL chooses, imagine the loss sequence that puts "loss weight" of 1 to this action and 0 to the other action. Then, FTL (as well as any algorithm of deterministic nature) will get "0 points", and choosing the best *fixed* action will achieve at least $\frac{T}{2}$. This observation implies the following theorem :

Theorem 2.4.4. *Any deterministic online learning algorithm has linear worst-case regret.*

Second Idea : Fair enough. Let's put some randomness in our initial approach :

Algorithm 11 Follow-the-Leader ("Randomized" Version)

Input: Action set $A = \{A_1, A_2\}$, loss vector ℓ with all the values/losses up to step $t-1$ (considering t as the current decision-making time step).

Compute $p_t = \arg\min_{p\in[0,1]} \sum_{t'=1}^{t-1} (p \cdot \ell_t^{A_1} + (1-p)\cdot \ell_t^{A_2})$

Choose a number $p' \in [0,1]$ uniformly at random

if $p' \leq p_t$ **then**

  return A_1

else

  return A_2

end if

To avoid repeating long expressions, let :

$$h(p_t; l_t) = p_t\ell_t^{A_1} + (1-p_t)\ell_t^{A_2} \tag{2.4}$$

be the expected loss of randomized FTL at time step t by playing action A_1 with probability p_t. Also, since we have switched from the deterministic context to the randomized one, just to be absolutely clear, let's define the expected version of regret. The definition arises naturally from the deterministic version of regret. More specifically :

$$\text{Expected-Regret}(T) = \sup_{\ell_1,\ldots,\ell_t} \left(\sum_{t=1}^{T} f(p_t;\ell_t) - \min_{p\in[0,1]} \sum_{t=1}^{T} f(p;\ell_t)\right)$$

For notational convenience, let $H_t(p)$ be the cumulative loss by playing the mixed strategy $\boldsymbol{p} = (p, 1-p)$ for actions A_1 and A_2 respectively in every time step up until t (inclusive), meaning :

$$H_t(p) = \sum_{t'=1}^{t} h(p;\ell_{t'}) \tag{2.5}$$

Alright, so now that we inserted randomness to reach no-regret, are we done? As we will see in the following subsection, the answer is no.

2.4.4 Is Randomization enough? No, we need stability too!

2.4.4.1 Correlation between Stability and Regret

After getting rid of strict determinism that made our algorithm vulnerable to certain adversarial loss sequences, there is one more pathology to fix in FTL's behavior in order to - finally - achieve no-regret. In loss sequences where the "current" best action changes very often, our algorithm should somehow "stick to the plan" and change its mind (i.e. its tendency to favor the one action over the other) not after the first indication, but after stronger evidence. For example, imagine the following loss sequences for the action set $A = \{A_1, A_2\} : A_1 = \{\overset{1234567}{1010101}\ldots\}$ and $A_2 = \{\overset{1234567}{0101010}\ldots\}$. An algorithm that changes its action too often, will suffer big losses too often! In fact, as we will see, the expected regret of the FTL algorithm is upper bounded by a stability quantity related to its changes in policy, aka the probability of choosing each action. More formally, it holds that :

Lemma 2.4.5. *For any sequence of losses $\ell_1, \ldots, \ell_T$, the expected regret of FTL Algorithm is bounded above by the stability quantity $\sum_{t=1}^{T} |p_t - p_{t+1}|$. More formally, it holds that :*

$$\overbrace{\sum_{t=1}^{T} h(p_t;\ell_t) - \min_{p\in[0;1]} \sum_{t=1}^{T} h(p;\ell_t)}^{\text{Expected Regret}} \le \overbrace{\sum_{t=1}^{T} |p_t - p_{t+1}|}^{\text{Stability}}$$

Proof. The high-level plan of the proof is the following: To upper-bound the expected regret of FTL by the stability term, we will first prove that it is upper-bounded by the expected regret of a tweaked version of FTL, say BTL (from Be-The-Leader), plus our stability term. After that, we will prove that the expected regret of BTL is non-positive, so the above statement holds.

To be more precise, BTL is a hypothetical structure/algorithm. It is some sort of "enhanced version" of FTL, if the online decision-maker knew the loss vector up to ***and including*** time-step t. In that situation, the chosen action would be the one that minimizes the loss up until time-step t(inclusive). For a concrete description of BTL, you can see Algorithm 12.

Algorithm 12 Be-the-Leader ("Randomized" Version)

Input: Action set $A = \{A_1, A_2\}$, loss vector ℓ with all the values/losses up to step t (considering t as the current decision-making time step).

1: Compute $p_t = \arg\min_{p\in[0,1]} \sum_{t'=1}^{t} (p \cdot \ell_t^{A_1} + (1 - p) \cdot \ell_t^{A_2})$
2: Choose a number $p' \in [0, 1]$ uniformly at random
3: **if** $p' \le p_t$ **then**
4:     **return** A_1
5: **else**
6:     **return** A_2
7: **end if**

Lemma 2.4.6.

Expected Regret of the FTL algorithm $\le$ *Expected Regret of the BTL algorithm* $+ \sum_{t=1}^{T} |p_t - p_{t+1}|$

Proof.

$$\sum_{t=1}^{T} h(p_t;\ell_t) = \sum_{t=1}^{T} h(p_t;\ell_t) + \sum_{t=1}^{T} h(p_t;\ell_t) - \sum_{t=1}^{T} h(p_t;\ell_t) =$$

$$= \sum_{t=1}^{T} h(p_t;\ell_t) + \sum_{t=1}^{T} (p_t \cdot \ell_t^{A_1} + (1 - p_t) \cdot \ell_t^{A_2}) - \sum_{t=1}^{T} (p_t \cdot \ell_t^{A_1} + (1 - p_t) \cdot \ell_t^{A_2})$$

$$= \sum_{t=1}^{T} h(p_t;\ell_t) + \sum_{t=1}^{T} ((p_t - p_t) \cdot \ell_t^{A_1} + (p_t - p_t) \cdot \ell_t^{A_2})$$

$$= \sum_{t=1}^{T} h(p_t;\ell_t) + \overbrace{\sum_{t=1}^{T} (p_t - p_t) \cdot (\ell_t^{A_1} - \ell_t^{A_2})}^{A} \tag{2.6}$$

Now, let's try to upper bound the expression A. Actually, this is quite easy to see, since from the definition of our setting, it holds that $\ell_t^{A_1}, \ell_t^{A_2} \in [0,1]$, hence :

$$A \le \sum_{t=1}^{T} |p_t - p_t| \tag{2.7}$$

Combining 2.6, 2.7 we end up with the following :

$$\begin{aligned}
\sum_{t=1}^{T} h(p_t;\ell_t) &\le \sum_{t=1}^{T} h(p_t;\ell_t) + \sum_{t=1}^{T} |p_t - p_t| = \\
&\overset{p_t = p_{t+1}}{=} \sum_{t=1}^{T} h(p_t;\ell_t) + \sum_{t=1}^{T} |p_t - p_{t+1}| \implies \\
&\implies \sum_{t=1}^{T} h(p_t;\ell_t) - \min_{p\in[0,1]} \sum_{t=1}^{T} h(p;\ell_t) \le \sum_{t=1}^{T} h(p_t;\ell_t) - \min_{p\in[0,1]} \sum_{t=1}^{T} h(p;\ell_t) + \sum_{t=1}^{T} |p_t - p_{t+1}| \implies \\
&\implies \mathrm{E}[\text{regret of FTL}] \le \mathrm{E}[\text{regret of BTL}] + \sum_{t=1}^{T} |p_t - p_{t+1}|
\end{aligned} \tag{2.8}$$

Lemma 2.4.7.

$$\textit{Expected Regret of the BTL algorithm} \le 0$$

Proof. In a more mathematical language, we wish to prove that :

$$\overbrace{\sum_{t'=1}^{t} h(p_{t'};\ell_{t'})}^{\text{BTL's loss until time step } t} \le \overbrace{\min_{p\in[0,1]} H_t(p)}^{\text{cumulative loss of best fixed action until time step } t}$$

We will prove this by induction.
Base Case : By BTL's definition, it holds that : $p_1 = \arg\min_{p\in[0,1]} h(p;\ell_t)$. Hence :

$$h(p_1;\ell_1) - \overbrace{\min_{p\in[0,1]} h(p;\ell_1)}^{=\min_{p\in[0,1]} H_1(p)} = 0$$

Induction Hypothesis : Suppose that the cumulative loss by BTL up until time step t is less than or equal the corresponding loss by the best fixed action in hindsight, aka :

$$\sum_{t'=1}^{t} h(p_{t'};\ell_{t'}) \le \min_{p\in[0,1]} H_t(p) \tag{2.9}$$

Induction Step :

$$\begin{aligned}
\sum_{t'=1}^{t+1} h(p_{t'};\ell_{t'}) &\leq h(p_{t+1};\ell_{t+1}) + \sum_{t'=1}^{t} h(p_{t'};\ell_{t'}) \\
&\overset{\text{Induction Hypothesis}}{\leq} h(p_{t+1};\ell_{t+1}) + \min_{p\in[0;1]} H_t(p) \\
&\leq h(p_{t+1};\ell_{t+1}) + H_t(p_{t+1}) = \\
&= H_{t+1}(p_{t+1}) \implies \sum_{t'=1}^{t+1} h(p_{t'};\ell_{t'}) \leq \sum_{t'=1}^{t+1} h(p_{t+1};\ell_{t'})
\end{aligned}$$

Combining the above lemmas, we conclude that :

$$\begin{aligned}
\mathrm{E}[\text{regret of FTL}] &\leq \mathrm{E}[\text{regret of BTL}] + \sum_{t=1}^{T} |p_t - p_{t+1}| \implies \\
&\implies \overset{\mathrm{E}[\text{regret of BTL}]\leq 0}{} \mathrm{E}[\text{regret of FTL}] \leq \sum_{t=1}^{T} |p_t - p_{t+1}|
\end{aligned}$$

2.4.4.2 Correlation between Stability and Convexity

As we saw in the previous subsection, besides randomization which is essential in our pursuit of no regret, there is one more - major - issue to figure out. How do we keep the stability quantity $\sum_{t=1}^{T} |p_t - p_{t+1}|$ small? Obviously, by ensuring that FTL stays or becomes stable. Yes, but how? In fact, we can't control the stability of FTL. Just think of the adversarial loss sequences $A_1 = \{\overset{1234567}{1010101}\ldots\}$ and $A_2 = \{\overset{1234567}{0101010}\ldots\}$ from the previous subsection. Hence, to exhaust the possibilities to achieve stability, we have to think of indirect methods , maybe by "tweaking" him a little bit. But before fixing the issue, we have to identify it first.
Recall that FTL's criterion as to which action should tend to play more at time step t is the minimization of the cumulative loss up until the previous time step, $t-1$. So, the instability is generated by big "jumps" between cumulative losses of consecutive time steps. Let's give the catchphrase: The problem with the randomized version of FTL is that the minima of two consecutive cumulative loss functions, say F_{t-1}, F_t - despite the fact that they differ only in $f(p_t, \ell_t)$ - can differ a lot. For a better understanding, Figure 2.5 illustrates the case of two linear function that express the cumulative losses of consecutive time steps, with very far minima.

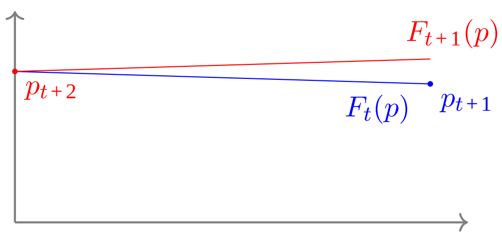


Figure 2.5: Two linear cumulative loss functions, close to each other but with far minima

Alright. Now we know the problem. Can we fix it? In fact, yes we can. If those 2 cumulative loss functions weren't linear but "very convex", then this pathology wouldn't appear. The following lemma formalises exactly this property, that 2 strictly convex functions that are close to each other have to have their minima colse to each other :

Lemma 2.4.8. *Consider 2 convex functions $f_1 : [0,1] \rightarrow \mathbb{R}$ and $f_2 : [0,1] \rightarrow \mathbb{R}$, such that $f_1''(x)$ and $f_2''(x) \geq \frac{1}{\eta}$ for all x, where their difference, say $d(x) = f_2(x) - f_1(x)$, is an $L-$ Lipschitz function, i.e. $|d(x) - d(x')| \leq L\,|x - x'|$. Then for their minima p_{f_1} and p_{f_2}, i.e. $p_{f_1} = \arg\min_{p \in [0,1]} f_1$ and $p_{f_2} = \arg\min_{p \in [0,1]} f_2$, it holds that :*

$$|p_{f_1} - p_{f_2}| \leq \eta L$$

Proof. The proof here is pretty straightforward, if we make use of a pictorial argument. More specifically, it is obvious from Figure 2.6 that it holds :

$$\Delta_1 - \Delta_2 = \Delta_3 + \Delta_4 \tag{2.10}$$

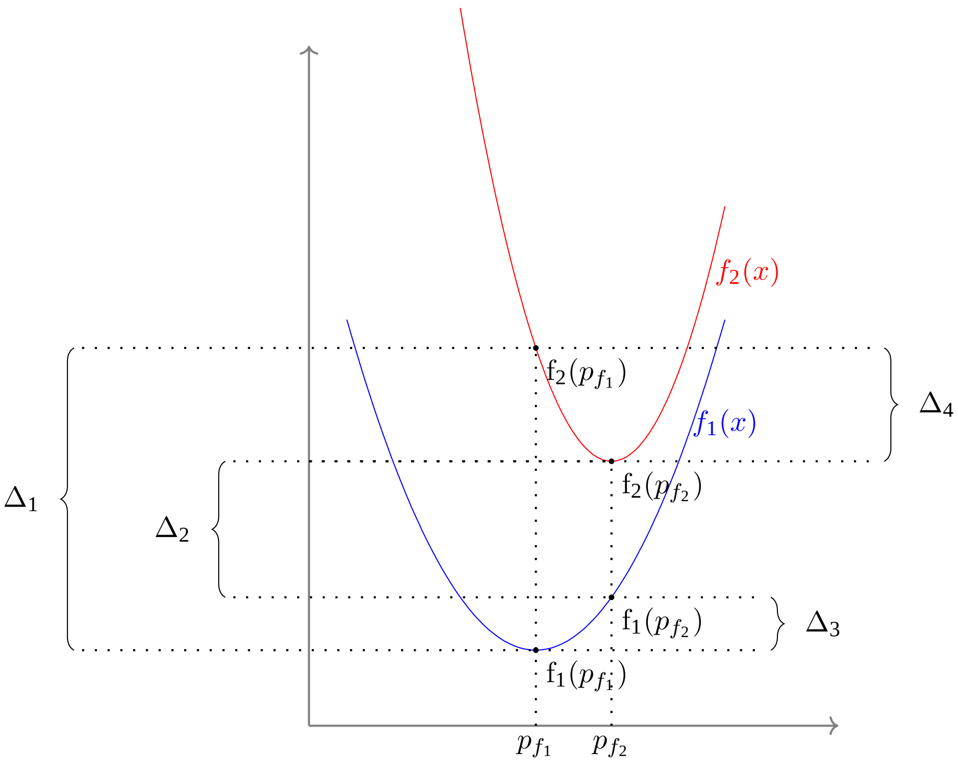


Figure 2.6: Pictorial proof of closeness of minima of convex functions

By the Lipschitzness of the distance function between f_1 and f_2, it holds that:

$$\Delta_1 - \Delta_2 = d(p_{f_1}) - d(p_{f_2}) \leq |d(p_{f_1}) - d(p_{f_2})| \overset{\text{hyp.}}{\Longrightarrow} \Delta_1 - \Delta_2 \leq L\,|p_{f_1} - p_{f_2}| \tag{2.11}$$

By the strong convexity of f_1 (more specifically from the property $f(x) \geq f(x^*) + \frac{\mu}{2}||x - x^*||^2$, where x^* is the global minimum of f and $\mu = \frac{1}{\eta}$), we have that :

$$\Delta_3 \quad \frac{1}{2\eta} \; (p_{f_1} \quad p_{f_2})^2 \tag{2.12}$$

With the same reasoning, for f_2 :

$$\Delta_4 \quad \frac{1}{2\eta} \; (p_{f_1} \quad p_{f_2})^2 \tag{2.13}$$

Combining all the above, we end up with the desired inequality :

$$\begin{aligned} \Delta_3 + \Delta_4 &\overset{2:10}{=} \Delta_1 \quad \Delta_2 \overset{2:11}{\quad} L \;\; j p_{f_1} \quad p_{f_2} j \overset{2:12;2:13}{=)} \\ &=) \;\; \frac{1}{\eta} \;\; (p_{f_1} \quad p_{f_2})^2 \quad L \;\; j p_{f_1} \quad p_{f_2} j \;\; =) \\ &=) \;\; j p_{f_1} \quad p_{f_2} j \quad \eta \;\; L \end{aligned}$$

Conclusively, what we need to do in order to achieve no-regret - besides randomizing FTL algorithm - is to "force"/modify its cumulative loss function so that it becomes strongly convex. We will present one common way to achieve this in the Entropy Regularization - related subsection 5.1.1.

2.4.5 Epimythium

What's the Epimythium?
Stability in Online Learning Version

- For the notion of regret, we switched our benchmark from the best action in hindsight to the best fixed action in hindsight(change from min $\overset{\mathbf{P}}{\underset{t=1}{}}\ \ell_t(\alpha_t)$ to $\overset{\mathbf{P}}{\underset{t=1}{}}\ \min \ell_t(\alpha_t)$). That's because at each time step, the adversary can assign weight loss equal to 1 to the action with higher probability and 0 to the other (obviously, the best action in hindsight would achieve 0 loss and our algorithm would achieve loss more than or equal $\frac{T}{2}$).
- Strict determinism hits the wall of linear regret, since the adversary can always assign weight loss equal to 1 to the chosen action and 0 to the other. Hence, the best fixed action in hindsight can always experience loss less than $\frac{T}{2}$.
- Hence, we need randomization to have any chance. But that's not enough.
- High-level, stability implies that the decision-making of the algorithm isn't very sensitive to new observations, i.e. doesn't "jump" from one action to another every time this happens : best action at time step $t \quad 1 \neq$ best action at time step t. In English (actually not so much!) that means that a stable algorithm doesn't overfit in the history.
- stability bounds regret
- strong convexity implies stability (closeness of minima of convex functions)
- randomized algorithm that fixes the stability issue of the FTL is no-regret!

Medusa

"Success is not final, failure is not fatal: It is the courage to continue that counts"
"If you are going through hell, keep going"
Uglqrml Afspafgjj

Chapter 3

The ”Proportional Allocation Rule” Principle

3.1 Motivation & setup

This section is by far the most ”Mechanism Design-oriented” section of our diploma thesis. Actually, our concern focuses on the first of the two major components of Mechanism Design, the allocation rule. For both the text and the technical part of the whole section, we are mostly based on the works of Chawla, Jagadeesan ([5]) and a joint work of them with Ilvento [21]. The context in which and for which they create their algorithmic framework is online advertising.
Before starting bombing with technicalities, why starting bombing with technicalities anyway? Is it worth it? In fact, yes it does. The algorithms produced in this line of work tend to encounter several ”fairness” issues that arise in ad auctions that take place on the Internet. Just a quick disclaimer before continuing. As you see, we chose to surround the word fairness with quotation marks. We will continue on the same pattern until defining formally this notion, and the same principle applies to ”similar” words, for reasons strictly related with the multi-aspect compulsive mentality of the author. I apologise for that in advance. Anyway, to continue our reasoning, there is a lot of empirical evidence of skewed delivery in online advertising. What do we mean by that? High-level, we want similarly ”qualified” people from different gender or racial group (or whatever - in this context), to end up seeing similar amounts of online ads - especially - in sensitive categories like housing and employment. So, skewed delivery refers to situations where this doesn't happen.
From empirical data, we can categorise the sources of unfairness to 2 main ”buckets” :

- **First Source :** When advertiser targeting is ”unfair” (maybe the advertising targeting parameters are noninclusive). This source of unfairness can be directly audited by looking at each advertiser's behavior in isolation.
- **Second Source :** When the platform's allocation algorithm creates ”unfairness”, even when the bids are ”fair”.

In this section, we will be interested exclusively with the second source of ”unfairness”. To highlight the importance of this problem in real-life settings, from empirical studies conducted in the recent past there are plenty of examples of ”unfairness” in online ads, introduced by the platforms' allocation mechanisms, like the following :

- **Employment :** It has been observed skewed delivery in ads even with gender-neutral advertisers. There is more context in the empirical study of Lambrecht and Tucker [24].
- **Housing :** It has been observed skewed delivery in ads even when advertiser targeting parameters are inclusive. Empirical evidence can be found in [1].

Before boosting our intuition on why a platform can introduce unfairness with the help of an example, let's sum up the problem that we will target in this section :

Takeaway from empirical studies : "Unfairness" can arise solely from the allocation algorithm of the platform!

As we will see (not so) later on, the contribution of both the Proportional and the Inverse Proportional Allocation algorithms is the creation of a "fairness" framework that eliminates "unfairness" introduced by the platform's allocation mechanism.
Ok, now as promised, let's see a very simple example of skewed delivery within a single ad category introduced by the platform, in order to draw some useful conclusions. Imagine the following scenario :

Example of skewed delivery within a single category :

- **2 Users :** Alice, Bob
- **2 Employment Tech Advertisers :** Multinational, Startup

First Attempt : Say that the platform's mechanism designer chooses a very standard auction format, the Highest-Bidder-Wins, as the allocation rule that determines the matchings between users and displayed ads. Now, assume the bidding scenario displayed in Figure 3.1 :

	Multinational	Startup
Alice	**1:01**	1
Bob	1	**1:01**

Figure 3.1: Example of skewed delivery within a category

What can we say, based on the above bidding profile of the 2 tech companies? Obviously, their bidding strategy is "fair" (since their bids on the 2 users are almost - but not exactly - identical - they have a slight preference over the one of the two users), but the allocation outcome isn't "fair" at all. Why is that? Because, since the highest bidder wins entirely the right to display his ad to the corresponding user, Alice sees only Multinational's online ads and Bob only Startup's online ads.
So, conclusively : *Highest-Bidder-Wins exaggerates small fuctuations in bids.*
Fine. Now we have seen a potential pathology that may arise from the platform's allocation mechanism, if the mechanism designer doesn't choose him wisely. So, let's analyze other allocation rules that may do the work better than Highest-Bidder-Wins did. Basically no. Before we do that, we haven't even defined concretely the rules of the "game" that we 'll try to figure out. In [5], the assumed model is the following :

The Model :

- Universe U of users that arrive sequentially
- n advertisers that have valuations for every user $u \in U$ (i.e. value vector on user $u, v_u = [v_u^1, v_u^2, \ldots, v_u^n]$)
- The auction allocates a single slot per user. Each advertiser gets a fraction of the slot (i.e. user u assigned allocation $[a_u^1, a_u^2, \ldots, a_u^n]$ where $\sum_{i=1} a_u^i \leq 1$)
- The allocation rule should satisfy weak monotonicity (so that incentive compatibility is achievable)

Now that we have established a concrete model to work with, let's direct our attention to the real elephant in the room. How do we get rid of the quotation marks?! Let's see.
As we mentioned earlier, our focus will be on possible "unfairness" introduced by the platform. That means that we only care about "fairness" from the perspective of the advertisers. We have assumed that somehow the advertisers' bids have been audited and are representative of their valuations on the users. To make it easy, if 2 users are identical for an advertiser, then we assume that he will place the same amount of bids on each.
In the line of work that we study, "fairness" (from the users perspective) is expressed as a stability condition. In fact, their starting point lies on the work of Dwork et al. ([15]), where they defined a notion of individual "fairness" as a situation where similar users receive similar allocations. On extending (or even better on adapting to our model) this notion, the stability requirement is the following : if 2 users receive similar bids/values from all the advertisers, then they should receive similar allocations.

Definition 3.1.1. (Value-stability). An allocation is value-stable with respect to function $f : [1, \infty) \rightarrow [0, 1]$ if the following condition is satisfied for every pair of value vectors v and v' :

$$|a_i(v) - a_i(v')| \leq f(\lambda) \text{ for all } i \in [n], \text{ where } \lambda \text{ is defined as } \max_{i \in [n]} \max\left\{\frac{v_i}{v'_i}, \frac{v'_i}{v_i}\right\}.$$

So, the catchphrase for the above definition in real-life settings (without assuming audit): The platform itself doesn't introduce any further unfairness than what may be present in the bidding profile of the advertisers.
Even though they are quite evident, just to be clear, let's summarize the main desiderata that our distance function $f(\cdot)$ should satisfy:

- $f(1)$ should be equal to 0 (since $\lambda = 1$ means that $v = v'$, so advertiser i should get the exact same amount under the 2 different instances/value vectors).
- $f(\infty)$ should be equal to 1 (since $\lambda \rightarrow \infty$ means that the 2 value vectors are arbitrarily different, hence we should allow each advertiser to get arbitrarily different allocations).
- $f(\cdot)$ should be increasing (since as λ grows, the 2 instances/value vectors become more and more different, so the stability/fairness constraint should be loosen).

Before analyzing the allocation mechanisms of this section, one more definition fist. Maybe it is useful, maybe not.

Definition 3.1.2. (Approximation Ratio). The approximation ratio of an (allocation) algorithm is the percentage of social welfare it achieved compared to the highest social welfare achievable. In our setting and in the language of maths :

$$\text{Approximation Ratio} := \frac{\sum_{u \in U} \sum_{i=1}^{n} a_u^i v_u^i}{\sum_{u \in U} \max_{i \in [n]} v_u^i}$$

Obviously, our goal is to design value-stable allocation algorithms that achieve a (near-)optimal approximation ratio.

3.2 Proportional Allocation Mechanisms

3.2.1 Some Intuition Pump

As we saw earlier, the choice of the Highest-Bidder-Wins allocation rule, despite the fact that it is welfare maximizing, is very far away from any notion of value stability - hence from our notion of value-stability, as defined above, too.
Second Idea : Allocate to each agent a fraction of user's ad slot proportional to her value. That means, agent i receives : $\frac{v_u^i}{\sum_{j=1} v_u^j}$. Alright, this approach at least doesn't seem doomed in the sense that it doesn't leave zero-allocation agents and also for each pair of agents i, j it holds that : $a_i \geq a_j \iff v_i \geq v_j$. So, there is some sort of fairness and efficiency involved. In fact, the following theorem holds :

Theorem 3.2.1. *The Proportional Allocation Mechanism is value-stable with respect to* $f(\lambda) = 1 - \frac{1}{\ }$

But, there is a major problem with this allocation rule. The fraction allocated to each agent grows (almost) inverse proportionally to n (where n denotes the number of agents). And why is that a problem? Because, obviously, as $n \to \infty$ every single agent gets almost nothing, so approximation ratio $\to 0$. To see this, let's consider the following "adversarial" instance : $v_u = [1, \epsilon, \ldots, \epsilon]$. Then, the proportional allocation rule will end up with the allocation : $a_u = [\frac{1}{1+(n-1)\epsilon}, \frac{\epsilon}{1+(n-1)\epsilon}, \ldots, \frac{\epsilon}{1+(n-1)\epsilon}]$.
Hmm, maybe we could get away with the same idea after some refinement. Like what ? While sticking to the proportional allocation logic, what if we increase the degrees of freedom, so that we can achieve higher efficiency (at the expense of "fairness")? Seems legit. Let's take a closer look.
Second Idea with some refinement : Allocate to each agent a "tuned" fraction of user's ad slot proportional to her value. That means, agent i receives : $\frac{(v_u^i)^\ell}{\sum_{j=1} (v_u^j)^\ell}$. This ℓ parameter generates in some degree stability-efficiency tradeoffs. That means that as ℓ grows, the allocation becomes more and more concentrated (uniform $\to$ Highest-Bidder-Wins). For experimental evidence that enhance the claims above we refer to Figures 5.1, 5.2, 5.3, 5.4. Actually, things are even worse. It is easy to see that this caveat appears no matter what function of v we choose. So, we need to escape from the "distribute somehow proportionally" logic.
Nevertheless, just for reasons of completeness and in order to create a stepping stone for the Inverse Proportional Allocation mechanism, we mention the generalised version of the Proportional Allocation Mechanism in the section below. After that, we will pass to the protagonist of the current section, the Inverse Proportional Allocation mechanism.

3.2.2 The Proportional Allocation Algorithm for the Single-Item Case

Below, we present the generalised version of the Proportional Allocation mechanism, aka the Proportional Allocation mechanism "equipped" with a "general" function of the values. Since this mechanism simply serves as a stepping stone for the Inverse Proportional Allocation Mechanism, we find it needless to analyse it furthermore.

Algorithm 13 Proportional Allocation Mechanism parameterized by $g()$

Input: Function $g() : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ continuous, supper-additive (i.e. $g(x) + g(y) \leq g(x + y)$), increasing function. Values $v_1, \ldots, v_n$.

for $i \in [n]$ **do**
  Set $a_i = \frac{g(v_i)}{\sum_{j=1} g(v_j)}$
end for
return a.

3.3 Inverse Proportional Allocation Mechanisms

3.3.1 Some Intuition Pump

At first, I have to apologise for the disproportion between the 2 subsections of this section, meaning the Proportional Allocation-related and the Inverse Proportional Allocation-related (which is quite ironic if you think that the topic of discussion is proportional allocations, but anyway). There are 2 main reasons for this. First, the concept of Inverse Proportionality is more obscure and - probably - interesting than the more "natural" and straightforward concept of Proportionality. Second, this algorithm (or more precisely this class of allocation algorithms) was the starting point of our diploma thesis. Both Mr.Fotakis & I have spent more working hours and brain power on analyzing the logic and performance of the Inverse Proportional Allocation algorithm than the Proportional Allocation one.
After this disclaimer of "procedural character", let's dive into our analysis. More specifically, our initial intuition was that the allocation mechanism produced by this family of algorithms - that we will call IPA, as acronym for Inverse Proportional Allocation - was a result of the "Linear Programming" principle (to avoid repeatability, for more context on this principle see the corresponding section). Nonetheless, given IPA's suboptimal performance (a topic that we will discuss in more details later on in this section), at the moment that this thesis was written we strongly believe that the "Linear Programming" principle *is not* the generative principle of IPA. This . Furthermore, our conjecture is that the class of IPA mechanisms was produced either by some ad hoc "inspiration" while trying different stuff or simply by a reasoning of the form "instead of giving proportionally, cut inverse proportionally". My personal opinion slants towards the second option. That is the main reason that I chose to place both PA and IPA mechanisms under the "Proportional Allocation Rule" umbrella. We have compunctions of technical nature on the proportional logic behind IPA, because - as we will see explicitly soon - the algorithms final outcome contains - possible "a lot" - of zero-allocation bidder, a feature that contradicts proportionality. Maybe, finding (and proving) the origins behind IPA could be an interesting line for future work.
But wait. This, definitely, isn't the right order to start the discussion. We don't have even

made clear how the IPA algorithm works, and we are talking about the principle behind him! Let's proceed to the mechanism's core idea and we will give more context afterwards.
Key Idea : Since allocating proportionally to the values doesn't work well, try to deduct proportionally to the inverse of the values. An attempt for visualization of this idea can be found in Figure 3.2, below.

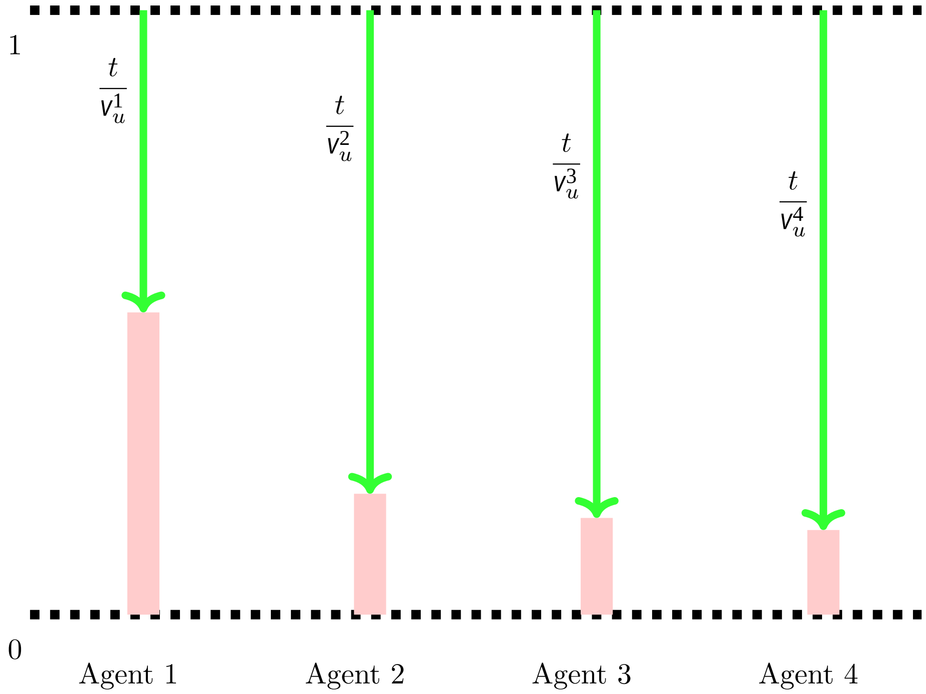


Figure 3.2: Allocation deductions "aligned" with the IPA logic

Before jumping into the main analysis and start analyzing a plethora of IPA-related theorems, maybe it is useful to give an informal description of our mechanism. Actually, we will give different formulations of the mechanism (probably the second one is conceptually easier to understand than the first, although both are quite straightforward).

- **Algorithmic Formulation of the IPA :** The quintessence of the IPA algorithm is to find the final number of active agents (meaning those who will receive non-zero fraction of the item). Obviously, the highest bidder should always belong to the set of active agents (in fact he should always receive the greatest fraction of the item among all the agents). Actually, let's put it this way : initially all the agents should be considered as (potentially) active. The final allocation of an active agent, say i, equals to $a_i = 1 \quad \frac{\#\text{active agents}}{\sum_{j \in \text{active}} g(v_j)} \quad g(v_i)$. Hence, the algorithmic version of IPA proceeds as follows : For convenience, let's relabel the agents such that : $v_1 \quad v_2 \quad \ldots \quad v_n$. Starting from the lowest bidder, the Algorithm checks if $a_n \quad 0$ (for #active agents $= n$) If yes, then the algorithm stops. Else, the $n \quad th$ bidder will receive zero allocation (so #active agents) and the algorithm proceeds to the next highest bidder. The whole process continues until finding an agent, say i, such that $a_i \quad 0$. The implementation of all the above in pseudocode can be found in Algorithm 14.

- **Alternative Formulation of the IPA:** Ok, so the formulation is the following: Let $\mathbf{v}$ be the standing value vector of the instance. Let $y_i(t) = \max(0, 1 - tg(v_i))$ for agent $i \in [n]$ and $y(t) = \sum_{i \in [n]} y_i(t)$, for $t \geq 0$. We are, essentially, looking for the "right" time t, where $\sum_{i \in [n]} y_i(t) = 1$. Let's denote this time by t^*. So, the final allocation is defined as $\mathbf{a}(\mathbf{v}) = (y_1(t^*), y_2(t^*), \ldots, y_n(t^*))$. To gain more intuition let's rephrase it in English: We start the whole process by allocating 1 to every agent. We gradually decrease the allocation of every agent, proportionally to $g(v_i)$ by increasing t. If an agent "hits" 0 allocation, he remains with 0 until the end. We stop when the total allocated mass equals to 1.

3.3.2 The Inverse Proportional Allocation Algorithm for the Single-Item Case

At first, we will present the Inverse Proportional Allocation Algorithm in a more concrete way, and afterwards we will analyze both its performance and its value-stability guarantees in a deep level.

Algorithm 14 Inverse Proportional Allocation Mechanism parameterized by $g()$

Input: Function $g() : \mathbb{R}_{\geq 0} \to (0, \infty]$ with $g(0) = \infty$ and $\lim_{x \to \infty} g(x) = 0$, decreasing function. Values $v_1, \ldots, v_n$.

```
Initialize a_i = 0 for 1 ≤ i ≤ n
Sort the values so that v_1 ≤ v_2 ≤ ... ≤ v_n.
if v_n = 0 then
    Set a_i = 1/n for all 1 ≤ i ≤ n.
    return a.
end if
Initialize s = min({i ∈ [n] | v_i > 0}).
while 1 − (n − s) g(v_s) / Σ_{j=s} g(v_j) ≤ 0 do
    s++
end while
for i ≥ s do
    Set a_i = 1 − (n − s) · g(v_i) / Σ_{j=s} g(v_j)
end for
return a.
```

So, after defining the protagonist of our section let's try to analyze its properties a bit further. One of the most fundamental properties of IPA mechanism is that it produces monotone allocations. Why is that so important? Besides giving us the possibility to apply Myerson's Lemma in order to design payment rules for truthful mechanisms, it helps us proceed with the value stability analysis. More precisely, for proving value stability, we are looking for value vectors that are multiplicatively close to the current value vector of our instance and changes the allocation to the greatest extent. In order to find that value vector we have to use the monotonicity property and argue that to minimize $i-$th agent's valuation we have to minimize its valuation. The following lemma expresses mathematically this property :

Lemma 3.3.1. *For any value vector $\mathbf{v}$ and any $i \in [n]$, let $\mathbf{v}' = (v_i', \mathbf{v}_{-i})$ be another value vector that differs from $\mathbf{v}$ only in coordinate i with $v_i > v_i'$. Then it holds that $a_i(\mathbf{v}) \geq a_i(\mathbf{v}')$ and $a_j(\mathbf{v}) \leq a_j(\mathbf{v}')$ for all $j \neq i$.*

To be more precise about what we call Inverse Proportional Allocation Mechanism, let's align our definition with the one used in the basic paper of this subsection ([5]).

Definition 3.3.1. (Inverse Proportional Allocation Algorithm). For $\ell \in (0, \infty)$, the Inverse Proportional Allocation Algorithm with parameter ℓ is Algorithm 14 with g defined as $g(x) = x^{-\ell}$ for $x \in [0, \infty)$.

The above allocation mechanism achieves value-stability, as stated in the theorems below :

Theorem 3.3.2. *For any $\ell > 0$ and any number $n > 0$ of advertisers, the inverse proportional allocation algorithm with parameter ℓ is value stable with respect to any function f that satisfies $f(\lambda) \geq f_\ell(\lambda) \equiv 1 - \lambda^{-2\ell}$ for all $\lambda \in [1, \infty)$.*

Theorem 3.3.3. *For any function f and parameter $\ell > 0$ such that over any number of advertisers $k > 0$, the IPA with parameter ℓ is value stable with respect to f, it holds that $f(x) \leq f_\ell(x)$ for all $x \geq 1$.*

But the question still remains. Does the transition from PA mechanisms to IPA really worth it? Since we have lost on value-stability, have we gained anything in terms of social welfare? The following theorem shows that the answer to the above is positive.

Theorem 3.3.4. *For any $\ell \in (0, \infty)$, the inverse proportional allocation algorithm with parameter ℓ obtains a worst case approximation ratio for social welfare of at least :*

$$a_\ell := \min_{x \in (0,1)} (1 - x^{\ell} + x^{\ell+1}) = 1 - \frac{1}{\ell+1}\left(\frac{\ell}{\ell+1}\right)^{\ell}$$

Just a semi-mathematical and semi-intuitive explanation for the above result. At first the key idea of the proof is to consider that the worst case value vector is of the form (after normalisation) $[1, \gamma, \gamma, \ldots, \gamma]$ (because every other value vector can be "transformed" to a value vector of this form without losing any more welfare, hence the greedy exchange argument holds).

A very important observation is that, as we see, IPA's approximation ratio is independent of the number n of advertisers, and this is one of the most - if not the most - important characteristics that distinguishes IPA allocation rules from the PA ones. More specifically, as we mentioned earlier the major pathology of PA mechanism is that every advertiser will necessarily receive some positive fraction of the item. This is the reason that PA's performance gets worse and worse as the number n of advertisers becomes larger. On the other hand, IPA by "creating" zero-allocation bidders (the lower ones), is a more robust mechanism.

Actually, often a good way to understand an algorithm's general behavior is to consider its limiting behavior. As far as it concerns IPA's behavior, we observe the following (reminder $g(x) = x^{-\ell}$):

- As $\ell \to 0$, IPA $\to$ Uniform, hence approximation ratio $\to 0$ and value-stability $\to$ super strong.
- As $\ell \to \infty$, IPA $\to$ Highest-Bidder-Wins, hence approximation ratio $\to 1$ and value-stability $\to$ super weak.

- Just to highlight another interesting case: For $\ell = 1$, IPA achieves a constant approximation ratio of $\frac{3}{4}$.

IPA's Value-Stability Proof Sketch :
High-Level Version :

What we need to do is to analyze how the allocation vector $\boldsymbol{a}$ changes as the valuation vector v changes. How?

- **Step 1:** Fix a valuation vector v. Find a valuation vector v' that is multiplicatively close to v (aka fix $n - 1$ agents, "perturb" the valuation of just a single agent, say i, such that $v'_i \in [\frac{1}{\lambda}, \lambda] \cdot v_i$) and changes the allocation by the greatest amount.
- **Step 2:** Bound the corresponding change to the allocation.

But : The challenging part is that the set of active agents changes along with the value vector. So, what can we do ?

High-Level Idea : Construct modified allocations where the sets of active agents are much easier to compare. Afterwards :

- Prove that the modified allocations are no closer than the real ones (hence, bounding the modified allocations gives us an upper bound to the original problem).
- Prove that the modified allocations are within $f(\lambda)$ with each other.

Now we have a quite good understanding of how IPA works and his value-stability and efficiency performance. One more loose end to tie up and then we will display our experimental results. Is it really worth it? I mean, obviously it is a finer idea than proportional allocation mechanisms, but those consist a somehow "vanilla" way of achieving value-stability anyway. Do they perform well compared to "fancier" mechanisms? Well, the short answer is yes.

More specifically, for the family of value-stability constraints as defined earlier, IPA achieves optimal approximation ratio for "large" number of agents (to be more precise, there is no prior-free, value-stable with respect to that family of constraints allocation mechanism that performs better than IPA's approximation ratio $- \frac{1}{n}$):

Theorem 3.3.5. *For the value-stability constraint $f(\lambda) = 1 - \lambda^{-2\ell}$, IPA with $g(x) = \frac{1}{x^\ell}$ achieves the optimal approximation ratio for social welfare as $k \to \infty$.*

Moreover, it can be proved that IPA achieves near-optimal approximation ratio for "general" value-stability constraints.

Theorem 3.3.6. *For any f satisfying a mild condition, there exists an IPA algorithm that is value-stable and achieves a worst case approximation factor :*

$$\Omega(\frac{a_f}{(1 + \log(1 + a_f))})$$

3.4 Epimythium

What's the Epimythium?

"Proportional Allocation Rule" Principle Version

- Core idea of PA mechanisms : Super simple. Allocate to every agent proportionally to (some increasing function of) her value.
- Caveat of Proportional Allocation Mechanism : Since every (non-zero) bidder will receive some fraction of the item, PA's performance depends on the number of advertisers. More specifically, *approximation ratio* ! 0 as n ! *1* .
- Core idea of IPA mechanisms : Rather than allocate proportionally to the values, deduct proportionally to the inverse of values.
- Comparative advantage of IPA to PA mechanisms : IPA's approximation ratio is independent of the number of advertisers. This is because, under PA allocation, even very low bidders will receive a fraction of the item, contrary to the IPA allocation, where they will receive nothing (and the proportion that they would receive from PA "shifts" to higher bidders). So, IPA achieves both value-stability and high social welfare, even if #*agents* ! 0.

Strive for Greatness

"The best way to predict the future is to create it"
"I do the very best I know how, the very best I can, and I mean to keep doing so until the end"
?`p_f_k Jglamjl

Chapter 4

The "Linear Programming" Principle

4.1 The setup

Now that we have familiarized ourselves with the "Proportional Allocation Rule" principle, and mainly with the Inverse Proportional Allocation mechanism, it's time to introduce another perspective for the same problem (i.e. ways to achieve highly efficient and value-stable allocations in auction settings), the "Linear Programming" principle. This section is the "conceptually" easier to understand section of our diploma thesis, and in the same time the one that contains the core of our work. For this section, We won't give any references to related literature, since it, actually, constitutes of an "ad hoc" attempt to approach the problem.
Just to give the chronological order of our thought process, we began our diploma thesis with a conjecture. The conjecture was that the IPA mechanism was produced as the optimal solution of the linear program displayed below (LP1). The program is super simple : maximize the social welfare while allocating fractions of a single item (or probabilities if you like), subject to a bunch of "fairness"/stability constraints. More specifically, these constraints enforce that 2 consecutive bidders (given an ordering based on their bids) should receive similar values from the mechanism. The degree of similarity depends on the value vector (the difference between their valuations) and also the "tightness" of the $g()$ function.
Before becoming more specific, just a sketch of the structure of this section. After defining the Linear Program that we wish to solve optimally, we introduce a new "family" of allocation mechanisms that will do the work, along with the corresponding proofs.

LP1(Single-Item)

$$\begin{aligned} \text{maximize} \quad & \sum_{i \in N} x_i \cdot v_i \\ \text{subject to} \quad & \sum_{i \in N} x_i = 1, \\ & x_i \cdot v_i - x_{i+1} \cdot v_{i+1} \leq g(v_i, v_{i+1}), \forall \text{ agent } i \in [0, \ldots, N-1], \\ & x_i \in [0,1], \forall \text{ agent i } \in N \end{aligned}$$

4.2 Our results

4.2.1 Theoretical Analysis

4.2.1.1 Introducing a new class of allocation algorithms for the Single-Item Case : The "Keep Consecutive Scales Tight" Family

In order to solve LP1, we introduce a new allocation mechanism. Besides solving optimally this problem, this mechanism will serve as benchmark for other allocation mechanisms that we examine in this diploma thesis (the Proportional, the Inverse Proportional and the Negative Entropy Regularizer).
So, the main logic is the following : Since we have the value-stability constraints, we are obviously obliged to satisfy them. But, since our "concern" - besides feasibility - is the welfare maximization, why not satisfy them tightly? I mean, from all the feasible solutions, we prefer the ones that favor the highest valuation players. So, why don't we give as much as possible fraction of the item to the "highest" player (we will talk in more details about this in a moment) - satisfy the first constraint tightly, i.e. between agents 1 and 2 - then, after satisfying the first constraint - give as much as possible fraction of the item to the second highest valuation player, who is the highest valuation *available* one and continue the same process until "exhausting" the item, leaving (potentially) the lowest bidders with zero allocations. So, high-level, what we try to do is to concentrate all the allocated mass to the highest bidders, but because of the stability constraints, some of this mass "gravitates away" from them and goes to the lower ones too. An attempt to visualise the allocation deductions based on this logic is shown in Figure 4.1.

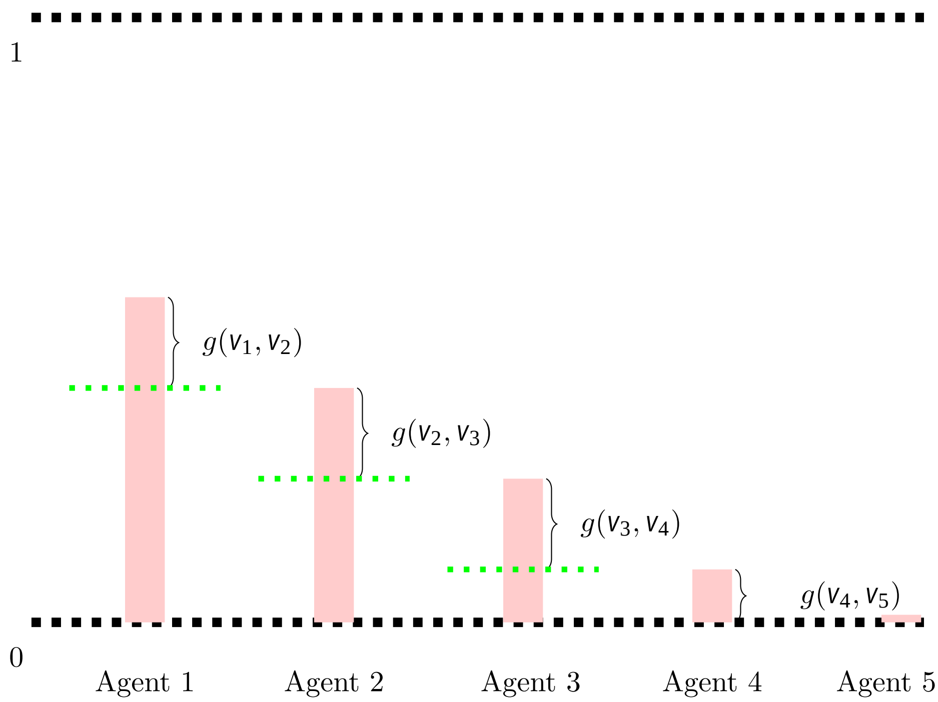


Figure 4.1: Allocation deductions "aligned" with the KCST logic

One more loose end to tie up and afterwards we will give a complete description of our algorithm, both in natural language and in pseudocode. What fraction of the item should we give to the highest (valuation) agent. If we give him too much, then as we "go down the stairs" satisfying marginally lower agents, we might run out of item without satisfying some of them. On the other hand, if we give him too little, there is obviously the danger of suboptimality/inefficiency. Actually, the right question isn't what to give but how to decide what to give. In fact, the answer arises naturally : "Binary search". Give him a random fraction of the item, say x_1, and keep satisfying the consecutive scales tightly. If some agents gets negative allocation, give him zero allocation and stop the process. If you run out of the item, decrease x_1 by half. If there are unallocated leftovers, increase x_1 by half. In both situations, try again. A formal description of our algorithm in pseudocode is shown below (Algorithm (15)).

Algorithm 15 "Keep Consecutive Scales Tight"(KCST) Allocation Mechanism parameterized by $g()$

Input: Matrix $1 \quad n$ of n non-negative players' valuations for the item
(WLOG we relabel the valuations, such that : $v_1 \quad v_2 \quad \dots \quad v_n$)
Output: Allocation $x(v)$

Let α be a random real number chosen uniformly from the interval $[0,1]$.
$lowerBound \quad 0$;
$upperBound \quad 1$;
while (true) **do**
    Set $x_1 = \alpha$;
    $s \quad 2$;
    $x_s \quad \frac{x_1 \, v_1 \quad g(v_1, v_s)}{v_s}$;
    while $x_s \quad 0$ **do**
        $s{+}{+}$;
        $x_s \quad \frac{x_{s \; 1} \, v_{s \; 1} \quad g(v_{s \; 1}, v_s)}{v_s}$;
    end while
    for $i \; 2 \; [s, n]$ **do**
        $x_i \quad 0$;
    end for
    if $\sum_{i=1} x_i > 1$ **then**
        $a \quad \frac{lowerBound + x_1}{2}$;
        $upperBound \quad x_1$;
    else if $\sum_{i=1} x_i < 1$ **then**
        $a \quad \frac{x_1 + upperBound}{2}$;
        $lowerBound \quad x_1$;
    else
        return $x(v)$
    end if
end while

4.2.1.2 Proof of optimality of "Keep Consecutive Scales Tight" for the Single-Item Case

Theorem 4.2.1. *"Keep Consecutive Scales Tight" algorithm outputs optimal allocations for LP1, for every distance function $g()$.*

Proof. Let SOL denote the solution that our algorithm returns, and OPT an optimal solution for LP1. Consider the allocations $\mathbf{x}^{\text{ALG}} = [x_1^{\text{ALG}}, \dots, x_n^{\text{ALG}}]$ and $\mathbf{x}^{\text{OPT}} = [x_1^{\text{OPT}}, \dots, x_n^{\text{OPT}}]$ that our algorithm and the optimal algorithm computes, respectively. Also, let m and m' denote the number of non-zero allocation/active bidders in SOL and OPT, respectively.
Before diving into the main analysis, we first prove the following - super useful - Lemma :

Lemma 4.2.2. *Every algorithm that solves optimally LP1, keeps tight the constraints between consecutive agents starting from the highest valuation one, until some agent gets satisfied by receiving zero allocation or we run out of agents.*

Proof. We will make use of the **"Greedy Exchange"** technique.
Let's assume that SOL and OPT differ for the first time between the $k-$th & the $(k+1)-$th agent, $k \in [1, n-1]$. Let $x_k^{\text{OPT}} \cdot v_k - x_{k+1}^{\text{OPT}} \cdot v_{k+1} = \alpha' < g(v_k, v_{k+1})$ (since we assumed that the constraint between $k-$th and $(k+1)-$th agents is not satisfied tightly). Let $\epsilon = \min\{\frac{g(v_k, v_{k+1}) - \overbrace{x_k^{\text{OPT}} \cdot v_k + x_{k+1}^{\text{OPT}} \cdot v_{k+1}}^{\alpha'}}{v_{k+1}}, \frac{x_{k+1}^{\text{OPT}} \cdot v_{k+1} - x_{k+2}^{\text{OPT}} \cdot v_{k+2}}{v_{k+1}}\}$ (a.k.a. the quantity of the item that we need to cut in order to make the constraint between the $k-$th & the $(k+1)-$th agent tight or/and the valuations between the $(k+1)-$th & the $(k+2)-$th agent equal). Let's follow the "Water-Level" algorithmic logic in order to redistribute the quantity ϵ. More precisely, for bidders $i, i+1$ with $i \leq k$, say that we increase the allocation of agent i by δ (let's imagine that δ expresses the rate of increase). We want to keep the constraints between consecutive bidders tight, hence the rate of increase of agent's $i+1$ valuation, should be the following :

$$
\begin{aligned}
(x_i + \delta') \cdot v_i - (x_{i+1} + \delta) \cdot v_{i+1} = g(v_i, v_{i+1}) \implies \\
\implies (x_i + \delta') \cdot v_i = g(v_i, v_{i+1}) + (x_{i+1} + \delta) \cdot v_{i+1} \implies \\
\implies \delta' = \frac{g(v_i, v_{i+1}) + (x_{i+1} + \delta) \cdot v_{i+1} - x_i \cdot v_i}{v_i}
\end{aligned} \tag{4.1}
$$

Hence, we reallocate the quantity ϵ between consecutive bidders in $\{1, \dots, k\}$ with the rates implied by 4.1 until we run out of quantity "for reallocation" (aka $\epsilon \to 0$). Let $\mathbf{x}'$ denote the new allocation that resulted from the redistribution of ϵ. Obviously this allocation is feasible, since :

- $\sum_{i=1}^{n} x'_i = \sum_{i=1}^{n} x_i^{\text{OPT}} = 1$(since we simply redistributed already allocated fraction of the item),
- $x'_i \cdot v_i - x'_{i+1} \cdot v_{i+1} = g(v_i, v_{i+1}), \forall i \in \{1, \dots, k-1\}$ (since we maintained the constraints tight between higher bidders),
- $x'_k \cdot v_k - x'_{k+1} \cdot v_{k+1} < g(v_k, v_{k+1})$, (since bidder k has lost some value and we didn't touch bidder's $k+1$ allocation) and

$x'_i v_i - x'_{i+1} v_{i+1} \geq g(v_i, v_i+1), \forall i \in \{k+1, \dots, n\}$ (since we didn't "touch" the allocations of lower bidders)

Also, the (feasible) allocation that results from the above redistribution, achieves higher social welfare than OPT, since we reallocate some fraction of the item to the same agent and to agents with higher valuations. The above violates the optimality of x^{OPT}, a contradiction.

Now, we can proceed to the main part of the proof. We consider the following cases :

Case 1 ($x_1^{\text{OPT}} > x_1^{\text{SOL}}$)(depending on $g()$, the allocation may be feasible) :

From the above Lemma, we conclude that the optimal allocation should keep tight the constraints between consecutive bidders. Furthermore, it must hold that $m^* < m$, since otherwise the allocation becomes infeasible. More precisely, fix an agent $i \in [m^*]$. It holds that :

$$\begin{aligned}
x_i^{\text{OPT}} &= \frac{x_{i-1}^{\text{OPT}} v_{i-1} - g(v_{i-1}, v_i)}{v_i} = \frac{\frac{x_{i-2}^{\text{OPT}} v_{i-2} - g(v_{i-2}, v_{i-1})}{v_{i-1}} v_{i-1} - g(v_{i-1}, v_i)}{v_i} = \\
&= \frac{x_{i-2}^{\text{OPT}} v_{i-2} - g(v_{i-2}, v_{i-1}) - g(v_{i-1}, v_i)}{v_i} = \dots = \\
&= \frac{x_1^{\text{OPT}} v_1 - g(v_1, v_2) - g(v_2, v_3) - \dots - g(v_{i-1}, v_i)}{v_i} > \\
&> \frac{x_1^{\text{SOL}} v_1 - g(v_1, v_2) - g(v_2, v_3) - \dots - g(v_{i-1}, v_i)}{v_i} \Rightarrow \\
&\Rightarrow x_i^{\text{OPT}} > x_i^{\text{SOL}}, \forall i \in [m^*]
\end{aligned} \tag{4.2}$$

Hence :

$$\sum_{i=1}^{m^*} x_i^{\text{OPT}} > \sum_{i=1}^{m^*} x_i^{\text{SOL}} \overset{[m] \subseteq [m^*]}{=} \sum_{i=1}^{m} x_i^{\text{SOL}} = 1, \text{ a contradiction.}$$

Therefore, $m^* < m$. But under this assumption, the fairness/stability constraints are violated. Indeed, for the m^*-th, (m^*+1)-th highest bidders, it holds that :

$$\overset{4.2}{\to} x_{m^*}^{\text{OPT}} > x_{m^*}^{\text{SOL}} \tag{4.3}$$

$$\begin{aligned}
&x_{m^*+1}^{\text{OPT}} = 0 < x_{m^*+1}^{\text{SOL}} (\text{since } m^* < m) \Rightarrow \\
&\qquad \Rightarrow -x_{m^*+1}^{\text{OPT}} v_{m^*+1} > -x_{m^*+1}^{\text{SOL}} v_{m^*+1}
\end{aligned} \tag{4.4}$$

$$\overset{4.3;4.4}{\to} x_{m^*}^{\text{OPT}} v_{m^*} - x_{m^*+1}^{\text{OPT}} v_{m^*+1} > x_{m^*}^{\text{SOL}} v_{m^*} - x_{m^*+1}^{\text{SOL}} v_{m^*+1} = g(v_{m^*}, v_{m^*+1}), \text{ a contradiction.}$$

Case 2 ($x_1^{\mathbf{OPT}} = x_1^{\mathbf{SOL}}$) :
This is the easiest case. By the above Lemma, the optimal allocation should keep tight the constraints between consecutive bidders (the exact same way SOL behaves), and since $x_1^{\text{OPT}} = x_1^{\text{SOL}}$, we conclude that OPT   SOL.

Case 3 ($x_1^{\mathbf{OPT}} < x_1^{\mathbf{SOL}}$) :
Again, with the aid of the above Lemma, in order to achieve optimal allocation, we should keep tight the constraints between active bidders. With the same reasoning as in Case 1, for every active bidder in SOL, it holds that :

$$x_i^{\text{OPT}} < x_i^{\text{SOL}}, 8i\ 2\ [m]$$

Also, using similar arguments as in Case 1, it holds that $m \ < m$.

Hence, the social welfare achieved by OPT is strictly less than the one achieved by SOL, which contradicts the optimality of OPT.

4.3 Epimythium

What's the Epimythium?
"Linear Programming" Principle Version

- **KCST in a nutshell :** Give the highest valuation agent an arbitrary fraction of the item, say x_1. Continue allocating by "keeping consecutive scales tight". If you reach a negative-allocation agent, give him zero allocation and stop allocating. If the total allocated mass is greater than 1, halve x_1 and repeat. If the total allocated mass is less than 1, double x_1 and repeat.
- KCST produces optimal and monotone allocations, with respect to LP1.

Ithaca

”The best way to pay for a lovely moment is to enjoy it”
”What the caterpillar calls the end, the rest of the world calls a butterfly”

Pgaf_pb @_af

Chapter 5

The "Slightly or Completely Modify the Objective Function" Principle

5.1 The Entropy Regularization Method

The main point of this "Entropy Regularizer"-related subsection is to understand why "enhancing" the cumulative loss function $H_t(\cdot)$ of the randomized version of Follow-the-Regularized-Leader (in the Online Learning Setting as we saw in the Preliminaries) with an additional entropy term, can lead to the highly anticipated stability! Actually, before that, just a quick reminder. For reasons of completeness, we mention that - as we proved in the Preliminaries section - the expected regret of FTL (which we wish to minimize) is upper-bounded by a stability term. This stability term is additive to the differences between "probabilities of playing a particular action" from consecutive time steps. To upper-bound this sum of differences, we would like the minima of consecutive time steps to be close to each other. This property is satisfied if we - somehow - manage to "tweak" $H_t(\cdot)$ to assure that it is strongly convex. Well, let's continue to reverse engineer a little more. From Lemma 2.4.8, - i.e. the closeness of minima of convex functions lemma in the Preliminaries section - if $F_t(\cdot)$ happened to be strictly convex, meaning had second derivatives bounded from below by $\frac{1}{\eta}$, then $\sum_{t=1}^{T} |p_t - p_{t+1}| \leq \eta \cdot T$. Why is that?

Since $H_t(p) - H_{t-1}(p) = h_t(p) = p_t \cdot \ell_t^{A_1} + (1 - p_t) \cdot \ell_t^{A_2} \overset{\ell_t \in [0;1]}{\leq} p_t + (1 - p_t) = 1 \cdot |t - (t-1)|$, meaning that $H_t(\cdot)$ is $1-$Lipschitz, we can apply Lemma 2.4.8 to the minimum of $H_t(\cdot)$ at each time step and conclude that : $|p_t - p_{t+1}| \leq \eta \cdot 1, \forall t \in [1, T]$. By adding each time step, we get the upper-bound for the stability term. Of course, η is a parameter that we can control. By choosing η to get values inversely proportional to a power of T, we get sub-linear regret for the "tweaked" version of FTL.
Now, that we have - somehow - explained why, let's see how. Let's consider a stricly convex function $R(p)$, who satisfies: $R''(p) \geq 1, \forall p \in [0, 1]$. If we "tweak" our initial cumulative function like : $\hat{H}_t(p) = H_t(p) + \frac{1}{\eta} R(p)$, then $H_t''(p) \geq \frac{1}{\eta}$.
Conclusively, if our FTL algorithm adapts his mixed strategy by considering the "tweaked" version of the initial cumulative loss function, i.e. $\hat{H}_t(p)$, then we have achieved strong convexity and as a result, at each time step t, p_t and p_{t+1} have to be close to each other. In the following

subsection, we will unify all the above in Algorithm 16 and afterwards we will proceed to some - brief - theoretical analysis in order to explain the benefits of the entropy regularizer.

5.1.1 Follow the Regularized Leader: Stability through Convexity through Regularization

The randomized version of Follow-the-Leader, modified such that at each time step, the expected loss of his mixed strategy includes its ("weighted") entropy, is shown in pseudocode below :

Algorithm 16 Follow-the-Regularized-Leader ("Randomized" Version)

Input: Action set $A = \{A_1, A_2\}$, loss vector ℓ with all the values/losses up to step $t-1$ (considering t as the current decision-making time step).

1: Compute $\hat{p}_t = \arg\min_{p \in [0,1]} \sum_{t'=1}^{t-1} (p \cdot \ell_t^{A_1} + (1-p) \cdot \ell_t^{A_2} + \frac{1}{\eta} R(p))$
2: Choose a number $p' \in [0,1]$ uniformly at random
3: **if** $p' \leq \hat{p}_t$ **then**
4: **return** A_1
5: **else**
6: **return** A_2
7: **end if**

For the expected regret that Follow-the-Regularized-Leader, theorem 5.1.1 holds. The proof is quite similar with the Follow-the-Leader's proof of worst-case regret (without the regularizer), using the Be-the-Regularized-Leader algorithm (for reasons of brevity, we find it pointless to give a detailed description of the algorithm in pseudocode form - in fact it is quite obvious). Just a brief mention on the proof. One difference between BTL and BTRL is that the second one experiences some small regret. The reason it "displays" regret is obvious, because of the regularizer, but why small though? This is also easy to see, since the regularization term is "loss-independent" and by extension "time step-independent", so we can view it as a fake loss in the beginning.

Theorem 5.1.1. *The expected regret of FTRL with a regularizer $R(p)$, satisfying $R''(p) \geq 1$ and a parameter η, is upper bounded:*

$$\textit{Expected-Regret}(T) \leq \frac{2\max_{p \in [0,1]} |R(p)|}{\eta} + \eta \cdot T$$

Based on the above theorem, it is easy to see that choosing the parameter η appropriately, leads - finally! - to sub-linear regret. More precisely, the following corollary holds:

Corollary 5.1.1. *Setting $\eta = \sqrt{\frac{\max_{p \in [0,1]} |R(p)|}{T}}$, the expected regret of FTRL is :*

$$\textit{Expected-Regret}(T) \leq 2 \cdot \sqrt{2 \max_{p \in [0,1]} |R(p)| \cdot T}$$

Alright, so adding as additional terms strictly convex functions to our initial cumulative function (aka "regularizing" $H_t()$) helped us achieve stability and indirectly no-regret. Now, let's take a closer look to a specific regularizer that does the job and displays some really nice properties, the negative entropy function :

$$R(p) = \sum_{i=1} p_i \log p_i \tag{5.1}$$

In order to fully understand the behavior of the negative entropy regularizer, let's examine the following theorem first. Besides stating the theorem, we also give its proof, since it will help us to draw the right conclusions in the "right depth" afterwards :

Theorem 5.1.2. *Let p, q be distributions in action set $A = \{A_1, A_2, \dots, A_n\}$, $H(p) = \sum_{i=1} p_i \log \frac{1}{p_i}$ the entropy of p and $D(p||q) = \sum_{i=1} p_i \log \frac{p_i}{q_i}$ the Kullback-Leibler divergence between p and q. Then the following hold :*

1. *For all $p : 0 \leq H(p) \leq \log n$*
2. *For all p and $q : D(p||q) \geq 0$*

Proof. 1. Given log()'s concavity, we have that :

$$H(p) = \sum_{i=1} p_i \log(p_i) = \mathrm{E}[\log p]$$

$$\overset{\text{Jensen's ineq.}}{\leq} \log \mathrm{E}[p] \overset{\sum_{i=1} p_i = 1}{\leq} \log \frac{1}{n} = \log n \implies$$

$$\implies H(p) \leq \log n \implies H(p) \leq \log n$$

Also :

$$H(p) = \sum_{i=1} p_i \log(p_i) = \mathrm{E}[\log p] \overset{\forall i \in [n]: p_i \leq 1}{\leq} 0 \implies H(p) \geq 0$$

2.

$$D(p||q) = \sum_{i=1} p_i \log \frac{p_i}{q_i} = \sum_{i=1} p_i \log \frac{q_i}{p_i} \tag{5.2}$$

Applying Jensen's inequality to (concave) function $log()$ with argument $g(x) = \frac{q(x)}{p(x)}$, it holds that :

$$\mathrm{E}[\log g(x)] \leq \log \mathrm{E}[g(x)] \overset{g(x) = \frac{q(x)}{p(x)}}{\implies} \mathrm{E}[\log \frac{q(x)}{p(x)}] \leq \log \mathrm{E}[\frac{q(x)}{p(x)}] \tag{5.3}$$

Combining 5.2, 5.3 we end up with the following :

$$D(p||q) \geq \log \sum_{i=1} p_i \frac{q_i}{p_i} \overset{\sum q_i \leq 1 \implies \log q_i \leq 0}{\implies} D(p||q) \geq 0$$

Did that theorem made it clear on what does the entropy regularizer brings to the table? Actually it will, if we think about the types of allocations that achieve the 2 extremes of the entropy values. More precisely, in order to have 0 entropy, all probabilities except one have to be equal to 0, and the - only - non-zero probability has to be equal to 1. Hence, allocations with 0 entropy are the highly concatenated ones. On the other hand, to achieve maximal entropy we have to saturate Jensen's inequality. Since $log()$ function is never linear, the only way to do so, is if and only if $p_1 = p_2 = \ldots = p_n$. So, only the uniform distribution can maximize the entropy. Now, it has to be clear what does the entropy regularizer brings to the table. Let's quote it in a more complete and concrete way, just to remember. ***The negative entropy regularizer penalizes distributions that are too highly concentrated (and favors the less concentrated/"more uniform" ones).*** Hence, the FTRL algorithm with the negative entropy regularizer tries not only to choose the best action in hindsight, but also to play with a distribution over actions that is not highly concatenated. So, it aligns with the "Randomization for Robustness!" principle. Also, the negative entropy regularizer shares a common characteristic with every strongly convex regularizer. It leads to stability, meaning it prevents the FTRL algorithm from overfitting to the history, meaning always jumping around to the best action. In fact, the negative entropy version of the FTRL algorithm makes only small modifications towards the action that performs best in hindsight, thus avoiding big cumulative losses from adversarial loss sequences.
Since in this diploma thesis we are interested in the mechanism design perspective (ok fine, just allocations without the payment part, but you get the point), it would be useful to make some more targeted observations when the objective function is the social welfare. Let's take a closer look :
Just to have something more concrete, let's imagine that we are in the single-item case. The objective modification is the following :

$$\sum_{i=1} x_i \quad v_i \; ! \; \sum_{i=1} x_i \quad v_i + \eta \quad (\quad \sum_{i=1} x_i \quad \log x_i)$$

The new objective function contains 2 sums. The first one expresses the "Highest Bidder Wins" part and forces all the mass to concentrate on the highest bidder. The second one expresses the "Allocate Proportionally" part and forces all the mass to be allocated equally to every bidder. Thus, tuning η generates trade-offs between fairness/stability and efficiency. One more comment that is really worth mentioning. One of the most - if not the most - important characteristics of this technique is its simplicity. More precisely, the negative entropy regularizer behaves like "ignoring" the existence of stability constraints and achieving these stability-efficiency trade-offs like the problem is unconstrained. The importance of this feature becomes more obvious in multi-item settings, where the impact factor of each stability constraint on the social welfare function is - in most cases - impossible to understand. Just to highlight this point a little bit more, think of the single-item case and more specifically the linear program LP1(Single-Item) in 4.1. In this example, the whole fairness/stability thing, which depends on n values/parameters (i.e. $g(v_{i\;1}, v_i)$) now becomes a single-parameter setting, depending only on η.

Now that we have established a solid theoretical background (or at least I hope so) on the whole negative entropy regularizer thing, let's jump into one of the most fundamental applications of this technique, the Multiplicative Weights Update Algorithm.

5.1.2 Killer App: Tell me about Multiplicative Weights Update Algorithm without telling me about Multiplicative Weights Update Algorithm

In order to be more concrete, let's rewrite the explicit formula that FTRL minimizes at each time step t, for the case where the action set is $A = \{A_1, A_2\}$ (along with some calculations) :

$$\hat{p}_t = \arg\min_{p\in[0,1]} H_{t-1}(p) + \frac{1}{\eta}(p\log(p) + (1-p)\log(1-p)) =$$

$$\stackrel{\text{def.}}{=} \arg\min_{p\in[0,1]} \sum_{t'=1}^{t-1} (p(\ell_{t'}^{A_1} - \ell_{t'}^{A_2}) + \ell_{t'}^{A_2}) + \frac{1}{\eta}(p\log(p) + (1-p)\log(1-p)) \Rightarrow$$

$$\stackrel{\ell_{t'}^{A_2}\text{ ind. of } p}{\Longrightarrow} \hat{p}_t = \arg\min_{p\in[0,1]} p \sum_{t'=1}^{t-1} (\ell_{t'}^{A_1} - \ell_{t'}^{A_2}) + \frac{1}{\eta}(p\log(p) + (1-p)\log(1-p)) \Rightarrow$$

$$\stackrel{\text{first order conditions}}{\Longrightarrow} \hat{p}_t = \frac{e^{-\sum_{t'=1}^{t-1} \ell_{t'}^{A_1}}}{e^{-\sum_{t'=1}^{t-1} \ell_{t'}^{A_1}} + e^{-\sum_{t'=1}^{t-1} \ell_{t'}^{A_2}}} \tag{5.4}$$

We can generalize 5.4 in the case where the action set consists of n actions, i.e. $A = \{A_1, A_2, \ldots, A_n\}$(for reasons of brevity and because it lies beyond the purposes of the current work, we omit any further explanation and formalism):

$$\hat{p}_t^{A_i} = \frac{e^{-\sum_{t'=1}^{t-1} \ell_{t'}^{A_i}}}{e^{-\sum_{t'=1}^{t-1} \ell_{t'}^{A_1}} + \ldots + e^{-\sum_{t'=1}^{t-1} \ell_{t'}^{A_n}}} \tag{5.5}$$

Wait! This looks familiar, right?
Yes, just take a moment to check again Algorithm 6 (i.e. the Exponential Mechanism that helps to select a "good" price for the item in the Digital Goods Auction). They are the same thing!
Moreover, there is a nice interpretation of equation 5.5, that yields to a very well known online algorithm (alternative to the Follow-the-Regularized-Leader that we just saw), called Multiplicative Weight Updates. Intuitively, this algorithm proceeds as follows: It keeps a weight $w_t^{A_i}$ for each action $A_i \in A$. At the beginning of the algorithm, before time step 1, we choose an action uniformly at random (aka $w_0^{A_i} = \frac{1}{n}, \forall i \in [n]$). Afterwards, at each time step we learn the current loss vector ℓ^t and update the weight of each action as follows:

$$w_{t+1}^{A_i} = w_{t+1}^{A_i} \cdot e^{-\ell_t^{A_i}}$$

At each iteration, we play each action with probability proportional to its weight. Therefore, the main idea is that gradually we penalize actions that incur higher losses by dropping their weights faster. For a more formal description of this algorithm and because of its importance, we provide Multiplicative Weights Update in pseudocode below.

Algorithm 17 Multiplicative Weights Update Algorithm

Input: action set $A = \{A_1, A_2, \ldots, A_n\}$, "exploration-exploitation" parameter $\eta \in (0, 1)$.

Initialize the weight of each action A_i as follows : $w_i^{(1)} \leftarrow 1$

for each time step $t = 1, 2, \ldots, T$ **do**

  Choose actions proportional to the weights $w_i^{(t)}$, i.e. use the distribution :

$$p^{(t)} = \left(\frac{w_1^{(t)}}{\sum_{i=1} w_i^t}, \frac{w_2^{(t)}}{\sum_{i=1} w_i^t}, \ldots, \frac{w_n^{(t)}}{\sum_{i=1} w_i^t}\right)$$

  Observe the loss vector ℓ_t of each action

  Update the weight of each action A_i as follows : $w_i^{(t+1)} \leftarrow w_i^{(t)} \cdot e^{-\eta \ell_t^{A_i}}$

end for

5.1.3 Our results

5.1.3.1 (Experimental) Comparison between IPA Mechanism, Proportional Mechanism, Entropy Regularized Objective Function Maximization Mechanism & LP Optimal Solver Mechanism for the Single-Item Case

Explaining our initial intuition & goals : Our first approach was to search for experimental indications that the Inverse Proportional Allocation Mechanism was produced by the "Linear Programming" Principle. Hence, we (strongly) believed that IPA algorithm was equivalent to KCST algorithm. Also, we would like to juxtapose graphically the "steep scales nature" of the IPA + KCST and the "smooth nature" of the Entropy Regularizer. Finally, we would like to confirm the "uniform nature" of the PA Mechanism.
Explaining the setup : The setup is quite simple. We fixed an arbitrary number of agents (100 in the plots displayed below). Since we are in the Single-Item case, each agent i has a single valuation v_i. We decided to generate valuations from the Single-Gaussian distribution (as well as the Uniform distribution) and we tune the variance, in order to observe how each allocation rule's behavior changes when agents' purchasing power goes from "all equal" to "rich vs poor guys". The agents are indexed in decreasing order with respect to their valuations, i.e. $v_1 \geq v_2 \geq \ldots \geq v_{100}$. In addition, in the Figures below, $x-$axis consists of the agents' labels (the leftmost agent is the highest valuation one) and $y-$axis consists of the corresponding fractions of the item, allocated to each agent depending on the allocation rule that was contextually running. Just a quick reminder and some clarifications, before jumping into the pictorial results :

- **"Keep Consecutive Scales Tight" Method :** A more analytical explanation about this method can be found in the corresponding subsection of the "The "Linear Programming" Principle" section (4.2.1.1). In a high-level, we chose as $g()$ function the following: $g(v_i, v_{i+1}) = 1 - (\frac{v_i}{v_{i+1}})^{-2\ell}, \forall i \in \{1, \ldots, n-1\}$. Because of the optimality of KCST, we expected to produce the most concentrated, yet efficient allocations. To be more precise, we expect this method to produce very "steeps scales" (since it satisfies tightly the stability constraints, trying to keep the differences in the allocations - at least on the highest bidders - as large as possible), thus achieving high efficiency.

- **Inverse Proportional Allocation Method :** As we explained earlier, the main functionality of this algorithm is to identify the number of non zero-allocation bidders. Assuming that the bidder's valuations is ordered as follows : $v_1 \leq v_2 \leq \ldots \leq v_n$, the mechanism initially sets all the agents as active and starts "trying" to allocate some positive fraction of the item from the lowest valuation bidder and on (sets index $s = 0$). On the i-th bidder, the quantity that the mechanism tries to allocate is the following :

$$a_i = \frac{(n - 1 - s) \cdot v_i^{\ell}}{\sum_{j=s}^{n} v_j^{\ell}}$$

If $a_i \leq 0$, the mechanism sets this bidder to the set of zero-allocation agents and proceeds to the next higher valuation agent, adjusting the index of still active agents appropriately (i.e. $s \to s+1$).

- **Proportional Allocation Method :** Probably the easier to explain method from the 4, since it is the easier to understand. We simply give each agent the following fraction of the item :

$$x_i = \frac{v_i^{\ell}}{\sum_{j=1}^{n} v_j^{\ell}}$$

- **Entropy Regularization Method :** As we explain in (way) more details in the entropy regularization-related subsections of the "The "Slightly or Completely Modify the Objective Function" Principle" section, the main purpose of the entropy regularizer is to generate trade-offs between "Highest Bidder Wins" allocations (the optimal unfair) and "Allocate Uniformly" allocations (the most fair). More technically, we modify the objective function (aka the "classical" social welfare function) and try to solve the following optimization problem :

$$\max\left\{\sum_{i=1}^{n} x_i \cdot v_i + \frac{1}{\eta} \cdot \left(-\sum_{i=1}^{n} x_i \cdot \log x_i\right)\right\}$$

Observe that the entropy part is multiplied by $\frac{1}{\eta}$ and not by simply η. The main reason for that choice is that we would like when increasing the ℓ or η parameter, to observe the same behaviors (meaning increasing ℓ or η the allocations, become more concatenated). The above maximization problem has a solution with closed form, as proved in Nesterov's paper [34]. After the small modification that we mentioned above (i.e. replacing $\eta \to \frac{1}{\eta}$), we allocate to each agent the following fraction of the item :

$$x_i = \frac{e^{\eta \cdot v(i)}}{\sum_{\ell=1}^{n} e^{\eta \cdot v(\ell)}}$$

The results :

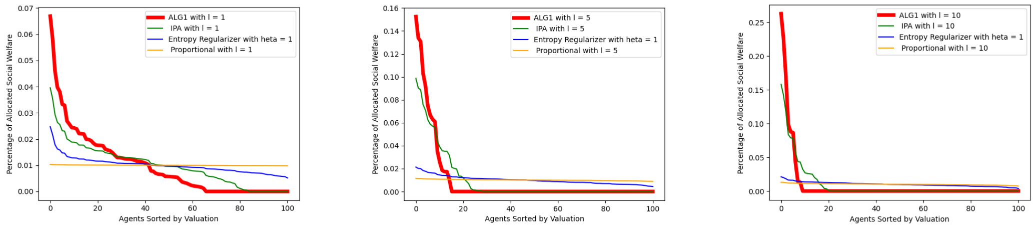


Figure 5.1: IPA vs KCST vs PA vs Entropy Regularizer for Single Gaussian($\mu = 30, \sigma = 0.3$)

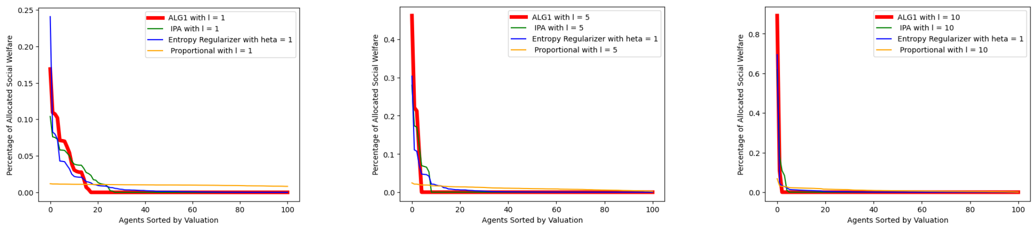


Figure 5.2: IPA vs KCST vs PA vs Entropy Regularizer for Single Gaussian($\mu = 30, \sigma = 3$)

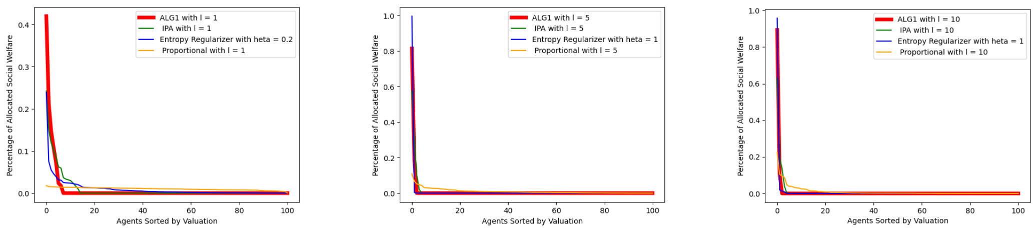


Figure 5.3: IPA vs KCST vs PA vs Entropy Regularizer for Single Gaussian($\mu = 30, \sigma = 9$)

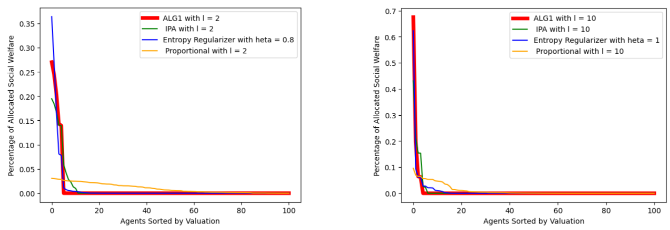


Figure 5.4: IPA vs KCST vs PA vs Entropy Regularizer for Uniform($\mu = 30$)

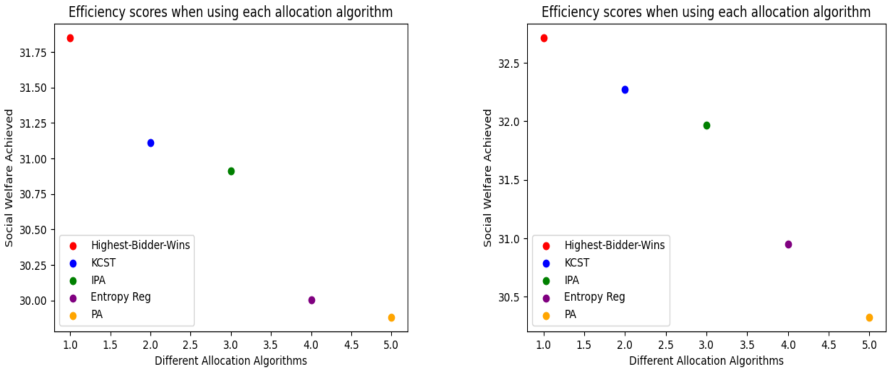


Figure 5.5: Efficiency of IPA vs KCST vs PA vs Entropy Regularizer vs HBW for Gaussian($\mu = 30, \sigma = 0.9$)

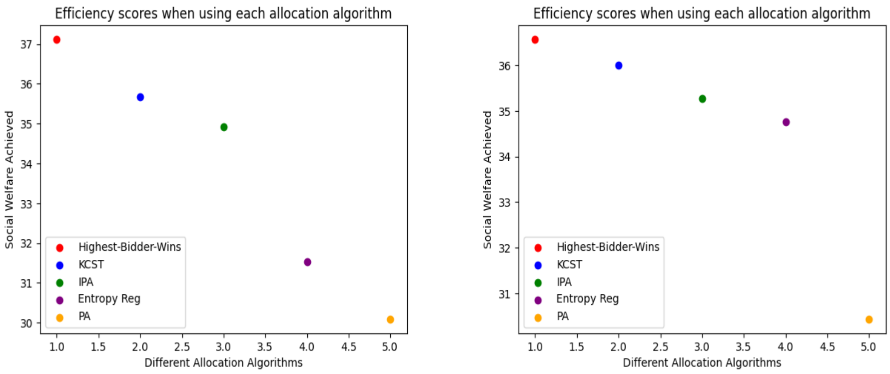


Figure 5.6: Efficiency of IPA vs KCST vs PA vs Entropy Regularizer vs HBW for Gaussian($\mu = 30, \sigma = 3$)

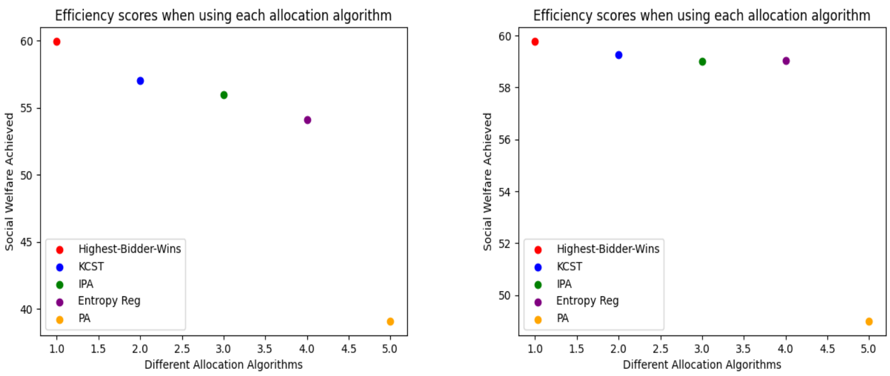


Figure 5.7: Efficiency of IPA vs KCST vs PA vs Entropy Regularizer vs HBW for Uniform($\mu = 30$)

Explaining the experimental results (or at least trying to) : Alright, some of the most important conclusions that can be drawn from Figures 5.1,5.2,5.3 and 5.4 are the following :

1. We have confirmed experimentally the nature/form of allocations for all the 4 different allocation mechanisms. For reasons of completeness, just a brief repetition : Both KCST and IPA rules produce allocations of "steep scales nature", Entropy regularizer produces allocations of "smooth nature" and PA rule of "uniform nature".
2. We have evidence of the difference between the IPA and the KCST mechanism. Moreover, since KCST produces systematically "steeper scales", it has to return higher social welfare compared to IPA's output.
3. The less efficient (always in terms of social welfare) but also the "fairest" (from the "equalizing sense") from the 4 mechanisms is the PA rule.
4. As ℓ grows, the allocations from KCST & IPA (from PA too, but - of course - in a lesser extent) become more concentrated, meaning less "fair" and more efficient (i.e. tend to the "Highest Bidder Wins"). Similar behavior as η grows (especially compared with the PA mechanism) is observed with the Entropy Regularizer method.
5. As the variance increases, all the allocation mechanisms produce allocations with higher concentrations (obviously with different rate of concentration).

5.1.3.2 Synopsis

In the table below, we present for our 4 allocation algorithms of concern, which of 3 important properties they display. The desired properties are the following :

- **Monotonicity:** In the sense of "making Myerson's Lemma work", in order to be able to design the appropriate payment rule and, in a second level, the corresponding truthful mechanism. More precisely, we would like - ideally - when increasing the valuation of an agent for the (single) item (keeping all the other agents' valuations fixed), the allocation that he receives from the algorithm to be no less that the one under the initial value vector.
- **Scale-Freeness :** This property is of interest when dealing with non-monetary settings (as in our case here). The main reason is that, when simply allocating fractions of an item for free, we consider value vectors of the form v and $\alpha \quad \mathsf{v}$, with $\alpha > 0$, as equivalent since the absolute values of the agents' valuations don't have any particular "natural meaning". They are simply a way to express the ordering between agents with criterion how much they would like to "purchase" the item (given their market force and -probably- other contributing factors as well), which is the only agents' feature that we are interested in in our diploma thesis. Obviously, to actually achieve scale-freeness (through the mechanisms that display this ability), we have to choose a scale-free distance function.
- **Value-Stability :** Since our research efforts are "aligned" in the direction of achieving some kind of value-stability/"fairness" in a welfare maximizing way, value-stability is - obviously - one of the most (if not the most) important desiderata. In our thesis, as we have already mentioned several times, we examine this value stability concept through 2 different kind of definitions. The self-referential global stability definition 3.1.1 (in the sense of referring to allocation constraints on the same agent under 2 different - but close to each other - instances/value vectors) from the line of work that initially triggered our research interest. And our own, non self-referential local stability definition (in the sense

of referring to value constraints on 2 different agents under the same instance/value vector), aka ”Scales” (for a more analytical explanation, we refer to the corresponding subsection 4.1 of the ”Linear Programming” Principle section).

		Allocation Algorithms			
		KCST	Entropy Reg.	IPA	PA
Desiderata	Monotonicity		X	X	X
	Scale-Freeness	X		X	X
	Value-Stability	”Scales”	”Scales”	Definition 3.1.1	Both

Table 5.1: Comparison between our 4 allocation algorithms of main interest

5.2 The Replacement with the Nash Social Welfare Objective Function Method

5.2.1 Multi-Objective Optimization vs Nash Social Welfare Optimization through an Example

5.2.1.1 Basic Definitions

Our main goal, here, is to better understand the Nash Social Welfare Objective. Most of our theoretical background and intuition on this topic was constructed with the aim of the work of Charkhgard et al. in [4] and Nakamura and Kaneko in [32] (which is by the way the most ”paper for economists” paper that i have studied so far!).

First things first : motivation! Why should we choose the Nash Social Welfare Objective over other objectives, hence why should we even consider replace our initial objective function ”equipped” with the stability constraints that we discussed earlier with this one in the first place? Before trying to answer that question, let's finish with the boring stuff first.

Initially, we will present some basic notions/definitions that will help us walk through the entire section. Let's begin by defining formally our protagonist, the Nash Social Welfare Objective/Program.

Nash Social Welfare Program (NSWP)

$$\begin{aligned} \text{maximize} \quad & \prod_{i \in S^+} (f_i(x) - d_i)^{w_i} \prod_{i \in S^-} (d_i - f_i(x))^{w_i} \\ \text{subject to} \quad & x \in X, \\ & f_i(x) \geq d_i, \quad \forall i \in S^+, \\ & f_i(x) \leq d_i, \quad \forall i \in S^- \end{aligned}$$

Now, let's define the 2 main Programs, that we will compare the Nash Social Welfare Program with, in order to highlight its usefulness. These 2 are the Single-Objective Optimization and Multi-Objective Optimization Programs.

Single-Objective Optimization Program
(SWP)

$$\max \left\{ \sum_{i=1}^{p'} w_i f_i(x) - \sum_{i=p'+1}^{p} w_i f_i(x) : x \in X \right\},$$

where $p \geq p' \geq 0$ and $X \subseteq \mathbb{R}^n$ represents the set of feasible solutions and is assumed to be bounded. Moreover, $\boldsymbol{f(x)} := (f_1(x), \ldots, f_p(x))$ is a vector of arbitrary functions and $\boldsymbol{w} := (w_1, \ldots, w_p)$ is a vector of weights with $w_i > 0$ for all $i = 1, \ldots, p$.

For the rest of the section, we will refer to the above problem as SWP (because the above objective function is called (weighted) social welfare function).

Multi-Objective Optimization Program
(MOOP)

$$\begin{aligned} &\max \{f_i(x) : i \in S^+\} \\ &\min \{f_i(x) : i \in S^-\} \\ &\text{s.t. } x \in X, \end{aligned}$$

where S^+ and S^- are the (index) sets of objective functions that need to be maximized and minimized, respectively. For simplicity, we assume that $S^+ := \{1, \ldots, p'\}$ and $S^- := \{p'+1, \ldots, p\}$ which implies that $S^+ \cup S^- := \{1, \ldots, p\}$.

For the rest of the section, we will refer to the above problem as MOOP. In order to better understand

Definition 5.2.1. (Pareto Improvement from its allocation-aspect). Given an initial allocation, a Pareto improvement is a new allocation where some agents will gain more value, while no agents will lose value.

Definition 5.2.2. (Pareto Optimality from its allocation-aspect). An allocation is called Pareto-optimal if no change in the allocation can lead to a Pareto improvement.

5.2.1.2 Weaknesses of Single-Objective Optimization

First, it is easy to observe that SWP is a solution approach for computing a Pareto-optimal point for the MOOP. Specifically, assume the following transformation (from MOOP to SWP):

$$\begin{array}{l} \max \{f_i(x) : i \in S^+\} \\ \min \{f_i(x) : i \in S^-\} \\ \text{s.t. } x \in X, \end{array} \implies \begin{array}{l} \max \{f_i(x) : i \in S^+\} \\ \max \{-f_i(x) : i \in S^-\} \\ \text{s.t. } x \in X, \end{array} \implies \max \left\{ \sum_{i \in S^+} w_i f_i(x) - \sum_{i \in S^-} w_i f_i(x) : x \in X \right\}$$

Setting $S^+ = \{1, \ldots, p'\}$ and $S^- = \{p'+1, \ldots, p\}$, we transform the weighted sum optimization problem into precisely the MOOP. So, trying to solve the SWP is basically equivalent to having p number of criteria/functions $f_1, \ldots, f_p$ and trying to find a desirable Pareto-optimal solution for MOOP by setting the weight w_i of $f_i(x)$, $\forall\, 1, \ldots, p$, where the weights indicate the degree of importance for each criterion.
Ok, now that he have established the relationship between a SWP and a MOOP, we are ready to jump into the main weaknesses of the Single-Objective Optimization, with the help of a very

simple example. But first, we have to define one more notion - the unsupported Pareto-optimal points - since it is highly correlated with those weaknesses.

Definition 5.2.3. (Unsupported Pareto-optimal points) For non-convex multi-objective optimization problems, there can exist - possibly infinite - Pareto-optimal points that cannot be obtained by optimizing a (positive) weighted summation of all objective functions over the feasible set. Such points are called unsupported Pareto-optimal points.

Alright, so now let's see the example: Say we would like to maximize 2 objective functions, $f_1(x), f_2(x)$. In addition, say that we end up with a Pareto-optimal frontier consisting of the points : $f(2, 5.5), (2.25, 4.5), (3.25, 3.5), (4.25, 2.5), (5.5, 2)g$ as illustrated in Figure 5.8. Based on this example, we can easily 2 detect to major weaknesses of SWP, correlated with the fairness of the its optimal solution. More precisely :

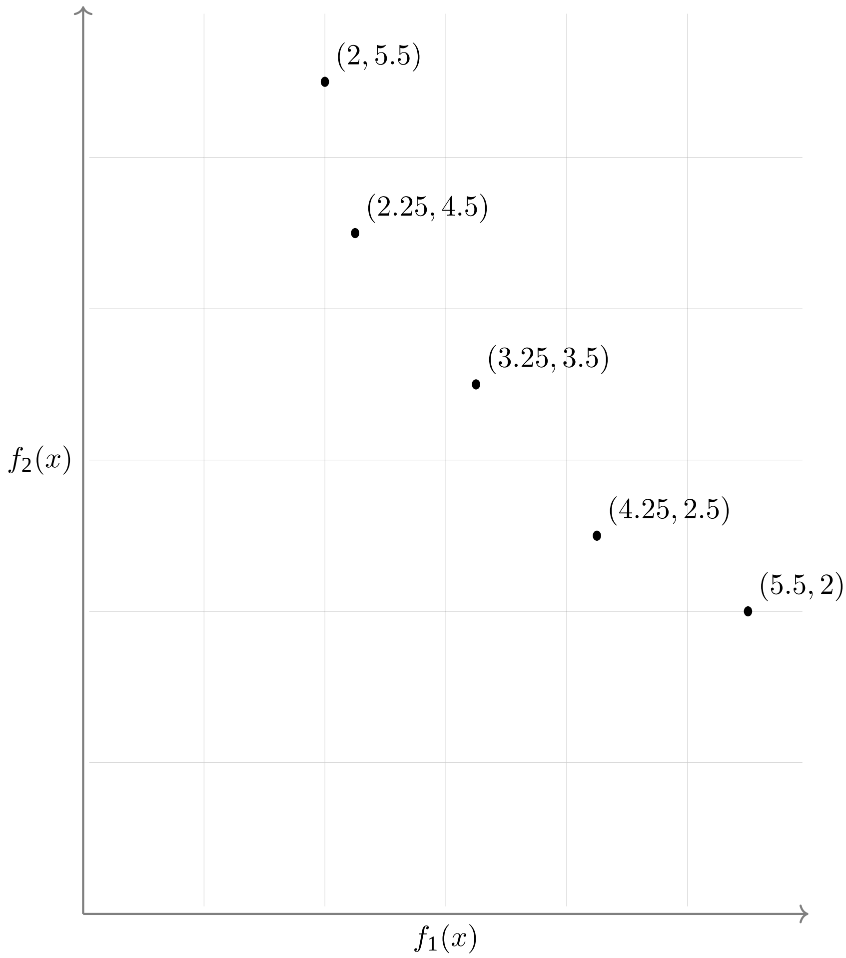


Figure 5.8: An illustration of Pareto-optimal frontier with 2 objective functions in the criterion space

- **Weakness 1 :** A weakness of the SWP is that it completely ignores the existence of unsupported Pareto-optimal points that can possibly better balance different objectives.

To see this, let's take a look at our example in Figure 5.8. If we use the SWP, then we will find either the point $(2, 5.5)$ or the point $(5.5, 2)$, thus completely ignoring each of the middle points (that balance better between the conflicting objectives - it can quite easily be proven, by observing that for each Pareto middle point (x_1, x_2) of the example, it holds that $7.5 \quad x_1 \quad x_2 > 0$).

- **Weakness 2 :** Another weakness of the SWP is that it does not necessarily ensure the fairness. In fact, this observation/weakness is on the same page with the first one. To see why, let's consider again Figure 5.8. Suppose that the 2 criteria/objectives are equally important (let $w_1 = w_2 = 1$). In this case, the SWP returns one of the 2 endpoints with equal propability $\frac{1}{2}$, and nothing else. However, it is quite obvious that the "fairest" solution is the middle point $(3.25, 3.5)$, that makes both players - almost - equally happy.

5.2.1.3 Weaknesses of Multi-Objective Optimization

Since every MOOP can be transformed to SWP - as shown in the previous subsection -, the fairness-related weaknesses of the single-objective optimization are present in the multi-objective optimization too. Besides, a lot of MOOP are too complex not only from a computational standpoint, but even from a desiderata one (meaning we may not even know our preference over the set of feasible solutions). Ok, to put it epigrammatically in a more concrete way :

- **Weakness 1 :** In some situations, computing even a single Pareto-optimal solution is computationally infeasible/expensive. So, selecting a desirable Pareto-optimal solution using the existing approaches can be infeasible.

- **Weakness 2 :** In some situations, either there is no decision maker or the decision makers do not know how to select a desirable solution. So, selecting a desirable Pareto-optimal solution using the existing approaches can be infeasible.

5.2.1.4 Good Properties of Nash Social Welfare Objective Function : Multiplying is better than Summing!

Let's revisit the initial example from the perspective of optimizing the Nash Social Welfare Objective. First, let's fix the reference point. Suppose that it is $d = (2, 2)$. Also, suppose that each of the 2 players has the same negotiating power (let's say $w_1 = w_2 = 1$). In that case, the NSWP that we defined earlier tries to maximize the area of the rectangle with endpoints on one of its diagonals d and the point under consideration. The points that we examine belong to the feasible space. As we observe in Figure 5.9, the solution that the NSWP returns, is the point $(3.25, 3.5)$. More specifically, we observe the following :

- The NSWP returned a Pareto-optimal point.
- The NSWP is a single-objective optimization problem (obviously, the product is the single objective).
- The NSWP returned an *unsupported* Pareto-optimal point.
- The NSWP returned the only fair point of the example.

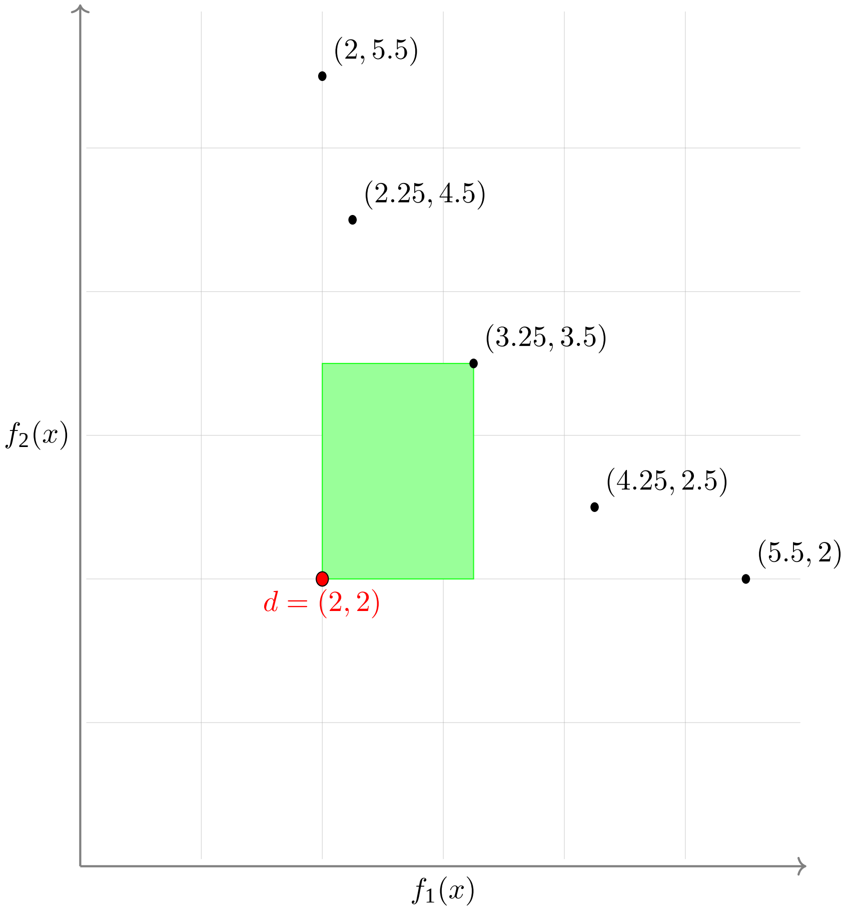


Figure 5.9: Selection of the Pareto-optimal point by the NSWP

The charm of Nash Social Welfare in a Nutshell

1. **Pareto Optimality:** As we explained earlier, this property implies that for a NSWO maximizing solution, there is no feasible state/solution where at least one individual will gain more value while no individual will lose value.
2. **Independence of Irrelevant Alternatives with Neutral Property:** This means that the social preference between two alternatives depends only on individuals' preferences between these alternatives, regardless of individuals' preferences relating to other alternatives.
3. **Anonymity:** The preferences of individuals and the value they receive should be considered without any regard to who those individuals are. More precisely, if we exchange the roles of two individuals - i.e. if two individuals exchange both their utility functions and the values they receive from the mechanism - the total social welfare will remain unchanged. This property assures that all individuals are treated the same, regardless of their identities.
4. **Continuity:** It is a self-referential notion. Imagine that an individual has a set of preferences for some outcomes that happen with certainty. Imagine that there are 2 alternatives. The first one is to get one of two outcomes with some probability distribution (for this case of two outcomes, we have a single parameter, say p) and the other alternative is to get a third outcome with certainty. Continuity implies that there is a value of p where the individual is indifferent between these 2 alternatives (necessary condition for the existence of Mixed Nash Equilibrium).

Just to summarize the information inside the box in a form of theorem, the following possibility theorem holds :

Theorem 5.2.1. *The Nash Social Welfare Function satisfies all 4 conditions: Pareto Optimality, independence of Irrelevant Alternatives with Neutral Property, Anonymity and Continuity.*

Actually, things are even better (for the Nash Social Welfare!). The following uniqueness theorem holds :

Theorem 5.2.2. *The Nash Social Welfare Function is the only social welfare that satisfies all 4 conditions: Pareto Optimality, independence of Irrelevant Alternatives with Neutral Property, Anonymity and Continuity.*

Another worth-mentioning property of Nash Social Welfare is its multiplicative scale-freeness. To be more precise,

- The NSWP is global-power-scale-free, i.e.,

$$\max_{y \in Y} \prod_{i \in S^+ \cup S} y_i^{w_i} \qquad \max_{y \in Y} \prod_{i \in S^+ \cup S} y_i \quad ^{w_i}$$

- The NSWP is local-benefit-scale-free, i.e.,

$$\max_{y \in Y} \prod_{i \in S^+ \cup S} y_i^{w_i} \qquad \max_{y \in Y} \prod_{i \in S^+ \cup S} (\alpha_i \quad y_i)^{w_i}$$

As Cole and Gkatzelis mention ([7]), the scale-freeness property means that choosing the desired allocation does not require interpersonal compatibility of the individual's preferences. Maybe a more straightforward explanation? Yes, actually this simply means that the agents should just report their relative valuations, i.e. how much more they like an item compared to another. Hence, maximizing NSWO using v_{ij} values (valuation of agent i for item j), is equivalent to using $a_i v_{ij}$ as values instead, where $a_i > 0$ is some constant for agent i. This property is particularly useful in settings where the agents are not paying for the items that they are allocated, in which case the scale in which their valuations are expressed may not have any real meaning.

5.2.2 Our results

5.2.2.1 Proof that "Proportional Allocation is the best Allocation" for Concave Program allocating probabilities with Nash Social Welfare Objective Function

Theorem 5.2.3. *The following concave program :*

AN ugrf L_qf Qmag_j Ucjd_pc M`hcargtc

$$
\begin{aligned}
\text{maximize} \quad & \sum_{i \in N} v_i \log x_i \\
\text{subject to} \quad & \sum_{i \in N} x_i = 1, \\
& x_i \geq 0, \forall i \in N
\end{aligned}
$$

has as optimal solution the proportional allocation.

Proof. It is easy to see that the objective function $f(x_1, \ldots, x_n) = \sum_{i \in N} v_i \log x_i$ is concave. Indeed, the Hessian matrix of f is $H(x_1, \ldots, x_n) = \begin{bmatrix} \frac{v_1}{x_1^2} & 0 & \cdots & 0 \\ 0 & \frac{v_2}{x_2^2} & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{v_n}{x_n^2} \end{bmatrix}$, so it is diagonal. Thus, it's eigenvalues are of the form : $-\frac{v_i}{x_i^2} \leq 0, \forall i \in N$, which means that H is negative semi-definite. We define the Lagrangian for CP, as :

$$L(x_1, \ldots, x_n, \lambda) = \sum_{i \in N} v_i \log x_i - \lambda \sum_{i \in N} x_i$$

$$\nabla L(x1, \ldots, x_n, \lambda) = 0 \implies$$

$$\nabla\Big(\sum_{i \in N} v_i \log x_i\Big) - \lambda \nabla\Big(\sum_{i \in N} x_i\Big) = 0 \implies$$

$$\left(\begin{bmatrix} \frac{v_1}{x_1} \\ \vdots \\ \frac{v_n}{x_n} \end{bmatrix} - \lambda \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}\right) = 0 \implies$$

$$x_i = \frac{v_i}{\lambda}, \quad \forall i \in N \xRightarrow[\lambda = \sum_{i \in N} v_i]{(\sum_{i \in N} x_i = 1 \implies)}$$

$$x_i = \frac{v_i}{\sum_{i \in N} v_i}, \quad \forall i \in N$$

5.3 Epimythium

What's the Epimythium?

"Slightly or Completely Modify the Objective Function" Principle Version

- The negative entropy regularizer penalizes distributions that are too highly concentrated (and favors the less concentrated/more uniform ones).
- The famous MWU Algorithm is simply the solution of the FTRL algorithm "equipped" with the negative entropy regularizer.
- Catchphrase for the MWU : Initially choose an action uniformly at random. Keep choosing actions, and if the action is proven to be "bad", "punish" her by decreasing exponentially the probability of choosing her in future rounds.
- One "On achieving no-regret" principle is the following: We need randomization, but not "completely random randomization". More specifically, past performance of actions should guide us on what action to choose in the current time step. This means that the probability of choosing some action should be decreasing to the cumulative loss.
- Another "On achieving no-regret" principle: The probability of choosing a poorly performance action should decrease in exponential rate.
- Probably the most important property of the NSWO that justify its existence: it returns Pareto optimal solutions, but that are somehow in the "medium" of the Pareto frontier, meaning that they achieve optimality by satisfying in some degree all the criteria/objective functions involved and not by optimizing only some of them.
- The concave program that allocates probabilities with the NSWO is solved optimally by the proportional allocation mechanism.

Bibliography

[1] Muhammad Ali et al. «Discrimination through optimization». In: *Proceedings of the 14th Innovations in Theoretical Computer Science (ITCS)*. 2019, pp. 1–30.

[2] M-F Balcan et al. «Mechanism design via machine learning». In: *FOCS 2005. 46th Annual IEEE Symposium* (2005), pp. 605–614.

[3] Ioannis Caragiannis and Alexandros A. Voudouris. «Welfare guarantees for proportional allocations». In: *Theory of Computing Systems* 59.4 (2016), pp. 581–599.

[4] Hadi Charkhgard et al. «THE MAGIC OF NASH SOCIAL WELFARE IN OPTIMIZATION: DO NOT SUM, JUST MULTIPLY!» In: *Anziam Journal, 64* 2 (2022), pp. 119–134.

[5] Suchi Chawla and Meena Jagadeesan. «Individual fairness in advertising auctions through inverse proportionality». In: *Proceedings of the 14th Innovations in Theoretical Computer Science (ITCS)*. 2021.

[6] George Christodoulou, Alkmini Sgouritsa, and Bo Tang. «On the efficiency of the proportional allocation mechanism for divisible resources». In: *Theory of Computing Systems* 59.4 (2016), pp. 600–618.

[7] Richard Cole and Vasilis Gkatzelis. «Approximating the Nash Social Welfare with Indivisible Items». In: *ACM SIGecom Exchanges* 14.1 (2015), pp. 84–88.

[8] J. R. Correa, A. S. Schulz, and N. E. Stier-Moses. «The price of anarchy of the proportional allocation mechanism revisited». In: *Proceedings of the 9th Conference on Web and Internet Economics (WINE), LNCS 8289*. Springer, 2013, pp. 109–120.

[9] Irit Dinur and Kobbi Nissim. «Revealing information while preserving privacy». In: *Proceedings of the twenty-second ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems*. 2003.

[10] Cynthia Dwork and Christina Ilvento. «Fairness under composition». In: *10th Innovations in Theoretical Computer Science Conference, ITCS 2019*. San Diego, California, USA, Jan. 2019, 33:1–33:20.

[11] Cynthia Dwork and Aaron Roth. «The algorithmic foundations of Differential Privacy». In: *Foundations and Trends® in Theoretical Computer Science* (2013), pp. 211–407.

[12] Cynthia Dwork, Guy N. Rothblum, and Salil Vadhan. «Boosting and differential privacy.» In: *Proceedings of the 51st Annual IEEE Symposium on Foundations of Computer Science,FOCS'10*. 2010, pp. 51–60.

[13] Cynthia Dwork et al. «Calibrating noise to sensitivity in private data analysis». In: *Theory of Cryptography* (2006), pp. 265–284.

[14] Cynthia Dwork et al. «Our data, ourselves: Privacy via distributed noise generation». In: *Advances in Cryptology - EUROCRYPT* (2006), pp. 486–503.

[15] Cynthia Dwork at al. «Fairness through awareness». In: *Proceedings of the 3rd Innovations in Theoretical Computer Science Conference (ITCS)*. 2012.

[16] Amos Fiat and Gerhard J. Woeginger. *Online Algorithms: The State of the Art*. Springer, 1998.

[17] Amos Fiat and Gerhard J. Woeginger. *Online algorithms: The state of the art*. Springer, 1998.

[18] Yoan Freund and Robert E Schapire. «A decision-theoretic generalization of on-line learning and an application to boosting». In: *Journal of Computer and System Sciences, 55* (1) (1997), pp. 119–139.

[19] Nika Haghtalab and Chara Podimata. *Incentive-Compatible and Incentive-Aware Learning*. Tutorial presented at EC'20, Conference on Economics and Computation. 2020.

[20] Lily Hu and Yiling Chen. «Fair classification and social welfare». In: *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency, FAT* '20*. New York, NY, USA: Association for Computing Machinery, 2020, pp. 535–545.

[21] Christina Ilvento, Meena Jagadeesan, and Suchi Chawla. «Multi-category fairness in sponsored search auctions». In: *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*. 2020.

[22] Ramesh Johari and John N. Tsitsiklis. «Efficiency loss in a network resource allocation game». In: *Mathematics of Operations Research* 29.3 (2004), pp. 407–435.

[23] Michael P. Kim et al. «Preference-informed fairness». In: *Proceedings of the 11th Innovations in Theoretical Computer Science, ITCS 2020*. Seattle, Washington, USA, Jan. 2020, to appear.

[24] Anja Lambrecht and Catherine e. Tucker. «Algorithmic bias? an empirical study into apparent gender-based discrimination in the display of STEM career ads». In: *SSRN Electronic Journal* (2016).

[25] Annie Liang, Jay Lu, and Xiaosheng Mu. «Algorithmic design: Fairness versus accuracy». In: *SSRN Electronic Journal* (2021).

[26] Richard J. Lipton and et al. «On Approximately Fair Allocations of Indivisible Goods». In: *Proceedings of the 5th ACM Conference on Electronic Commerce*. [Preprint]. 2004.

[27] Evangelos Markakis. «Approximation Algorithms and Hardness Results for Fair Division with Indivisible Goods». In: *Trends in Computational Social Choice*. Ed. by Ulle Endriss. AI Access, 2017. Chap. 12, pp. 231–247.

[28] Evangelos Markakis and Charis-Angelos Psomas. «On Worst-Case Allocations in the Presence of Indivisible Goods». In: *Lecture Notes in Computer Science*. 2011, pp. 278–289.

[29] Frank McSherry and Kunal Talwar. «Mechanism design via Differential Privacy». In: *Proceedings of the 48th annual IEEE Symposium on Foundations of Computer Science, Guide Proceedings*. 2007.

[30] Michael Mitzenmacher and Eli Upfal. *Probability and Computing: Randomized Algorithms and Probabilistic Analysis*. Cambridge University Press, 2005.

[31] R. Motwani and P. Raghavan. *Randomized Algorithms*. Cambridge University Press, 1995.

[32] Kenjiro Nakamura and Mamoru Kaneko. «The Nash Social Welfare function». In: *Econometrica, 47* (2) (1979), pp. 423–435.

[33] John F Nash et al. «Equilibrium points in n-person games». In: *Proceedings of the national academy of sciences* 36.1 (1950), pp. 48–49.

[34] Yurii Nesterov. «Smooth minimization of non-smooth functions». In: *Mathematical Programming, 103* (1) (2004), pp. 127–152.

[35] T. Nguyen and M. Vojnovic. «Weighted proportional allocation». In: *Proceedings of the 2011 ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Systems*. 2011, pp. 173–184.

[36] Emerson M. S. Niou and Peter C. Ordeshook. «A game-theoretic interpretation of Sun Tzu's,31». In: *Journal of Peace Research* 2 (1994), pp. 161–174.

[37] Aaron Roth. *Diferential Privacy in Game Theory and Mechanism Design*. Lecture Notes. CIS700: University of Pennsylvania, Spring 2014.

[38] Shai Shalev-Shwartz. «Online learning and online convex optimization». In: *Foundations and Trends in Machine Learning, 4* (2) (2012), pp. 107–194.

[39] Shai Shalev-Shwartz and Shai Ben-David. *Understanding machine leanring: From theory to algorithms: From theory to algorithms*. Cambridge University Press, Cambridge, 2022.

[40] Shai Shalev-Shwartz and Yoram Singer. «A primal-dual perspective of online learning algorithms.» In: *Machine Learning, 69* (2-3) (2007), pp. 115–142.

[41] Shai Shalev-Shwartz and Yoram Singer. *Online learning: Theory, algorithms and applications*. 2007.

[42] Vasilis Syrgkanis. *Online Learning*. Lecture Notes. 6.853: Algorithmic Game Theory and Data Science, MIT EECS, 15 February. 2017.

[43] John Von Neumann and Oskar Morgenstern. «Theory of games and economic behavior». In: (1947).

[44] Susan Wei and Marc Niethammer. «The Fairness-Accuracy Pareto Front». In: *Statistical Analysis and Data Mining: The ASA Data Science Journal* (2021).

[45] Crystal Yang and Will Dobbie. «Equal Protection under algorithms: A new statistical and legal framework». In: *Michigan Law Review* (2020).

[46] Muhammad Bilal Zafar et al. «From parity to preference-based notions of fairness in classification». In: *Advances in Neural Information Processing Systems 30: NeurIPS 2017*. Long Beach, CA, USA, Apr. 2017, pp. 229–239.